# Health Technology Assessment

## Spine Fusion and Discography

### Clinical Committee Meeting

November 16, 2007

## Contributors and Participants

<table>
<thead>
<tr>
<th>Name</th>
<th>Role</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Richard N. Wohns, MD</td>
<td>Past Chair AANS/CNS Section on Disorders of the Spine and Peripheral Nerves</td>
<td>Columbia University, New York, NY</td>
</tr>
<tr>
<td>Jens R. Chapman, MD</td>
<td>Chair-Elect AANS/CNS Section on Disorders of the Spine and Peripheral Nerves</td>
<td>University of Wisconsin, Madison, WI</td>
</tr>
<tr>
<td>Ray M. Baker, MD</td>
<td>Chair, AANS/CNS Section on Disorders of the Spine and Peripheral Nerves</td>
<td>University of Utah, Salt Lake City, UT</td>
</tr>
<tr>
<td>Theodore A. Wagner, MD</td>
<td>Chair, AANS/CNS Section on Disorders of the Spine and Peripheral Nerves</td>
<td>University of Washington, Seattle, WA</td>
</tr>
<tr>
<td>Paul C. McCormick, MD</td>
<td>Past Chair AANS/CNS Section on Disorders of the Spine and Peripheral Nerves</td>
<td>Columbia University, New York, NY</td>
</tr>
<tr>
<td>Daniel K. Resnick, MD</td>
<td>Chair, AANS/CNS Section on Disorders of the Spine and Peripheral Nerves</td>
<td>Columbia University, New York, NY</td>
</tr>
<tr>
<td>Andrew T. Dailey, MD</td>
<td>Chair, AANS/CNS Section on Disorders of the Spine and Peripheral Nerves</td>
<td>University of Utah, Salt Lake City, UT</td>
</tr>
<tr>
<td>Joseph T. Alexander, MD</td>
<td>Chair, AANS/CNS Section on Disorders of the Spine and Peripheral Nerves</td>
<td>University of Washington, Seattle, WA</td>
</tr>
</tbody>
</table>

### Washington State Health Care Authority

Health Technology Assessment

Spine Fusion and Discography

Clinical Committee Meeting

November 16, 2007

**Contributors and Participants**

- Richard N. Wohns, MD, AANS/CNS, Tacoma, WA
- Jens R. Chapman, MD, AAOS, University of Washington, Seattle, WA
- Ray M. Baker, MD, NASS, University of Washington, Seattle, WA
- Theodore A. Wagner, MD, AAOS, University of Washington, Seattle, WA
- Paul C. McCormick, MD, Past Chair, AANS/CNS Section on Disorders of the Spine and Peripheral Nerves, Columbia University, New York, NY
- Daniel K. Resnick, MD, Chair-Elect, AANS/CNS Section on Disorders of the Spine and Peripheral Nerves, University of Wisconsin, Madison, WI
- Andrew T. Dailey, MD, AANS/CNS Section on Disorders of the Spine and Peripheral Nerves, University of Utah, Salt Lake City, UT
- Joseph T. Alexander, MD, Chair, AANS/CNS Section on Disorders of the Spine and Peripheral Nerves, Scarborough, ME
This discussion is limited to the performance of lumbar fusion for axial back pain thought to be due to degenerative disc disease (DDD).

The utility of fusion as an adjunct to decompression or as a means of deformity correction in the North American patient population has been firmly established:

- SPORT II, 2007; Ghogowala and Benzel, 2006; Resnick et al, 2005; many others

Efficacy studies for surgery (4 European RCTs - Fritzell, Brox(2), Fairbank)

Study limitations (design, methods, assumptions)-biased to null against surgery

Significant improvement in surgical groups compared to non-operative groups

Failure of any type of nonoperative treatment to achieve MCID in ECRI meta-analysis or adequately powered RCT
Outline

- Outcomes of North American FDA IDE for degenerative disc disease
- Unavailability of structured PT in the United States
- ECRI’s calculation of “true impact of surgery”

4 European RCTs

- Single hypothesis-Is surgery better than nonoperative management.

Assumptions:
1. Low back pain is a homogeneous condition.
2. Each patient equally likely to achieve mean improvement associated with either treatment.
3. The mean outcome score for each treatment provides a consistently accurate estimate of every patient's response to either treatment.
4. Singular conclusion “which treatment is better” applies to all patients
Low back pain

- Heterogeneous- cause, structural basis, natural hx, clinical aspects and patient effects, patient’s treatment preference and response
- Many patients do well with nonoperative care, some don’t.
- Many patients do well with surgery, some don’t.
- Wide variation in magnitude of Rx. Effect reflected in SD of outcomes in European RCTs.
- Which treatment is better is not the question.
- Which treatment strategy is likely to have the best probability of providing the greatest benefit to each individual patient.

4 European RCTs: Differences in Patient Population

- Patients were randomized at the time of presentation as opposed to after an adequate trial of non-operative measures.
  -Not similar to NA clinical practice

Assumption

Surgery and nonoperative care are competitive/interchangeable treatments and utilized under similar circumstances.
<table>
<thead>
<tr>
<th>4 European RCTs: Differences in Patient Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Surgery and nonoperative management are not competitive treatments</td>
</tr>
<tr>
<td>- Surgery only considered in select patients following a failure to improve with nonoperative treatment.</td>
</tr>
<tr>
<td>- Surgery and nonoperative management have different mechanisms of action and are applied under different circumstances</td>
</tr>
<tr>
<td>- Complementary treatments utilized in series not in parallel.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4 European RCTs: Differences in Patient Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>- The selection criteria used for these studies was primarily based on plain films. No imaging in Fairbank RCT.</td>
</tr>
</tbody>
</table>
  - Not similar to NA clinical practice |
| Assumption |
| - Most, if not all, patients with chronic low back pain (mean 8 years) are surgical candidates. |
4 European based RCTs

- Dichotomous hypothesis based on flawed assumptions
- Unclear inclusion/exclusion criteria
- Eligible patients who refused randomization
- Variability in treatments
- Limited assessment of operative objective
- Crossovers without ‘as treated’ analysis
- Underpowered (Brox et. al.-2)
- Unavailable control treatment in the USA in ¾ trials

Biases in Study Design:

- Intent to treat analyses significantly bias results against surgery
- Non-standardized diagnostic criteria and non-standardized treatment limits any conclusions that can be drawn
- Back and leg pain measures were only secondary outcome measures
Example of Bias

- Fairbank Study:
  - Diagnostic criteria poorly described
  - Surgical intervention all over map
    - Many patients not fused
  - Intent to treat analysis with 30% crossover to surgery
  - Significant loss of power due to 25% lost to follow-up in surgical group

Surgical Results

- All of these studies demonstrated significant (statistically and clinically important) improvement in the surgical group that exceeded the \textit{a priori} stated MCID in the primary and some secondary outcome measures compared to baseline
- Net improvement surgery vs. nonoperative care exceeded \textit{a priori} MCID in the two adequately powered RCTs (Fritzell, Fairbank)
Surgical Results from European Studies

<table>
<thead>
<tr>
<th>Study</th>
<th>ODI</th>
<th>SF-36 PCS/GFS</th>
<th>Pain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fairbank, 2006</td>
<td>12.5</td>
<td>15 (PCS)</td>
<td>19.5 (SF-36)</td>
</tr>
<tr>
<td></td>
<td>(SD-21.1)</td>
<td></td>
<td>(SD-26.4)</td>
</tr>
<tr>
<td>Brox, 2003</td>
<td>15.6</td>
<td>NA</td>
<td>20.7 (VAS)</td>
</tr>
<tr>
<td></td>
<td>(SD-16.4)</td>
<td></td>
<td>(SD-27.3)</td>
</tr>
<tr>
<td>Fritzell, 2001</td>
<td>11.6</td>
<td>15 (GFS)</td>
<td>21 (VAS)</td>
</tr>
<tr>
<td></td>
<td>(SD-18)</td>
<td></td>
<td>(SD-25.2)</td>
</tr>
</tbody>
</table>

USA based FDA IDE trials

- Prospective, randomized
- Strict, explicit inclusion/exclusion criteria
- Failed 6 months nonoperative care
- Standardized surgical treatments
- Validated outcomes assessment
- High compliance
- Low attrition
A Prospective, Randomized, Multicenter Food and Drug Administration Investigational Device Exemptions Study of Lumbar Total Disc Replacement With the CHARITÉ™ Artificial Disc Versus Lumbar Fusion
Part I: Evaluation of Clinical Outcomes

Table 2. Inclusion and Exclusion Criteria

<table>
<thead>
<tr>
<th>Inclusion Criteria</th>
<th>Exclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age 18 to 64 yrs</td>
<td>Previous thoracic or lumbar fusion</td>
</tr>
<tr>
<td>Symptomatic DDD confirmed by discography</td>
<td>Current or prior Paracentesis at L4, L5, or S1</td>
</tr>
<tr>
<td>Single-level DDD at L4-5 or L5-S1</td>
<td>Symptomatic multilevel degeneration</td>
</tr>
<tr>
<td>Disability score &lt; 30</td>
<td>High-contrast herniated nucleus pulposus</td>
</tr>
<tr>
<td>VAS score ≤ 40 or 100</td>
<td>Spondylolisthesis &gt; 5 mm</td>
</tr>
<tr>
<td>Failed ≥ 6 times of appropriate nonsurgical care</td>
<td>Scoliosis &gt; 11°</td>
</tr>
<tr>
<td>Back, neck, leg pain with no nerve root compression</td>
<td>Nonsurgical, thoracic, or lumbar herniated nucleus pulposus</td>
</tr>
<tr>
<td>Able to tolerate anterior approach</td>
<td>Spondylolisthesis &gt; 5 mm</td>
</tr>
<tr>
<td>Able and willing to comply with follow-up schedule</td>
<td>Spinal tumor</td>
</tr>
<tr>
<td>Writing to give written informed consent</td>
<td>Osteoporosis</td>
</tr>
<tr>
<td></td>
<td>Osteopenia, osteoporosis, or metabolic bone disease</td>
</tr>
<tr>
<td></td>
<td>Infection</td>
</tr>
<tr>
<td></td>
<td>Recurrent herniated disc</td>
</tr>
<tr>
<td></td>
<td>Psychosocial disorder</td>
</tr>
<tr>
<td></td>
<td>Mental instability</td>
</tr>
<tr>
<td></td>
<td>Mental disorder</td>
</tr>
<tr>
<td></td>
<td>Malignant cell disease</td>
</tr>
<tr>
<td></td>
<td>Use of a bone growth stimulator</td>
</tr>
<tr>
<td></td>
<td>Participation in another study</td>
</tr>
<tr>
<td></td>
<td>Avascular disease</td>
</tr>
<tr>
<td></td>
<td>Chronic steroid use</td>
</tr>
<tr>
<td></td>
<td>Autoimmune disorder</td>
</tr>
<tr>
<td></td>
<td>Pregnancy</td>
</tr>
<tr>
<td></td>
<td>Other surgical surgery at affected level (except discotomy, laminotomy/foraminotomy, without accompanying laminotomy or foraminotomy at the same level to be treated)</td>
</tr>
</tbody>
</table>

These Results are Generally Inferior to those seen in North American IDE Studies

- ALIF Control from Charite Trial: 35 point improvement in VAS, 25 point improvement in ODI
- BMP Bone Dowel Control: 29 point ODI improvement
- Predisc control group- 30 point improvement in ODI, 21 point VAS
- All with at least 2 year f/u
- Differences likely due to patient selection
ECRI and Intensive Multidisciplinary Nonoperative Management

- ECRI review implies that regimens similar to Brox regimen exist in Washington State (page 39 reference 114)
- Treatment at multidisciplinary centers has not been associated with improved outcomes (notably in Washington State)

Outcomes of Pain Center Treatment in Washington State Workers’ Compensation

James P. Robinson, MD, PhD, Deborah Fulton-Kehoe, MPH, Donald C. Martin, PhD, and Gary M. Franklin, MD, MPH

Results: Univariate analysis revealed that at 2-year follow-up, 35% of treated subjects were receiving time loss payments vs. 40% of evaluated only subjects ($P < 0.05$). Subjects who were younger, female, and less chronic were more likely to undergo pain center treatment and were less likely to be on time loss at 2-year follow-up. In multivariate analyses, which statistically controlled baseline differences between the two groups, there was no difference between treated subjects and evaluated only subjects.

ECRI estimation of “true impact of surgery”

We consider the estimates in this study to be the best empirical estimates of clinically important change in ODI and VAS. Although the pre-post change in ODI is technically a different concept than the between-group difference in ODI, we consider the two concepts similar enough that the pre-post change can be used as a surrogate for the between-group difference. We view the control group as a surrogate for the active-treatment group: what change would the surgical patients have experienced if they had received non-surgical treatment? Taking this view, the between-group difference at followup is therefore an unbiased estimate of the surgical group’s change score after factoring out the non-surgical treatment effect. Thus, our use of 10 units on Oswestry is, technically speaking, the change score at which the true impact of surgery is considered clinically significant. Accordingly, we used a difference of 10 for the ODI and a difference of 20 for the VAS as the minimal clinically important difference in our assessment of these outcomes.

- Mean surgery outcome - mean nonoperative outcome = “true impact of surgery”
ECRI Institute evidence assessments:

We did not find sufficient evidence that lumbar fusion surgery is more effective to a clinically meaningful degree than nonsurgical treatments for any of the following patient populations, comparisons and outcomes:

* Homogeneity of condition/uniform Rx. response
* Surgery/nonoperative are competitive treatments
* Variable/unclear inclusion criteria
* High crossover rate; no ‘as treated’ analysis
* Nonoperative treatment not available or effective in the US

Analysis Flawed

Notable critics of lumbar fusion for axial back pain have reviewed the same literature and concluded that the trials were inadequate to support any firm conclusions regarding the superiority of fusion compared to “control” measures.
"Methodological limitations of the randomized trials prevent firm conclusions."

"These trials do not allow a general statement regarding the efficacy of fusion over nonoperative care for discogenic back pain."

Spine 32: 816-823, 2007

Fusion Results

- Lumbar fusion is consistently associated with clinically relevant and durable improvements in validated outcomes measures in properly selected patients.
Alternative Treatments

- If fusion is not performed, what alternatives exist for patients with disabling low back pain?
- ECRI analysis showed no benefit of either intensive cognitive or unstructured PT that reached MCID in any RCT or meta-analysis (Brox/Fairbank)

Conclusion

- Lumbar Fusion works in properly selected patients with disabling LBP who have failed an adequate trial of non-operative treatment and who have appropriate physical examination and imaging characteristics.
<table>
<thead>
<tr>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>- No effective alternative treatment modality exists for these patients in the United States, including Washington State.</td>
</tr>
</tbody>
</table>
Spinal Fusion and Discography for Chronic Low Back Pain and Uncomplicated Lumbar Degenerative Disc Disease

Health Technology Clinical Committee Meeting
Washington State Health Technology Assessment Program
James Reston, Ph.D., M.P.H.
Jonathan Treadwell, Ph.D.
Karen Schoelles, M.D., S.M.
Seattle, Washington
November 16, 2007
Scope of Report

This report evaluates relevant published research describing use of lumbar fusion and discography in patients with chronic low back pain and uncomplicated degenerative disc disease (DDD).

The word “uncomplicated” in the title of this report is intended to exclude patients who had fusion for the following conditions:

- Radiculopathy
- Functional neurologic deficits (motor weakness or EMG findings of radiculopathy)
- Spondylolisthesis (>Grade 1)
- Isthmic spondylolysis
- Primary neurogenic claudication associated with stenosis
- Fracture, tumor, infection, inflammatory disease
- Degenerative disease associated with significant deformity
Scope of Report

Therefore, the conclusions of this report are not necessarily applicable to patients undergoing fusion for any of the excluded conditions:
(radiculopathy, functional neurologic deficits, spondylolisthesis > grade 1, isthmic spondylolysis, primary neurogenic claudication associated with stenosis, fracture, tumor, infection, inflammatory disease, degenerative disease associated with significant deformity)

Epidemiology of Low Back Pain

• Most common cause of disability in patients under age 45
• Causes greater loss of productivity than any other medical condition
• 1.2 million patients in the U.S. disabled by chronic low back pain
Degenerative Disc Disease and Low Back Pain

- DDD can be identified by plain radiograph, CT, or MRI
- DDD can occur at any level, but is not always associated with pain
- No clear case definition for “discogenic back pain”

Non-Operative Therapies

- Back education
- Medications (including epidural injections)
- Weight reduction
- Exercise
- Physical therapy
- Cognitive behavioral therapy
- Chiropractic manipulation
- Acupuncture
- Therapeutic massage
Spinal Fusion

- DDD in association with chronic low back pain that has not responded to conservative therapy is considered by many surgeons as an indication for spinal fusion.
- Goal is to permanently immobilize the spinal column vertebrae surrounding the disc(s) that is(are) diagnosed as the cause of chronic low back pain.

Several fusion procedures are currently used in practice. They differ by surgical approach (e.g. anterior, posterior, posterolateral, transforaminal, circumferential) and instrumentation used (e.g. various types of cages, pedicle screws, rods). All methods may have advantages and disadvantages.
Rates and Regional Variation

- HCUP NIS data revealed a 220% increase in rates of lumbar fusion for degenerative conditions in the U.S. between 1990 and 2001 (Deyo et al. 2005)

- Medicare data revealed a nearly 20-fold variation in the range of regional rates of lumbar fusion (for any indication) among Medicare enrollees in the U.S. in 2002 and 2003 (Weinstein et al. 2006)

Spinal Fusion - Key Questions

1) Does lumbar fusion surgery reduce pain and improve functional status/quality of life more effectively than nonsurgical treatments?
2) What are the rates of adverse events (perioperative, long-term events, and reoperations) for lumbar fusion surgery and nonsurgical treatments?

3) What patient characteristics (i.e., workers’ compensation population, patients with chronic pain, psychological distress, and age-groups) are associated with differences in the benefits and adverse events of lumbar fusion surgery?
Inclusion Criteria

• Peer-reviewed full-length publications (no meeting abstracts or supplements)
• English language publications
• Studies with data on relevant outcomes
• If multiple publications of the same study, only the largest and most recent publication was used (unless other reports have non-overlapping data)

Inclusion Criteria

• Patients had chronic (3+ months) lumbar pain. At least 80% did not have excluded conditions
• At least 80% of patients must have contributed follow-up data to a given time point
• At least 80% of patients enrolled for fusion must have received fusion
• For outcome data related to pain, quality of life, or functional status, the study must have used a previously validated instrument
Inclusion Criteria

- Key Question 1 and 3: only RCTs comparing lumbar fusion to a nonsurgical approach with at least 10 patients per treatment arm (1990 or later)
- Key Question 2: Studies that met criteria for KQ1 or KQ3, OR other studies of lumbar fusion that enrolled at least 100 patients (1990 or later)

Key Question 1

- Subjectivity of pain, function, and QoL creates vulnerability to measurement biases, placebo effects or regression to the mean
- Several types of surgery or invasive treatments have shown evidence of placebo effects
- Parallel randomized treatment group is the best method to control for potential biases
Literature Search – Spinal Fusion

- Medical librarians searched 15 databases
- Last search date 10/15/07
- 482 articles identified
- 243 retrieved
- 30 included (27 unique studies)
- 4 for KQ1, 27 for KQ2, 1 for KQ3

Strength of Evidence

- Quality, quantity, consistency, robustness, and magnitude of effect
- Strong, Moderate, Weak, or Insufficient
- Separately assessed for different outcomes
Studies Comparing Spinal Fusion and Non-Operative Therapy

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Design</th>
<th>No. of patients randomized</th>
<th>Followup</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brox et al. 2006</td>
<td>Norway</td>
<td>RCT</td>
<td>60</td>
<td>1 year</td>
</tr>
<tr>
<td>Fairbank et al. 2005</td>
<td>UK</td>
<td>RCT</td>
<td>349</td>
<td>2 years</td>
</tr>
<tr>
<td>Brox et al. 2003</td>
<td>Norway</td>
<td>RCT</td>
<td>64</td>
<td>1 year</td>
</tr>
<tr>
<td>Fritzell et al. 2001</td>
<td>Sweden</td>
<td>RCT</td>
<td>294</td>
<td>2 years</td>
</tr>
</tbody>
</table>

Internal Validity

- All randomized
- None had blinding of patients, providers, and outcome assessors
- 2 studies had >15% of patients that did not receive assigned treatment
Internal Validity

• 2 studies had baseline between-group differences in important patient characteristics
• Possible differences between groups in ancillary treatments (only one study reported information on ancillary treatments)
• Subjectivity of outcomes (pain, functional status, QoL)
• All studies scored as moderate quality

Generalizability

• Non-operative treatment: 3 RCTs used an intensive rehabilitation program (including CBT), while one used non-intensive physical therapy
• Diagnostic criteria: 3 RCTs used a less stringent diagnostic criterion (DDD on plain radiograph) than is typically used in the U.S.
Generalizability

- Fusion strategies: different surgical approaches and instrumentation used in different studies
- Patients: Average age 40-45 years, but other characteristics differed among studies. One study excluded patients with prior back surgery, one study included only patients with prior back surgery, remaining studies included mostly patients without prior back surgery. One study had 11% patients with spondylolisthesis.

Key Question 1

Does lumbar fusion surgery reduce pain and improve functional status/quality of life more effectively than nonsurgical treatments?
Key Question 1: Analysis

- Due to differences among studies, data from the 4 RCTs could not be combined
- One study (Fritzell et al. 2001) that used non-intensive physical therapy as a control was analyzed separately from studies that used intensive rehabilitation
- One study (Brox et al. 2006) that included only patients with prior back surgery was analyzed separately from other studies of intensive rehabilitation

Three Analysis Groups

- Fusion versus Intensive Exercise/Rehabilitation Plus CBT in Patients without Prior Back Surgery
  - 2 studies, total of 413 patients
- Fusion versus Intensive Exercise/Rehabilitation Plus CBT in Patients with Prior Back Surgery
  - 1 study, 60 patients
- Fusion versus Non-intensive Physical Therapy in Patients without Prior Back Surgery
  - 1 study, 294 patients
Selection of Functional Status Measure

- Return to work – problems with this outcome.
  - Ability to work is not synonymous with return to work
  - Return to same work or less physically demanding work? (not all workers can easily change their work conditions)
  - Studies have shown return to work often largely governed by factors other than spinal symptoms
  - Not a relevant outcome for certain subgroups of patients (homemakers, students, pensioners)

- For these reasons, we did not consider return to work to be the best measure of functional status

Key Outcome for Functional Status

**Oswestry Disability Index (ODI)** – measures functional ability in patients treated for back pain

- Well-validated and is measurable in all patients (unlike return to work)
- Controversy in literature regarding minimum clinically important difference (MCID) in ODI – range of estimates from 4 to 18 (FDA uses 15)
- We considered 10 points to be the best empirical estimate in the literature (Hagg et al. 2003)
Other Key Outcomes

- **Visual Analog Scale (VAS)** for back pain
  - Based on literature, we defined a difference of 20 as the minimum clinically important difference (MCID) in VAS for back pain

- **Quality of life**: Short-form (SF) 36
  - Based on literature, we defined a difference of 5 as the MCID for the SF-36 instrument

Fusion vs Intensive Rehabilitation – Patients without Prior Back Surgery

- 2 moderate-quality RCTs with 413 patients
- One reported data at 1 year followup, the other reported data at 2 years’ followup
- Both studies reported the between-group difference in the pre-post change in ODI score, adjusted for baseline values
- The data were combined in a random effects meta-analysis
Fusion vs Intensive Rehabilitation – Patients without Prior Back Surgery

- Meta-analysis of difference in ODI change scores

<table>
<thead>
<tr>
<th>Study name</th>
<th>Statistics for each study</th>
<th>Difference in means and 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Difference in means</td>
<td>Lower limit</td>
</tr>
<tr>
<td>Brox 2003</td>
<td>2.70</td>
<td>-6.61</td>
</tr>
<tr>
<td>Fairbank 2005</td>
<td>4.10</td>
<td>0.12</td>
</tr>
<tr>
<td>Summary</td>
<td>NC</td>
<td>0.22</td>
</tr>
</tbody>
</table>

No clinically meaningful difference was observed between fusion and intensive rehabilitation plus CBT in patients without prior back surgery (95% CI 0.2 to 7.5 points on ODI, with minimum clinically important difference defined \textit{a priori} as 10 points), although the difference slightly favored fusion.

Strength of evidence: Weak
Fusion vs Intensive Rehabilitation – Patients without Prior Back Surgery

- Only 1 of the 2 RCTs reported VAS for back pain
- No statistically significant difference in change in VAS scores between fusion and rehabilitation groups
- Because this is a single moderate quality study, the evidence is insufficient to allow a conclusion for this outcome
- The only study to report quality-of-life data lacked 80% followup for this outcome

Fusion vs Intensive Rehabilitation – Patients with Prior Back Surgery

- 1 moderate-quality RCT with 60 patients
- Inconclusive findings for ODI and VAS
- The evidence was insufficient to determine the relative benefits of lumbar fusion compared to intensive rehabilitation plus CBT in patients with prior back surgery
Fusion vs Non-Intensive Physical Therapy – Patients without Prior Back Surgery

- 1 moderate-quality RCT with 294 patients
- No clinically significant change in mean ODI or VAS in fusion group compared to physical therapy group, although differences were statistically significant
- The evidence was insufficient to determine the relative benefits of lumbar fusion compared to conventional physical therapy in patients without prior back surgery

Key Question 2

What are the rates of adverse events (perioperative, long-term events, and reoperations) for lumbar fusion surgery and nonsurgical treatments?
Fusion vs Non-Operative Therapy – Adverse Events

- 4 RCTs with 767 patients
- All trials calculated adverse event rates on a per protocol basis (only patients who actually received surgery were included in calculations)
- No adverse events reported in non-operative control groups in any of these trials
- Most early and late events reported in surgical groups could not have occurred in the absence of surgery

Additional Studies – Adverse Events

- 23 studies with 5,639 patients – reported adverse events of fusion, but no non-operative control groups
- 14 studies were prospective, 9 were retrospective
- Some studies reported all adverse events, others focused on specific events (such as reoperation)
Adverse Events Reported in ≥Two Studies

<table>
<thead>
<tr>
<th>Adverse event</th>
<th>No. of studies reporting event</th>
<th>Range of reported event rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reoperation</td>
<td>18/27 (1 reported 0 events)</td>
<td>0% to 46.1%</td>
</tr>
<tr>
<td>Infection (deep or superficial)</td>
<td>14/27 (1 reported 0 events)</td>
<td>0% to 9%</td>
</tr>
<tr>
<td>Neurologic</td>
<td>12/27 (no study reported 0 events)</td>
<td>0.7% to 25.8%</td>
</tr>
<tr>
<td>Bleeding/vascular injury</td>
<td>10/27 (2 reported 0 events)</td>
<td>0% to 12.8%</td>
</tr>
<tr>
<td>Thrombosis</td>
<td>11/27 (1 reported 0 events)</td>
<td>0% to 4%</td>
</tr>
<tr>
<td>Dural injury</td>
<td>10/27 (no study reported 0 events)</td>
<td>0.5% to 29%</td>
</tr>
<tr>
<td>Hematoma</td>
<td>7/27 (no study reported 0 events)</td>
<td>1% to 4%</td>
</tr>
<tr>
<td>Retrograde ejaculation</td>
<td>6/27 (no study reported 0 events)</td>
<td>0.7% to 6%</td>
</tr>
<tr>
<td>Device-related</td>
<td>13/27 (1 reported 0 events with a specific type of fusion)</td>
<td>0% to 17.8%</td>
</tr>
<tr>
<td>Death</td>
<td>4/27 (the other 22 studies were assumed to have 0 surgically-related deaths)</td>
<td>0% to 2%</td>
</tr>
</tbody>
</table>

Key Question 3

What patient characteristics (i.e., workers’ compensation population, patients with chronic pain, psychological distress, and age-groups) are associated with differences in the benefits and adverse events of lumbar fusion surgery?
Key Question 3

- 1 RCT with 294 patients
- Used multiple logistic regression to correlate patient characteristics with certain outcomes in the surgical and nonsurgical groups
- Because this was a single moderate-quality study with no large associations, the evidence was considered insufficient to allow a conclusion

Discography

- Diagnostic procedure
- Is the disc itself the source of pain?
- Injection of dye directly into the spinal disc:
  - Typical pain reproduced?
  - Abnormal morphology?
Pain provocation

- Subjective response
- An emphasis on “typical” pain
- Considered by most practitioners to be the key discography finding

Morphology

- Integrity of the disc annulus
- Dye leakage
- Dallas Discogram Description:
  - Grade 0: Normal
  - Grade 1: Leak to inner 1/3 of annulus
  - Grade 2: Leak to inner 2/3 of annulus
  - Grade 3: Through annulus
  - Grade 4 or 5: Beyond annulus
False positives

- Abnormal morphology despite no prior lumbar pain
- Typical *non-back* pain reproduced by discography
- Walsh criteria for a positive discography
- Adjacent disc tests

---

Discography - Key Questions

**Reliability**

4) In patients being considered for lumbar fusion surgery, what is the reliability of discography?
   - Test-retest reliability
   - Inter-reader reliability
Discography - Key Questions

**Prediction**

5) In patients undergoing lumbar fusion surgery, do the results of pre-surgical discography predict the degree of pain reduction or improvement in functional status/quality of life after lumbar fusion surgery?

---

Discography - Key Questions

**Impact**

6) In patients being considered for lumbar fusion surgery, do patients who receive discography that influences the treatment choice have better treatment outcomes than patients who do not receive discography?
Inclusion Criteria

- Similar general inclusion criteria as for fusion
- Exceptions:
  - No control group required for the reliability question
  - Randomization not required
  - No date restriction

Discography - Key Questions

Reliability

4) In patients being considered for lumbar fusion surgery, what is the reliability of discography?
   - Test-retest reliability
   - Inter-reader reliability
Evidence on Reliability

• No studies reported any of the following:
  - Reliability of discography result when different people perform the injection
  - Reliability of discography on the same disc at different times
  - Reliability of patients’ reports of pain provocation

• But some data exist on whether a given discogram is judged to have the same morphology grade:
  - By the same reader at different times (1 study, N=72)
  - By different readers (2 studies, N=72 and N=45)

• Moderate quality
### Test-Retest Reliability Data

<table>
<thead>
<tr>
<th>Study</th>
<th>Discs</th>
<th>Test-retest kappa (95% CI)</th>
<th>Rater 1</th>
<th>Rater 2</th>
<th>Rater 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agorastides</td>
<td>133</td>
<td></td>
<td>0.80</td>
<td>0.85</td>
<td>0.80</td>
</tr>
<tr>
<td>(2002)</td>
<td></td>
<td>(0.71 to 0.89)</td>
<td>(0.77 to 0.93)</td>
<td>(0.70 to 0.90)</td>
<td></td>
</tr>
</tbody>
</table>

Not enough data to permit a conclusion

### Inter-Rater Reliability Data

<table>
<thead>
<tr>
<th>Study</th>
<th>Discs</th>
<th>System</th>
<th>Kappa (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agorastides</td>
<td>133</td>
<td>Adams classification</td>
<td>0.77 (0.66 to 0.87)</td>
</tr>
<tr>
<td>(2002)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milette (1999)</td>
<td>132</td>
<td>DDD degeneration</td>
<td>0.67 (0.55 to 0.78)</td>
</tr>
<tr>
<td>Milette (1999)</td>
<td>132</td>
<td>DDD disruption</td>
<td>0.66 (0.56 to 0.76)</td>
</tr>
</tbody>
</table>

Not enough data to permit a conclusion
Discography - Key Questions

Prediction

5) In patients undergoing lumbar fusion surgery, do the results of pre-surgical discography predict the degree of pain reduction or improvement in functional status/quality of life after lumbar fusion surgery?

Evidence on Prediction

- 3 studies, all Low quality
- Different definitions of a positive test:
  - Willems (2007) had 2 groups, based on pain provocation in adjacent discs (total N=82)
  - Gill (1992) had 3 groups, based on morphology of suspected disc (total N=53)
  - Colhoun (1988) had 4 groups, based on both pain provocation and morphology of suspected disc (total N=195).
Evidence on Prediction

- Different outcomes were assessed:
  - Willems (2007) measured VAS pain scores at mean 6.7 years’ followup
  - Gill (1992) measured a composite outcome (based on ODI, VAS, and pain drawing) at mean 3 years’ followup
  - Colhoun (1988) measured a composite outcome of “success” at mean 3.6 years’ followup

Evidence on Prediction

- Different results were found:
  - Willems (2007): No difference in surgical outcomes between those with positive discography (+) and those with negative discography (-)
  - Gill (1992): Inconclusive findings
  - Colhoun (1988): Surgical outcomes were better among those with positive discography (+)
Summary of Prediction

Due to low quality, as well as major differences in designs, assessments and outcomes, no conclusion is warranted.

Discography - Key Questions

**Impact**

6) In patients being considered for lumbar fusion surgery, do patients who receive discography that influences the treatment choice have better treatment outcomes than patients who do not receive discography?
Evidence on Impact

- Only one study: N=32 who received discography and N=41 who did not
- All patients received fusion
- Retrospective, non-concurrent, non-randomized, unblinded, poor matching at baseline
- Very low quality

Summary of Discography

- Insufficient evidence for each of the 3 Key Questions pertaining to discography:
  - For #4 (reliability), the primary reason was low quantity
  - For #5 (prediction), there were major inconsistencies in definitions, post-surgical variables assessed, and reported results
  - For #6 (impact), there was only one very-low quality study
Washington State Health Care Authority

Health Technology Assessment
Spine Fusion and Discography
Clinical Committee Meeting
November 16, 2007

Contributors and Participants

Richard N. Wohns, MD, AANS/CNS, Tacoma, WA

Jens R. Chapman, MD, AAOS, University of Washington, Seattle, WA

Ray M. Baker, MD, NASS, University of Washington, Seattle, WA

Theodore A. Wagner, MD, AAOS, University of Washington, Seattle, WA

Paul C. McCormick, MD, Past Chair, AANS/CNS Section on Disorders of the Spine and Peripheral Nerves, Columbia University, New York, NY

Daniel K. Resnick, MD, Chair-Elect, AANS/CNS Section on Disorders of the Spine and Peripheral Nerves, University of Wisconsin, Madison, WI

Andrew T. Dailey, MD, AANS/CNS Section on Disorders of the Spine and Peripheral Nerves, University of Utah, Salt Lake City, UT

Joseph T. Alexander, MD, Chair, AANS/CNS Section on Disorders of the Spine and Peripheral Nerves, Scarborough, ME
Outline

- This discussion is limited to the performance of lumbar fusion for axial back pain thought to be due to degenerative disc disease (DDD)
- The utility of fusion as an adjunct to decompression or as a means of deformity correction in the North American patient population has been firmly established
  - SPORT II, 2007; Ghogowala and Benzel, 2006; Resnick et al, 2005; many others

Outline

- Efficacy studies for surgery (4 European RCTs- Fritzell, Brox(2), Fairbank)
- Study limitations (design, methods, assumptions)-biased to null against surgery
- Significant improvement in surgical groups compared to non-operative groups
- Failure of any type of nonoperative treatment to achieve MCID in ECRI meta-analysis or adequately powered RCT
Outline

- Outcomes of North American FDA IDE for degenerative disc disease
- Unavailability of structured PT in the United States
- ECRI’s calculation of “true impact of surgery”

4 European RCTs

- Single hypothesis-Is surgery better than nonoperative management.

Assumptions:
1. Low back pain is a homogeneous condition.
2. Each patient equally likely to achieve mean improvement associated with either treatment.
3. The mean outcome score for each treatment provides a consistently accurate estimate of every patient's response to either treatment.
4. Singular conclusion “which treatment is better” applies to all patients
### Low back pain

- Heterogeneous- cause, structural basis, natural hx, clinical aspects and patient effects, patient’s treatment preference and response
- Many patients do well with nonoperative care, some don’t.
- Many patients do well with surgery, some don’t.
- Wide variation in magnitude of Rx. Effect reflected in SD of outcomes in European RCTs.
- Which treatment is better is not the question.
- Which treatment strategy is likely to have the best probability of providing the greatest benefit to each individual patient.

### 4 European RCTs: Differences in Patient Population

- Patients were randomized at the time of presentation as opposed to after an adequate trial of non-operative measures.
  - Not similar to NA clinical practice

Assumption

Surgery and nonoperative care are competitive/interchangeable treatments and utilized under similar circumstances.
### 4 European RCTs: Differences in Patient Population

- Surgery and nonoperative management are not competitive treatments.
- Surgery only considered in select patients following a failure to improve with nonoperative treatment.
- Surgery and nonoperative management have different mechanisms of action and are applied under different circumstances.
- Complementary treatments utilized in series not in parallel.

---

### 4 European RCTs: Differences in Patient Population

- The selection criteria used for these studies was primarily based on plain films. No imaging in Fairbank RCT.
  - Not similar to NA clinical practice
- Assumption
  - Most, if not all, patients with chronic low back pain (mean 8 years) are surgical candidates.
4 European based RCTs

- Dichotomous hypothesis based on flawed assumptions
- Unclear inclusion/exclusion criteria
- Eligible patients who refused randomization
- Variability in treatments
- Limited assessment of operative objective
- Crossovers without ‘as treated’ analysis
- Underpowered (Brox et. al.-2)
- Unavailable control treatment in the USA in ¾ trials

Biases in Study Design:

- Intent to treat analyses significantly bias results against surgery
- Non-standardized diagnostic criteria and non-standardized treatment limits any conclusions that can be drawn
- Back and leg pain measures were only secondary outcome measures
Example of Bias

- Fairbank Study:
  - Diagnostic criteria poorly described
  - Surgical intervention all over map
    - Many patients not fused
  - Intent to treat analysis with 30% crossover to surgery
  - Significant loss of power due to 25% lost to follow-up in surgical group

Surgical Results

- All of these studies demonstrated significant (statistically and clinically important) improvement in the surgical group that exceeded the a priori stated MCID in the primary and some secondary outcome measures compared to baseline
- Net improvement surgery vs. nonoperative care exceeded a priori MCID in the two adequately powered RCTs (Fritzell, Fairbank)
### Surgical Results from European Studies

<table>
<thead>
<tr>
<th>Study</th>
<th>ODI</th>
<th>SF-36 PCS/GFS</th>
<th>Pain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fairbank, 2006</td>
<td>12.5</td>
<td>15 (PCS)</td>
<td>19.5 (SF-36)</td>
</tr>
<tr>
<td></td>
<td>(SD-21.1)</td>
<td></td>
<td>(SD-26.4)</td>
</tr>
<tr>
<td>Brox, 2003</td>
<td>15.6</td>
<td>NA</td>
<td>20.7 (VAS)</td>
</tr>
<tr>
<td></td>
<td>(SD-16.4)</td>
<td></td>
<td>(SD-27.3)</td>
</tr>
<tr>
<td>Fritzell, 2001</td>
<td>11.6</td>
<td>15 (GFS)</td>
<td>21 (VAS)</td>
</tr>
<tr>
<td></td>
<td>(SD-18)</td>
<td></td>
<td>(SD-25.2)</td>
</tr>
</tbody>
</table>

### USA based FDA IDE trials

- Prospective, randomized
- Strict, explicit inclusion/exclusion criteria
- Failed 6 months nonoperative care
- Standardized surgical treatments
- Validated outcomes assessment
- High compliance
- Low attrition
These Results are Generally Inferior to those seen in North American IDE Studies

- ALIF Control from Charite Trial: 35 point improvement in VAS, 25 point improvement in ODI
- BMP Bone Dowel Control: 29 point ODI improvement
- Prodisc control group- 30 point improvement in ODI, 21 point VAS
- All with at least 2 year f/u
- Differences likely due to patient selection
ECRI and Intensive Multidisciplinary Nonoperative Management

- ECRI review implies that regimens similar to Brox regimen exist in Washington State (page 39 reference 114)
- Treatment at multidisciplinary centers has not been associated with improved outcomes (notably in Washington State)

Outcomes of Pain Center Treatment in Washington State Workers’ Compensation

James P. Robinson, MD, PhD, Deborah Fulton-Kehoe, MPH, Donald C. Martin, PhD, and Gary M. Franklin, MD, MPH

Results: Univariate analysis revealed that at 2-year follow-up, 35% of treated subjects were receiving time loss payments vs. 40% of evaluated only subjects ($P < 0.05$). Subjects who were younger, female, and less chronic were more likely to undergo pain center treatment and were less likely to be on time loss at 2-year follow-up. In multivariate analyses, which statistically controlled baseline differences between the two groups, there was no difference between treated subjects and evaluated only subjects.

ECRI estimation of “true impact of surgery”

We consider the estimates in this study to be the best empirical estimates of clinically important change in ODI and VAS. Although the pre-post change in ODI is technically a different concept than the between-group difference in ODI, we consider the two concepts similar enough that the pre-post change can be used as a surrogate for the between-group difference. We view the control group as a surrogate for the active-treatment group: what change would the surgical patients have experienced if they had received non-surgical treatment? Taking this view, the between-group difference at followup is therefore an unbiased estimate of the surgical group’s change score after factoring out the non-surgical treatment effect. Thus, our use of 10 units on Oswestry is, technically speaking, the change score at which the true impact of surgery is considered clinically significant. Accordingly, we used a difference of 10 for the ODI and a difference of 20 for the VAS as the minimal clinically important difference in our assessment of these outcomes.

- Mean surgery outcome - mean nonoperative outcome = “true impact of surgery”

p. 27
ECRI Institute evidence assessments:

We did not find sufficient evidence that lumbar fusion surgery is more effective to a clinically meaningful degree than nonsurgical treatments for any of the following patient populations, comparisons and outcomes:

Fusion versus Intensive Exercise/Rehabilitation Plus CBT in Patients without Prior Back Surgery

- Meta-analysis of postoperative changes in Oswestry disability scores from two moderate-quality RCTs (n = 413 patients) revealed no clinically meaningful difference between fusion and intensive exercise/rehabilitation plus cognitive behavioral therapy (CBT) in patients without prior back surgery (95% CI 0.2 to 7.5, a priori 10 point difference defined as clinically meaningful), although the difference slightly favored fusion. Strength of evidence: Weak.
- The evidence was insufficient to determine whether lumbar fusion provides a greater improvement in back pain (one moderate-quality RCT, n = 64 patients) or quality of life (no acceptable evidence) compared to intensive exercise/rehabilitation plus CBT in patients without prior back surgery.
  - Homogeneity of condition/uniform Rx. response
  - Surgery/nonoperative are competitive treatments
  - Variable/unclear inclusion criteria
  - High crossover rate; no ‘as treated’ analysis
  - Nonoperative treatment not available or effective in the US

Analysis Flawed

- Notable critics of lumbar fusion for axial back pain have reviewed the same literature and concluded that the trials were inadequate to support any firm conclusions regarding the superiority of fusion compared to “control” measures.
“Methodological limitations of the randomized trials prevent firm conclusions.”

“These trials do not allow a general statement regarding the efficacy of fusion over nonoperative care for discogenic back pain.”

Spine 32: 816-823, 2007

Fusion Results

Lumbar fusion is consistently associated with clinically relevant and durable improvements in validated outcomes measures in properly selected patients.
Alternative Treatments

- If fusion is not performed, what alternatives exist for patients with disabling low back pain?
- ECRI analysis showed no benefit of either intensive cognitive or unstructured PT that reached MCID in any RCT or meta-analysis (Brox/Fairbank)

Conclusion

- Lumbar Fusion works in properly selected patients with disabling LBP who have failed an adequate trial of non-operative treatment and who have appropriate physical examination and imaging characteristics.
Conclusion

- No effective alternative treatment modality exists for these patients in the United States, including Washington State.
What about the HTA program?

Goal: Achieve better health by paying for technologies that work.
Focus: safety, efficacy and cost

Process: As defined in the WAC when making a coverage determination, committee members shall review and consider the health technology assessment. The committee may also consider other information it deems relevant, including other information provided by the administrator, reports and presentations from an advisory group, submission or comments from the public.

Is this evidence-based medicine?
TODAY’S DISCUSSION

• The HTA is probably not the optimal process for achieving better health for our patients.

• Evidence Based Medicine (EBM)
  – 3 components
  – Relationship to LBP and Spine Fusion
  – Critical Appraisal and Surgical Trials

• A proposal for an improved process for achieving better health for our patients.

EBM: What is it?

• “the explicit, judicious, and conscientious use of current best evidence from health care research in decisions about the care of individuals and populations.”

• Integrating this best evidence with clinical expertise and patient values

Guyatt 92, 2004,2006; Sackett EBM Text 2005
EBM

Clinical State/Circumstances
Clinical Expertise
Patient Preference
Research Evidence

Hatynes and Guyatt, 2002, 2006
Components of EBM
1. Best available evidence, 2. Clinical Expertise
3. Patient Preference

Question/s

1. Does lumbar fusion surgery reduce pain and improve function more effectively than non-surgical treatments for DDD?
2. Rates of Adverse Events?
3. Patient Characteristics that are associated with 1 and 2?

• What about the disease and intervention?

HTA ECRI REPORT
Oct. 19, 2007

“However, there is currently no clear case definition for discogenic back pain.”

Spine: Volume 31(18) 15 August 2006 pp 2151-2161
What is Intervertebral Disc Degeneration, and What Causes It?
[Literature Review]
Adams, Michael A. PhD*; Roughley, Peter J. PhD†
Components of EBM

Question/s

1. Does lumbar fusion surgery reduce pain and improve function more effectively then non surgical treatments for DDD?
2. Rates of Adverse Events?
3. Patient Characteristics that are associated with 1 and 2?
Quality of the Evidence

• Study design
• Quality
• Consistency
• Directedness

“The average quality of studies was moderate due to several limitations, most notably lack of blinding of patients, providers, and outcome assessors (for the majority of outcomes) in all studies.”

HTA Report Oct. 2007

Levels of Evidence and the Question

• Critical Relationship
  – ? and study design

• Surgical versus Medical
  – Are the rules different?
  – New Technology
Recent Examples

Surgical versus Nonsurgical Treatment for Lumbar Degenerative Spondylolisthesis


Surgical vs Nonoperative Treatment for Lumbar Disk Herniation
The Spine Patient Outcomes Research Trial (SPORT): A Randomized Trial

JAMA 2007

Quality of the Evidence

Question/s

1. Does lumbar fusion surgery reduce pain and improve function more effectively than nonsurgical treatments for DDD?
2. Rates of Adverse Events?
3. Patient Characteristics that are associated with 1 and 2?

• Study design
• Consistency
• Quality
• Directedness
3. Patient Characteristics that associated with questions 1 and 2?

"To achieve a balanced view when formulating recommendations, a multidisciplinary panel with broad representation, including clinicians, methodologists, generalists, patient representatives, and experienced guideline developers, should be assembled and proper group processes for reaching consensus on guidelines should be followed."
What is the correct process?

- HTA methodology is excellent, but only a component of EBM.
- Integrate Clinical Expertise and Patient Values.
- Determine appropriate research methodology for this population.
- Partner with Key Opinion Leaders.
- Partner with Industry.

Summary

- Literature Appraisal is only a component of EBM.
- DDD is ill defined and poorly understood.
- Rules of evidence for surgical trials is unique.
- Fusion for LBP is effective but the indications are not clear.
- A new process for HTA for fusion and LBP (DDD) should be initiated.
THANK YOU
WA Health Technology Assessment
Lumbar Fusion and Discography Review
November 16, 2006

Paul Schwaegler, MD
Orthopedics International
Seattle, WA

Outline

- U.S. Rates and Variations for Lumbar Fusion – Medicare Patients
- Review of WA State Workers’ Compensation Lumbar Fusion Rates and Outcomes
- Examination of the Effects of Delayed Surgical Care
- Washington Patients – Workers Compensation vs. other payors?
US Trends and Regional Variation in Lumbar Spine Surgery

- Weinstein, JN et al. – Spine, 31 (23):2707-2714 2006
  - 1992-2003 Rates of discectomy and fusion have increased in Medicare Patients, as well as regional variations – why??
  - Caveats – Based on Claims data, inability to determine primary diagnosis/indications, number of disabled patients or comorbidity status, number of revision patients included, which fusion tx works best, or per capita spine surgeon availability in U.S… that being said…
    - Access to better technology and tx –
      - More treatment options than 15 yrs ago – biologics, fusion, open, MIS, etc.
      - Relatively short pt. wait time and surgery access following failure of conservative care → best outcomes (Braybooke 2007)
      - Better informed (internet) and more aggressive patients – unwilling to live in pain and longer life spans
      - Willingness to travel – (regional variation may be due to patients traveling further to see a specialist) – we even have patients leave the US for other countries to obtain surgery

Washington State Workers’ Compensation Lumbar Fusion Rates and Outcomes

- Key Elements WA Workers’ Compensation Studies:
    - Claims Based Study for 1986-1987
    - Fusion incidence 41.7/100,000
    - Time of injury to tx was mean of 2.5 yrs
    - Unable to determine fusion type, and assess if one tx more effective
    - Patient satisfaction survey via phone 4-5 yrs post surgery
    - Inability to distinguish primary diagnosis, and those with prior surgery 1986-1987 (revisions) lumped into primary cohort
    - 65% patients contacted successfully about RTW and satisfaction (4-5 yrs post surgery)
      - Unknown if those not contacted RTW or disabled
      - No description in paper about partial RTW
      - No description in paper about those near retirement age and most likely not going back to work following surgery
    - Long wait time related to poor outcome and satisfaction – suggesting earlier surgery would be beneficial
    - Long wait time and poor outcomes also documented in (Franklin Am J Indust Med 1996, Braybooke, Eur Spine J 2007)
Washington State Workers’ Compensation Lumbar Fusion Rates and Outcomes

- Key Elements WA Workers’ Compensation Studies:
    - Claims Based Study for 1994-2001 (during the time period of cage introduction, new technology) more recent period would be better reflection of current uses
    - Fusion incidence 14.6/100,000 – 1994 and 19.6/100,000 in 2001
      - 2001: 19.6/100,000 vs. 1986-1987: 41.7/100,000 – Tremendous reduction in fusions!
    - Restrictions in workers compensation fusion guidelines?
  - Time of injury to tx was mean of 3.0 yrs
  - Unable to determine fusion type, and assess if one tx more effective
  - Inability to distinguish primary diagnosis, and those with prior surgery (revisions) lumped into cohort
  - RTW questions
    - Unknown if those not contacted RTW or disabled
    - No description in paper about partial RTW
    - No description in paper about those near retirement age and most likely not going back to work following surgery, so
  - Long wait time related to poor outcome and satisfaction – suggesting earlier surgery would be beneficial also documented in (Franklin Am J Indust Med 1996, Braybooke, Eur Spine J 2007)

Importance of the Effects of Delayed Care

- How does delayed care (limited patient access) affect WA patients?
  - Washington State and Other published literature clearly show that delayed care increases chances of poor outcome.
- Decisions by the HTA should consider patient outcomes, patient access and the effect delay of care will have.
Washington Patients – Workers Compensation vs. Other Payors

- Restrictive state WC guidelines create a second rate of care for injured workers
- Wait time is exorbitant for WC patients needing lumbar fusion vs. other WA patients
- HTA should be cautious when considering taking decisions away from WA spine surgeons that could delay patient care

Thanks!
Outcomes of Lumbar Fusion

Variation in Utilization, Efficacy, and Safety

Sohail K. Mirza, MD MPH
Professor, Department of Orthopedics and Joint Professor, Department of Neurological Surgery
University of Washington

Financial Conflict Disclosure Form

<table>
<thead>
<tr>
<th>Direct or indirect remuneration</th>
<th>NASS Format</th>
<th>Company/Companies</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Royalties</td>
<td></td>
<td>UW Tech Transfer (Synthes)</td>
</tr>
<tr>
<td>b. Stock ownership (options, warrants)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. Consulting fees</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. Loans from the sponsor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. Speaking arrangements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Holding a position in a company</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. Board of Directors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>g. Scientific Advisory Board</td>
<td></td>
<td></td>
</tr>
<tr>
<td>h. Other office in the company</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Receiving support from sponsors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>i. Endowments</td>
<td></td>
<td>Depuy, Surgical Dynamics, Synthes (to UW Dept. of Orthopedics)</td>
</tr>
<tr>
<td>j. Equipment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k. Biomaterials</td>
<td></td>
<td></td>
</tr>
<tr>
<td>l. Discretionary funds</td>
<td></td>
<td></td>
</tr>
<tr>
<td>m. Support of staff or training</td>
<td></td>
<td>Depuy, Synthes (to Dept of Orthop)</td>
</tr>
<tr>
<td>n. Trips</td>
<td></td>
<td></td>
</tr>
<tr>
<td>o. Other sponsorship</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

November 16, 2007
Disclosure

• I hold the University of Washington Surgical Dynamics Endowed Chair for Spine Outcomes Research (approx $90k in 2006).
• I receive royalties for surgical drills licensed by Synthes Spine through UW Office of Technology Transfer (approx. $16k in 2006).
• UW Department of Orthopedics receives spine fellowship support, research support, and endowments from Synthes Spine and Depuy Spine. I work with the spine fellows and am involved with two of the research projects supported by these funds.
• I prepared all the slides.

Rationale for Fusion

• Treat infection of tumor
• Correct deformity
• Stability after decompressing nerves?
• Excise pain??
Treatment Options for Discogenic Pain

- Observation, rehabilitation, pain management
- Intradiscal electrothermal coagulation (IDET)
- Posterolateral in situ fusion (no hardware)
- Instrumented posterior fusion (pedicle screws)
- Anterior lumbar interbody fusion (ALIF)
- Laparoscopic / Minimally Invasive fusion
- Posterior lumbar interbody fusion (PLIF)
- Combined anterior and posterior (360°) fusion
- Artificial Disc Replacement

**Variation**
Ratio of Back Surgery Rates

Geographical Variations in Spine Surgery Rates
(rate per 1,000 enrollees within the 2001 U.S. Medicare population)

<table>
<thead>
<tr>
<th>Low-rate states</th>
<th>High-rate states</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hawaii</td>
<td>Montana</td>
</tr>
<tr>
<td>Vermont</td>
<td>Oregon</td>
</tr>
<tr>
<td>New Jersey</td>
<td>Idaho</td>
</tr>
<tr>
<td>New York</td>
<td>Wyoming</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Low-rate cities</th>
<th>High-rate cities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Terre Haute, IN</td>
<td>Ft. Collins, CO</td>
</tr>
<tr>
<td>Bronx, NY</td>
<td>Eugene, OR</td>
</tr>
<tr>
<td>Honolulu, HI</td>
<td>Idaho Falls, ID</td>
</tr>
<tr>
<td>Wilkes-Barre, PA</td>
<td>Slidell, LA</td>
</tr>
<tr>
<td>Manhattan, NY</td>
<td>Amarillo, TX</td>
</tr>
<tr>
<td>McAllen, TX</td>
<td>Newport News, VA</td>
</tr>
<tr>
<td>Huntington, WV</td>
<td>Billings, MT</td>
</tr>
<tr>
<td>Hackensack, NJ</td>
<td>Greeley, CO</td>
</tr>
<tr>
<td>Lebanon, NH</td>
<td>Rapid City, SD</td>
</tr>
<tr>
<td>Newark, NJ</td>
<td>Casper, WY</td>
</tr>
<tr>
<td>East Long Island, NY</td>
<td>Boise, ID</td>
</tr>
<tr>
<td>Paterson, NJ</td>
<td>Bend, OR</td>
</tr>
</tbody>
</table>

Overall U.S. Rate 4.5
Variation in Lumbar Fusion Rates

Variation in Lumbar Surgery Rates

**Variation in Regional Rates**

Laminectomy: 8X

Fusion: 20X

Weinstein, Lurie et al Spine 2006
Causes of Variation

- Lack of scientific evidence
- Financial Incentives and Disincentives
- Clinical Training and Professional Opinion
- New technology

Weinstein, Lurie et al. Spine 2006

Annual Number of Operations in U.S.

Data from National Inpatient Sample, HCUP/AHRQ

Deyo, Nechemson, Mirza. NEJM 2004
Weinstein, Lurie et al Spine 2006

Inpatient Medicare Reimbursement

<table>
<thead>
<tr>
<th>Year</th>
<th>Rate of Lumbar Fusion</th>
<th>1992</th>
<th>30 per 100k</th>
<th>2003</th>
<th>110 per 100k</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spending for Lumbar Fusion</td>
<td>$75 million</td>
<td>$482 million</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Percent Spending for Fusion</td>
<td>14%</td>
<td>47%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Weinstein, Lurie et al Spine 2006
Efficacy

New Clinical Knowledge:

5 RCTs have compared fusion to non-operative treatment for chronic back pain:

Moller 2000*
Fritzell 2001
Brox 2003
Fairbank 2005
Brox 2006

*spondylolisthesis
### RCT: fusion vs. non-op results

<table>
<thead>
<tr>
<th></th>
<th>Fritzell</th>
<th>Brox 03</th>
<th>Fairbank</th>
<th>Brox 06</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final follow-up interval</td>
<td>2 years</td>
<td>1 year</td>
<td>2 years</td>
<td>1 year</td>
</tr>
<tr>
<td>Follow-up rate</td>
<td>98%</td>
<td>97%</td>
<td>82%</td>
<td>97%</td>
</tr>
</tbody>
</table>

**Surgery Group**

- **n=201**
- **n=35**
- **n=176**
- **n=29**
- Baseline Oswestry Index: 47
- Final Oswestry Index: 35
- Change (Final – Baseline): -12
- Percent Improvement: 24%

**Nonoperative Group**

- **n=63**
- **n=26**
- **n=173**
- **n=31**
- Baseline Oswestry Index: 48
- Final Oswestry Index: 45
- Change (Final – Baseline): -3
- Percent Improvement: 6%

Differential improvement across treatments (∆Surg – ∆Nonop)

- Change in Oswestry Index: 8.8 (Surg) vs. 2.3 (Nonop) vs. 3.8 (Surg) vs. (-3.9) (Nonop)
- Percent benefit with surgery: 19% (Surg) vs. 7.0% (Nonop) vs. 8% (Surg) vs. (-10%) (Nonop)

Mirza, Deyo, Spine 2007
Outcomes in Washington State

- Fusion (510 / 2546 reoperations)
- Non-fusion (4142 / 22767 reoperations)

Outcomes in Washington State

- Herniated Disc
  - Fusion (113 / 462 reoperations)
  - Non-fusion (3111 / 16071 reoperations)

- Degenerative
  - Fusion (103 / 515 reoperations)
  - Non-fusion (121 / 634 reoperations)

- Spinal Stenosis
  - Fusion (117 / 609 reoperations)
  - Non-fusion (824 / 5091 reoperations)

- Spondylolisthesis
  - Fusion (146 / 634 reoperations)
  - Non-fusion (42 / 156 reoperations)

Martin, Mirza, Deyo et al. Spine 2007

WA - Health Technology Clinical Committee
Outcomes in Washington State

- '90–93 (2752 / 25313 reoperations)
- '97–98 (1505 / 12520 reoperations)

Martin, Mirza, Deyo et al. Spine 2007

Outcomes in Washington State

- '90–93 (304 reops out of 2345 fusions)
- '97–00 (653 reops out of 4821 fusions)

Martin, Mirza, Deyo et al. Spine 2007
### Outcomes in Injured Workers

**Multivariate Analysis**

<table>
<thead>
<tr>
<th>Outcomes OR (95% CI)</th>
<th>Neither (Reference)</th>
<th>Cage alone</th>
<th>Instrumentation alone</th>
<th>Both</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Work disability</strong></td>
<td>1</td>
<td>1.46</td>
<td>1.07</td>
<td>1.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.98-2.16</td>
<td>0.78-1.47</td>
<td>0.72-1.57</td>
</tr>
<tr>
<td><strong>Postoperative complications</strong></td>
<td>1</td>
<td>1.98*</td>
<td>1.86*</td>
<td>2.20*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.02-3.80</td>
<td>1.07-3.22</td>
<td>1.16-4.16</td>
</tr>
<tr>
<td>Reoperation</td>
<td>1</td>
<td>0.82</td>
<td>0.97</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.53-1.27</td>
<td>0.69-1.36</td>
<td>0.50-1.26</td>
</tr>
</tbody>
</table>

*p<0.05

---

Juratli, Franklin, Mirza et al Spine 2006

---

Juratli, Franklin, Mirza et al Spine 2006
Safety

Number of subjects per group

<table>
<thead>
<tr>
<th>Difference in frequency, Treated - Untreated</th>
<th>Frequency in Untreated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
</tr>
</tbody>
</table>

Number of subjects per group vs. Difference in frequency, Treated - Untreated
**Systematic Review of Stand Alone Cages for Back Pain**

**Number of Studies Reporting Specified Complications by Surgical Approach**

<table>
<thead>
<tr>
<th>Complication</th>
<th>Anterior</th>
<th>Posterior</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
<td>n (%)</td>
</tr>
<tr>
<td>Non-union</td>
<td>15 (83)*</td>
<td>10 (100)</td>
<td>23 (85)</td>
</tr>
<tr>
<td>Reoperation</td>
<td>14 (74)*</td>
<td>7 (70)</td>
<td>20 (74)</td>
</tr>
<tr>
<td>Major Vessel Injury</td>
<td>12 (55)</td>
<td>1 (10)</td>
<td>12 (40)</td>
</tr>
<tr>
<td>Retrograde Ejaculation</td>
<td>12 (55)</td>
<td>1 (10)</td>
<td>12 (40)</td>
</tr>
<tr>
<td>Visceral Injury</td>
<td>3 (14)</td>
<td>0</td>
<td>3 (10)</td>
</tr>
<tr>
<td>Transfusion</td>
<td>3 (14)</td>
<td>1 (10)</td>
<td>4 (14)</td>
</tr>
<tr>
<td>Neurologic Complication</td>
<td>4 (18)</td>
<td>6 (60)</td>
<td>10 (33)</td>
</tr>
<tr>
<td>Dural Injury</td>
<td>2 (9)</td>
<td>7 (70)</td>
<td>8 (27)</td>
</tr>
<tr>
<td>Infection</td>
<td>3 (14)</td>
<td>5 (50)</td>
<td>8 (27)</td>
</tr>
</tbody>
</table>

(N=22 studies) (N=10 studies) (N=30 studies)

*Denominator excludes four studies that had less than six months of follow-up for fusion status.

^Denominator excludes three studies that reported outcomes only through the perioperative period.

_Fenton, Mirza, Deyo et al. Spine 2007_

---

**Reported non-union rates in studies of lumbar interbody fusion with stand-alone cage devices**

_Fenton, Mirza, Deyo et al. Spine 2007_
Potential Financial Conflicts of Interest

Favorable Results in Industry-Sponsored Research

<table>
<thead>
<tr>
<th>Sponsor</th>
<th>Odds Ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sponsor of study</td>
<td>3.6</td>
<td>2.6 to 4.9</td>
</tr>
<tr>
<td>For-profit organizations</td>
<td>5.3</td>
<td>2.0 to 14.4</td>
</tr>
<tr>
<td>Manufacturer of drugs</td>
<td>8.0</td>
<td>1.1 to 53.2</td>
</tr>
<tr>
<td>Spinal device manufacturer</td>
<td>3.3</td>
<td>2.4 to 4.5</td>
</tr>
</tbody>
</table>

Jacobs, Gatante, Mirza, Zdeblick JBJS 2006
Favorable Results

<table>
<thead>
<tr>
<th>Field</th>
<th>Industry-funded</th>
<th>Independent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spine</td>
<td>73</td>
<td>44</td>
</tr>
<tr>
<td>Hip</td>
<td>93</td>
<td>37</td>
</tr>
<tr>
<td>Knee</td>
<td>75</td>
<td>20</td>
</tr>
</tbody>
</table>

Conclusions

- Rates of lumbar fusion for chronic back pain have increased despite lack of efficacy data.
- Lumbar fusion for chronic low back pain offers little or no benefit compared to structured non-operative treatment.
- Safety data are limited and highly variable.
- Advances in technology have not improved outcomes.
- Investigator-sponsor financial conflicts are common.
Thank you.
HTCC Coverage and Reimbursement Determination Analytic Tool

HTA’s goal is to achieve better health care outcomes for enrollees and beneficiaries of state programs by paying for proven health technologies that work.

To find best outcomes and value for the state and the patient, the HTA program focuses on these questions:

1. Is it safe and effective?
2. Is it more effective or safer?
3. Is it equally effective and safe, and more cost-effective?

The principles HTCC uses to review evidence and make determinations are:

**Principle One: Determinations are Evidence based**

HTCC requires scientific evidence that a health technology is safe, effective and cost-effective as expressed by the following standards.

- Persons will experience better health outcomes than if the health technology was not covered and that the benefits outweigh the harms.
- The HTCC emphasizes evidence that directly links the technology with health outcomes. Indirect evidence may be sufficient if it supports the principal links in the analytic framework.
- Although the HTCC acknowledges that subjective judgments do enter into the evaluation of evidence and the weighing of benefits and harms, its recommendations are not based largely on opinion.
- The HTCC is explicit about the scientific evidence relied upon for its determinations.

**Principle Two: Determinations result in health benefit**

The outcomes critical to HTCC in making coverage and reimbursement determinations are health benefits and harms.

- In considering potential benefits, the HTCC focuses on absolute reductions in the risk of outcomes that people can feel or care about.
- In considering potential harms, the HTCC examines harms of all types, including physical, psychological, and non-medical harms that may occur sooner or later as a result of the use of the technology.
- Where possible, the HTCC considers the feasibility of future widespread implementation of the technology in making recommendations.
- The HTCC generally takes a population perspective in weighing the magnitude of benefits against the magnitude of harms. In some situations, it may make a determination for a technology with a large potential benefit for a small proportion of the population.
- In assessing net benefits, the HTCC subjectively estimates the indicated population's value for each benefit and harm. When the HTCC judges that the balance of benefits and harms is likely to vary substantially within the population, coverage or reimbursement determinations may be more selective based on the variation.
- The HTCC considers the economic costs of the health technology in making determinations, but costs are the lowest priority.

---

1 Based on Legislative mandate: See RCW 70.14.100(2).
2 The principles and standards are based on USPSTF Principles at: http://www.ahrq.gov/clinic/ajpmsuppl/harris3.htm
3 The principles and standards are based on USPSTF Principles at: http://www.ahrq.gov/clinic/ajpmsuppl/harris3.htm
HTCC Evaluation Factors

HTCC implements the program mandate and key principles that the decision be evidence based and that it be weighted most importantly on whether a given technology is safe and improves health through a decision tool.

Using Evidence as the basis for a Coverage Decision

Evaluate the primary coverage question by identifying for each primary factor (Safety, Effectiveness, and Cost) whether (1) evidence is available, (2) the confidence in the evidence, and (3) applicability to decision.

1. Availability of Evidence:
   Committee members decide whether information is available - Yes/No

2. Confidence in the Evidence:
   Committee members decide how confident they are in the scientific evidence by identifying the type and quality of evidence for consideration such as:
   - Type of evidence as reported in the technology assessment or other evidence presented to committee (randomized trials, observational studies, case series, expert opinion);
   - the amount of evidence (sparse to many number of evidence or events or individuals studied);
   - consistency of evidence (results vary or largely similar);
   - recency (timeliness of information);
   - directness of evidence (link between technology and outcome);
   - relevance of evidence (applicability to agency program and clients);
   - bias (likelihood of conflict of interest or lack of safeguards).

<table>
<thead>
<tr>
<th>Not Confident</th>
<th>Confident</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appreciable uncertainty exists. Further information is needed or further information is likely to change confidence.</td>
<td>Very certain of evidentiary support. Further information is unlikely to change confidence</td>
</tr>
</tbody>
</table>

3. Factors for Consideration - Importance

Committee members also consider the degree of importance that particular evidentiary information has to the policy and coverage decision. Factors used to assess level of importance are topic specific but most often include, for areas of safety, effectiveness, and cost:
   - risk of event occurring;
   - the degree of harm associated with risk;
   - the number of risks; the burden of the condition;
   - burden untreated or treated with alternatives;
   - the importance of the outcome (e.g. treatment prevents death vs relief of symptom);
   - the degree of effect (e.g. relief of all, none, or some symptom, duration, etc.);
   - value variation based on patient preference.

---

4 Based on GRADE recommendation: [http://www.gradeworkinggroup.org/FAQ/index.htm](http://www.gradeworkinggroup.org/FAQ/index.htm)
DIAGNOSTIC HEALTH TECHNOLOGY

Effectiveness / Accuracy

Compared to current/alternative methods of diagnosis, does the scientific evidence confirm that use of the technology is more accurate? That is, does the use of the technology more accurately identify both those with the condition being evaluated and those without the condition being evaluated? Does the use of the technology result in better sensitivity and better specificity? Or do gains in sensitivity outweigh a reduction in specificity or vice versa such that on balance the diagnostic technology is thought to be more accurate than current diagnostic testing?

If the evidence does not show that use of the technology is more accurate, does the scientific evidence confirm that use of the technology is equally accurate – compared to currently available diagnostic testing? That is, does the use of the technology identify both those with the condition being evaluated and those without the condition being evaluated with accuracy equivalent to current diagnostic testing? Does the use of the technology result in equivalent sensitivity and specificity? Or are gains in sensitivity countered by loss of specificity or vice versa such that on balance the diagnostic technology is thought to be of equivalent accuracy to current diagnostic testing?

Or, finally compared to current/alternative methods of diagnosis, does the scientific evidence show that use of the technology is less accurate?

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Diagnostic Outcome</th>
<th>Level of Confidence Technology is Beneficial**</th>
<th>Level of Confidence Technology is Equivalent**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degenerative Disc Disease</td>
<td>More accurate</td>
<td>□ Not Confident</td>
<td>□ Confident†</td>
</tr>
<tr>
<td></td>
<td>Equally accurate</td>
<td>□ Not Confident</td>
<td>□ Confident†</td>
</tr>
<tr>
<td></td>
<td>Less Accurate</td>
<td>□ Not Confident</td>
<td>□ Not</td>
</tr>
<tr>
<td></td>
<td>Inconclusive</td>
<td>□ Not Confident</td>
<td>□ Not</td>
</tr>
<tr>
<td>Source of Pain</td>
<td>More accurate</td>
<td>□ Not Confident</td>
<td>□ Confident†</td>
</tr>
<tr>
<td></td>
<td>Equally accurate</td>
<td>□ Not Confident</td>
<td>□ Confident†</td>
</tr>
<tr>
<td></td>
<td>Less Accurate</td>
<td>□ Not Confident</td>
<td>□ Not</td>
</tr>
<tr>
<td></td>
<td>Inconclusive</td>
<td>□ Not Confident</td>
<td>□ Not</td>
</tr>
<tr>
<td>Patients that will improve with lumbar fusion</td>
<td>More accurate</td>
<td>□ Not Confident</td>
<td>□ Confident†</td>
</tr>
<tr>
<td></td>
<td>Equally accurate</td>
<td>□ Not Confident</td>
<td>□ Confident†</td>
</tr>
<tr>
<td></td>
<td>Less Accurate</td>
<td>□ Not Confident</td>
<td>□ Not</td>
</tr>
<tr>
<td></td>
<td>Inconclusive</td>
<td>□ Not Confident</td>
<td>□ Not</td>
</tr>
<tr>
<td></td>
<td>More accurate</td>
<td>□ Not Confident</td>
<td>□ Not</td>
</tr>
<tr>
<td></td>
<td>Equally accurate</td>
<td>□ Not Confident</td>
<td>□ Not</td>
</tr>
<tr>
<td></td>
<td>Less Accurate</td>
<td>□ Not Confident</td>
<td>□ Not</td>
</tr>
<tr>
<td></td>
<td>Inconclusive</td>
<td>□ Not Confident</td>
<td>□ Not</td>
</tr>
</tbody>
</table>

*Beneficial – Technology is more accurate
**Equivalent – Technology is equivalent in accuracy
†Confident – Generally supported by moderate or strong evidence
Does the scientific evidence confirm that use of the technology can safely and effectively replace other tests?

<table>
<thead>
<tr>
<th>Test</th>
<th>Can the Technology Replace Other Test?</th>
<th>Level of Confidence Technology can replace other test?</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRI</td>
<td>□ Yes</td>
<td>□ Confident†</td>
</tr>
<tr>
<td></td>
<td>□ No</td>
<td>□ Not Confident</td>
</tr>
<tr>
<td>Plain Radiographs</td>
<td>□ Yes</td>
<td>□ Confident†</td>
</tr>
<tr>
<td></td>
<td>□ No</td>
<td>□ Not Confident</td>
</tr>
</tbody>
</table>

†Confident – Generally supported by moderate or strong evidence

Overall Efficacy: Does scientific evidence confirm that use of the technology results in better health outcomes than management without use of the technology?

- □ Yes
- □ No
- □ Not Studied/No Evidence

- Level of confidence that the evidence confirms that use of the technology results in better health outcomes?
  - □ Not Confident
  - □ Confident
  - □ Not applicable: The evidence does not show that the technology results in better health outcomes.
DIAGNOSTIC HEALTH TECHNOLOGY

Safety

Morbidity

- Does scientific evidence confirm that use of the technology is free of or unlikely to produce significant morbidity? (either directly related to the diagnostic test or long term)?

  Significant morbidity: Frequent adverse effect on health, but unlikely to result in lasting harm or be life-threatening, or:
  Adverse effect on health that can result in lasting harm or can be life-threatening.

### Morbid Outcome

<table>
<thead>
<tr>
<th>Morbid Outcome</th>
<th>Significant Morbidity</th>
<th>Level of Confidence Technology is Safe?*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short term/direct complication: Pain provocation complication</td>
<td>Morbidity Unlikely</td>
<td>Confident †</td>
</tr>
<tr>
<td></td>
<td>Morbidity Likely</td>
<td>Not Confident</td>
</tr>
<tr>
<td></td>
<td>Inconclusive</td>
<td></td>
</tr>
<tr>
<td>Short term/direct complication: ____________________________</td>
<td>Morbidity Unlikely</td>
<td>Confident †</td>
</tr>
<tr>
<td></td>
<td>Morbidity Likely</td>
<td>Not Confident</td>
</tr>
<tr>
<td></td>
<td>Inconclusive</td>
<td></td>
</tr>
<tr>
<td>Short term/direct complication: ____________________________</td>
<td>Morbidity Unlikely</td>
<td>Confident †</td>
</tr>
<tr>
<td></td>
<td>Morbidity Likely</td>
<td>Not Confident</td>
</tr>
<tr>
<td></td>
<td>Inconclusive</td>
<td></td>
</tr>
<tr>
<td>Long term complication: ____________________________</td>
<td>Morbidity Unlikely</td>
<td>Confident †</td>
</tr>
<tr>
<td></td>
<td>Morbidity Likely</td>
<td>Not Confident</td>
</tr>
<tr>
<td></td>
<td>Inconclusive</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Morbidity Unlikely</td>
<td>Confident †</td>
</tr>
<tr>
<td></td>
<td>Morbidity Likely</td>
<td>Not Confident</td>
</tr>
<tr>
<td></td>
<td>Inconclusive</td>
<td></td>
</tr>
</tbody>
</table>

*Safe – significant morbidity is unlikely
†Confident – Generally supported by moderate or strong evidence

- Does scientific evidence confirm that use of the technology is free of or unlikely to produce significant morbidity directly related to the diagnostic test?
  - [ ] Yes
  - [ ] No
  - [ ] Not Studied/No Evidence

- In terms of short term morbidity, level of confidence that the evidence confirms use of the technology is safe:
  - [ ] Not confident
  - [ ] Confident
Mortality

- Does scientific evidence confirm that use of the technology is not likely to increase mortality?
  - Yes
  - No
  - Not Studied / No or Inconclusive Evidence

- In terms of mortality, level of confidence that use of the evidence confirms the technology is safe:
  - Not confident
  - Confident (Generally supported by moderate or strong evidence)

Overall

- Considering short and long term morbidity and mortality, does scientific evidence confirm that use of the technology is safe?
  - Yes
  - No

- Level of confidence that the evidence confirms that use of the technology is safe?
  - Not confident
  - Confident (Generally supported by moderate or strong evidence)
DIAGNOSTIC TEST TECHNOLOGY

Cost Impact

- Are independent cost analyses (cost benefit; cost effectiveness; or other cost analysis) identified?
  - Yes
  - No

  If Yes:
  - Do the cost analyses show that use of the new technology will result in costs that are greater, equivalent or lower than management without use of the technology?
    - Greater
    - Equivalent
    - Lower
    - Not applicable: No independent cost analysis identified

  If No:
  - Does the evidence available to the committee indicate that use of the new technology will result in costs that are greater, equivalent or lower than management without use of the technology?
    - Short term costs (Cost of first year)
      - Greater
      - Equivalent
      - Lower
      - Inconclusive
    - Long term costs (Costs beyond first year)
      - Greater
      - Equivalent
      - Lower
      - Inconclusive
Based on the current level of evidence regarding the technology’s safety and effectiveness relative to currently available diagnostic methods, is use of the technology likely to have a net benefit, an equivalent benefit, less benefit or a net harm?

- Net Benefit
- Equivalent Benefit
- Less Benefit
- Net Harm
- The available evidence does not permit a conclusion

Based on the current level of evidence regarding the technology’s cost impact relative to currently available diagnostic methods, is use of the technology likely to increase cost, result in equivalent cost or reduce cost?

- Increase Cost
- Equivalent Cost
- Lower Cost

Relative to currently available diagnostic methods, into which category does the evidence indicate use of the new technology will fall?

<table>
<thead>
<tr>
<th>Less Benefit Increased Cost</th>
<th>Equivalent Benefit Increased Cost</th>
<th>Net Benefit Increased Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less Benefit Equivalent Cost</td>
<td>Equivalent Benefit Equivalent Cost</td>
<td>Net Benefit Equivalent Cost</td>
</tr>
<tr>
<td>Less Benefit Reduced Cost</td>
<td>Equivalent Benefit Reduced Cost</td>
<td>Net Benefit Reduced Cost</td>
</tr>
</tbody>
</table>
DIAGNOSTIC TEST TECHNOLOGY

Coverage Determination

The HTCC has reviewed and considered the technology assessment and information provided by the administrator, reports and/or testimony from an advisory group, and submissions or comments from the public. The committee has given greatest weight to the evidence it determined, based on objective factors, to be the most valid and reliable.

- Based on the evidence that regarding the technology’s safety, effectiveness, and cost-effectiveness, the use of the technology should be covered?

  □ No. Evidence is insufficient to conclude that the health technology is safe, efficacious, and cost-effective or the evidence is sufficient to conclude that the health technology is unsafe, ineffectual, or not cost-effective

  or

  □ Yes. The evidence is sufficient to conclude that the health technology is safe, efficacious, and cost-effective for all indicated conditions; evidence is sufficient to conclude that

  or

  □ Yes, under certain conditions. Coverage is allowed with special conditions (e.g. population, conditions, timing, adjunct services, qualifications, etc.) because the evidence is sufficient to conclude that the health technology is safe, efficacious, and cost-effective only when:

  _______________________________________________________

  _______________________________________________________

  _______________________________________________________

  _______________________________________________________

□ This determination is consistent with the identified Medicare decisions and expert guidelines.

□ Based on the evidence, this determination is inconsistent with either the identified Medicare decisions or expert guidelines.