SPINE Volume 33, Number 11, pp 1262–1269

An Evidence-Based Medicine Approach in Determining Factors That May Affect Outcome in Lumbar Total Disc Replacement

Michael R. Zindrick, MD,*† Michael N. Tzermiadianos, MD,† Leonard I. Voronov, MD,† Mark Lorenz, MD,* and Alexander Hadjipavlou, MD‡

Study Design. Literature research.

Objective. To analyze the available evidence about a variety of factors that might affect outcome of lumbar artificial disc replacement.

Summary of Background Data. Evaluating the scientific merit of new technology is important for a clinician considering incorporating these techniques. An evidence-based medicine approach can aid in this decision-making process.

Methods. Eleven questions were asked about patient selection issues, surgical accuracy of placement, and evidence that motion preservation alters the natural history of degeneration. Studies where answers were found were ranked according to their level of evidence.

Results. The majority of studies found were level IV, with only limited numbers of higher level studies. Only lower level studies with conflicting results assess the effect on outcomes of single *versus* multilevel surgery, L4–L5 *versus* L5–S1 implantations, patient's age, and history of previous surgery. One lower level study suggests that mild-to-moderate facet degeneration does not influence outcomes. The extent of preoperative facet degeneration that can be accepted remains unclear, as level IV studies report degradation of facet degeneration after implantation. Higher level studies support the importance of surgical precision on clinical outcome and lower level studies give mixed results on the same issue. A level III prognostic study suggests that higher range of motion of the implanted segment may be associated with better outcomes, whereas 2 level IV therapeutic studies provide conflicting results. The incidence of adjacent level degeneration in lower level studies ranges between 17% and 28.6%, and can require additional surgery in 2% to 3% of patients. Two level IV studies suggest that preservation of motion may have a prophylactic effect on adjacent discs.

Conclusion. Existing evidence does not provide definite conclusions in the majority of the questions regarding indications and factors that may affect outcomes. Where feasible, conclusions are mainly drawn from lower level, least reliable evidence. Highest quality data are short-term whereas longer-term data are of lower quality and in many instances conflicting. More high level studies with long-term follow-up are necessary to shed light to important clinical issues.

From the *Hinsdale Orthopaedic Associates, SC, Hinsdale, IL; †Loyola University Medical Center, Maywood, IL; and ‡University of Crete, Heraklion, Greece.

Acknowledgment date: August 10, 2007. Revision date: October 10, 2007. Acceptance date: December 17, 2007.

The device(s)/drug(s) is/are FDA-approved or approved by corresponding national agency for this indication.

No funds were received in support of this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.

Address correspondence and reprint requests to Michael R. Zindrick, MD, Hinsdale Orthopaedic Associates, SC, 550 West Ogden Avenue, Hinsdale, IL 60521; E-mail: Zindrickmike@cs.com

Key words: artificial disc replacement, evidence-based medicine, factors affecting outcome. **Spine 2008;33: 1262–1269**

The role of artificial disc replacement (ADR) in the treatment of spinal disorders still remains unclear. Evaluating the scientific merit of any new technology early in its life cycle is important for a clinician considering incorporating these techniques into practice. An evidence-based medicine (EBM) approach can aid in this decision-making process. The method of EBM integrates the best clinical research evidence with the clinician's experience and the patient values.[1,2] With any new technology, however, the relative value of these 3 key components of EBM becomes corrupted. Physician experience may be minimal or lacking; in this information age, patients' values are distorted as they are being bombarded from every angle. Patients often craving for the latest and the greatest technology are unable to distinguish between marketing hype and true scientific evidence. Under these circumstances, fully understanding strengths or limitations of the remaining component of EBM, the best available evidence assumes added importance. To find the best available evidence, a specified problem faced by the clinician is translated into an answerable question, and a systematic retrieval of evidence from the medical literature is performed.[3] Clinical research studies are classified into level I–IV based on strength of scientific methodology employed in performing the study. The reliability of the answer to a posed question is based on the level of study of which it was derived.

Reports of early implantations in the lumbar spine that as experience progressed the indications became more narrow.[4] Recent studies brought to attention that ADR may not represent an alternative to fusion, showing that 100% of fusion patients had one or more recognized contraindication to ADR.[5,6] Therefore, it is expected that the future growth of ADR will come either from indications for surgery not present today, or from elimination of current contraindications. In an attempt to understand what is the strength of the best evidence available concerning a variety of indications and other factors that might affect the outcomes of ADR, the authors undertook a systematic review of the current lumbar ADR literature. Specific questions about factors that might affect the outcome after ADR were posed. These questions included patient selection issues, surgical accuracy of implantation, and existing evidence

that motion preservation alters the natural history of spinal segmental degeneration. Studies where the answers were found were then reviewed and ranked according to the EBM methodology to answer these important clinically questions.

■ Materials and Methods

The 11 posed questions regarding lumbar ADR were grouped in 3 main categories: (a) Questions about patient selection issues: (i) effect of single level *versus* multilevel implantation, (ii) effect of the level of implantation, (iii) effect of patient's age, (iv) effect of prior surgery, (v) effect of preoperative facet degeneration. (b) Questions about surgical technique issues: (i) effect of prosthesis positioning. (c) Questions about motion technology issues: (i) effect of the implanted segment range of motion (ROM) on outcome, (ii) fate of facets after ADR, (iii) incidence of heterotopic ossification and effect on ROM and outcome, (iv) incidence of adjacent level degeneration, (v) effect of motion preservation on adjacent segment degeneration.

A thorough review of clinical literature between 1990 and May 2007 was carried using electronic and manual search. Only peer-reviewed literature in English language was included. Data that was in abstract form only was not used. Duplicate reports were eliminated, if there was prior studies that presented the same group of patients, the most current report was used. If authors reported a subset of a multicenter study, the largest multicenter series data available was used. Only studies including data addressing the above framed questions were included in this review.

Studies were ranked into appropriate evidence levels using the modification of Sackett grading system provided in the *J Bone Joint Surg Am*, January 2003.[7] Briefly summarized, Level I studies are randomized, controlled clinical trials. Level II are prospective nonrandomized comparative studies. Level III are retrospective comparative studies or case-controlled studies. Level IV includes case series, with no comparison group. Level V evidence, which includes expert opinions, was not included in the present study. Previous reviews on ADR were also not included. Articles were graded according to the type of study (therapeutic, prognostic, *etc.*) and the level of evidence (I–IV) by 2 reviewers. The quality of reports, especially for Level IV studies, was also assessed by their study design, follow-up, and the use of validated outcome measures. The findings from individual studies were aggregated to produce a "bottom line" conclusion.

■ Results

We reviewed 76 clinical papers to find evidence addressing the formulated questions. From those papers, 49 were excluded from this review as they did not include relevant information or were duplicates. The remaining 27 papers were evaluated (Table 1).

Patient Selection Issues

Is the Outcome After Single Segment Implantation Similar to Multisegmental Implantation? Ten level IV studies were found (Table 2). Three studies report inferior results

Table 1. Articles Included in This Study, Ranked by Level of Evidence, Study Design, Follow-up, and Outcome Measures

Author	Level	No. Pts	Study Design	FU	Lost at FU (%)	Outcome Measures
CHARITÉ						
Tortolani, 2007[8]	I Prognostic	276	Prosp	2 yr		Heterotopic ossification
Trouillier, 2006[9]	I Prognostic	13	Prosp	6 mo		Facet subchondral bone density
McAfee, 2005[10]	I Therapeutic	205–99	Prosp	2 yr	8.5	ODI, SF36
Shim, 2007[11]	III Therapeutic	61	Retro	3 yr	6.5	ODI
David, 2007[12]	IV Therapeutic	108	Retro	13.2 yr	2	Nonvalidated
Putzier, 2006[13]	IV Therapeutic	71	Retro	17 yr	25	ODI
Regan, 2005[14]	IV Therapeutic	100	Prosp	6–24 mo		ODI
Lemaire, 2005[4]	IV Therapeutic	107	Retro	11.3 yr	7	Nonvalidated
Van Ooij, 2003[15]	IV Therapeutic	27	Retro	7.5 yr		Nonvalidated
Scott, 2000[16]	IV Therapeutic	14	Retro	18–68 mo	28.50	Nonvalidated
Zeegers, 1999[17]	IV Therapeutic	50	Prosp	2 yr	8	Nonvalidated
Lemaire, 1997[18]	IV Therapeutic	105	Retro	4 yr		Nonvalidated
Cinotti, 1996[19]	IV Therapeutic	46	Retro	3.2 yr		Nonvalidated
ProDisc						
Patel, 2006[20]	I Prognostic	52	Prosp	2 yr		ODI, CT scan
Huang, 2006[21]	III Prognostic	64	Retro	8.7 yr	34	Radiographic review
Huang, 2005[22]	III Prognostic	64	Retro	8.6 y	41	Stauffer-Coventry score, ODI
Siepe, 2007[23]	IV Therapeutic	99	Prosp	2 yr		ODI
Siepe, 2006[24]	IV Therapeutic	94	Prosp	3 yr	2	ODI, SF36
Chung, 2006[25]	IV Therapeutic	38	Prosp	37 mo	5	ODI
Bertagnoli, 2006[26]	IV Therapeutic	22	Prosp	2 yr	0	ODI
Bertagnoli, 2005[27]	IV Therapeutic	118	Prosp	2 yr	12	ODI
Bertagnoli, 2005[28]	IV Therapeutic	29	Prosp	2 yr	14	ODI
Tropiano, 2005[29]	IV Therapeutic	64	Retro	8.7 yr	14	Nonvalidated
Tropiano, 2003[30]	IV Therapeutic	53	Prosp	1.4 yr		ODI
Bertagnoli, 2002[31]	IV Therapeutic	108	Prosp	3 mo–2 y		ODI
Mayer, 2002[32]	IV Therapeutic	34	Prosp	1 yr	23.5	ODI
Maverick						
Le Huec, 2005[33]	IV Therapeutic	64	Prosp	2 yr	0	ODI

ODI indicates Oswestry Disability Index; Prosp, prospective study; Retro, retrospective study.

Table 2. Effect of Number of Levels Implanted in Clinical Outcomes

Author	Level	Study Design	FU	No. Pts	Effect of Multisegmental Implantation on Outcome
CHARITÉ					
Cinotti, 1996[19]	IV	Retro	3.2 yr	1 level: 36 2 levels: 10	Inferior results
Lemaire, 2005[4]	IV	Retro	>10 yr	1 level: 54 2 level: 45	No difference
Zeegers, 1999[17]	IV	Prosp	2 yr	1 level: 29 2 level: 18	No difference
ProDisc					
Siepe, 2007[23]	IV	Prosp	2 yr	1 level: 79 2 level: 20	Inferior results
Chung, 2006[25]	IV	Prosp	2 yr	1 level: 25 2 level: 11	Inferior results
Bertagnoli, 2005[27]	IV	Prosp	2 yr	1 level: 106	No difference
Bertagnoli, 2005[28]	IV	Prosp	2 yr	≥2 levels: 25	
Tropiano, 2005[29]	IV	Prosp	8.7 yr	1 level: 35 ≥2 levels: 20	No difference
Tropiano, 2003[30]	IV	Prosp	1–2 yr	1 level: 40 ≥2 levels: 13	No difference
Mayer, 2002[32]	IV	Prosp	1 yr	1 level: 31 ≥2 level: 3	No difference

with multisegmental implantations, whereas 6 studies report similar results. Bertagnoli *et al* report the results of single level and multilevel implantations in separate studies allowing only indirect comparisons.[27,28]

Does Spinal Level of ADR Affect Outcome? Two prospective, level IV studies were found. Regan *et al*,[14] in a study of 100 patients implanted with CHARITÉ, report no statistical difference in outcome when L4–L5 was compared to L5–S1 at 6 to 24 months of follow-up. Siepe *et al*,[23] in a study of 99 patients with ProDisc II with a mean 2-year follow-up report a trend towards better outcomes at L4–L5 compared with L5–S1.

Does Patient's Age Affect Outcome? Eight level IV studies were found (Table 3). Younger age was a favorable predictive factor in 3 studies,[17,24,33] whereas was a negative factor in 1 study.[29] Patient age did not affect outcome in 4 studies.[16,25,26,30] Some authors report higher complication rate in older patients. Lordosis enhancement after implantation can exacerbate spinal stenosis, and compromised bone quality can increase the risk of subsidence.[26] Therefore, routine open prophylactic vertebroplasty, and posterior decompressive laminectomy in cases with radiographic evidence of spinal stenosis has been suggested by some authors.[26]

Does Prior Surgery Affect Outcome? Twelve level IV studies were found (Table 4). Prior surgery had a negative effect on outcome in 6 studies,[12,19,28–30,33] whereas it had no effect on outcome in 5 studies.[18,24,25,27,32] In 1 study, prior surgery had a negative effect on outcome at 1 year and no effect at 2 years follow-up.[17] Most of the studies use non validated outcome measures.[12,17–19,29]

Does Preoperative Facet Degeneration Affect Outcome? Only 1 level IV study was found. Le Huec *et al*,[33] in a prospective study of 64 Maverick ADR reported that mild or moderate facet osteoarthritis (grade 1 or 2, on the 0–3 Fujiwara scale), did not influence outcome at 2 years follow-up. Patients with severe facet arthrosis had worse outcome, but their number was small to reach conclusive evidence.

Surgical Technique Issues

Does Prosthesis Positioning Affect ROM or Outcome? One level I study and 6 level IV studies were found (Table 5). There is level I evidence that accuracy of placement af-

Table 3. Effect of Patient's Age in Clinical Outcomes

Author	Level	Study Design	FU	No. Pts	Effect of Age on Outcome
CHARITÉ					
Zeegers, 1999[17]	IV	Prosp	2 yr	46	Patients <45 yr had better outcome
Scott, 2000[16]	IV	Retro	4 yr	14	Age >45 does not affect outcome
ProDisc					
Siepe, 2006[24]	IV	Prosp	3 yr	92	Patients <40 yr had better outcome
Tropiano, 2005[29]	IV	Prosp	8.7 yr	55	Patients >45 yr had better outcome
Chung, 2006[25]	IV	Prosp	2 yr	36	Age does not affect outcome
Bertagnoli, 2006[26]	IV	Prosp	2 yr	22	Age does not affect outcome
Tropiano, 2003[30]	IV	Prosp	1.4 yr	53	Age >50 does not affect outcome
Maverick					
Le Huec, 2005[33]	IV	Prosp	2 yr	64	Young patients had better outcome

Table 4. Effect of Prior Surgery on Patients' Outcome

Author	Level	Study Design	FU	No. Pts With (+) or Without (−) Previous Surgery	Effect of Previous Surgery on Outcome
CHARITÉ					
Cinotti, 1996[19]	IV	Retro	3.2 yr	(+): 24 (−): 22	Negative effect
David, 2007[12]	IV	Retro	13.2 yr	(+): 44 (−): 62	Negative effect in patients with >2 previous surgeries
Zeegers, 1999[17]	IV	Prosp	2 yr	(+): 27 (−): 33	Negative effect at 1 yr No effect at 2 yr
Lemaire, 1997[18]	IV	Retro	4 yr	(+): 55 (−): 50	No effect
ProDisc					
Bertagnoli, 2005[28]	IV	Prosp	2 yr	(+): 17 (−): 12	Negative effect
Tropiano, 2005[29]	IV	Prosp	8.7 yr	(+): 28 (−): 27	Negative effect
Tropiano, 2003[30]	IV	Prosp	1.4 yr	(+): 11 (−): 33	90% satisfactory results 97% satisfactory result
Mayer, 2002[32]	IV	Prosp	1 yr	(+): 9 (−): 25	No effect
Bertagnoli, 2005[27]	IV	Prosp	2 yr	(+): 60 (−): 46	No effect
Siepe, 2006[24]	IV	Prosp	3 yr	(+): 17 (−): 75	No effect
Chung, 2006[25]	IV	Prosp	2 yr	(+): 7 (−): 29	No effect
Maverick					
Le Huec, 2005[33]	IV	Prosp	2 yr	64	Negative effect

fects both clinical outcome and ROM after ADR.[10] Data from level IV studies are conflicting; 3 studies report that placement can affect long-term outcome[4,12] or ROM,[19] whereas 4 studies show no effect.[17,20,25,33]

Motion Technology Issues

Does ROM of the Implanted Segment Affect Outcome? One level III and 2 level IV studies were found (Table 6). A level III prognostic study reports that segmental ROM >5° was associated with a statistically significant but clinically modest better clinical outcome and a trend toward improved low back pain scores when compared with ROM ≤5°.[22] Similarly, a level IV prospective study reports that higher segmental motion after implantation was associated with better clinical outcomes.[25] On the contrary, another level IV retrospective study reports that patients with functional implants were significantly less satisfied than those with spontaneous ankylosis.[13]

What Is the Fate of Facets After the Implantation? Two level I, 1 level III, and 3 level IV studies were found (Table 7). Level I studies suggest no facet encumberment, as measured by CT osteoabsorptiometry of subchondral bone density,[9] or facet changes measured on CT examination.[20] However, follow-up in both studies was short, ranging from 6 to 24 months. Level III and level IV studies with longer follow-up suggest progression of facet arthrosis over time. Lemaire et al[4] reported that patients that developed facet arthrosis had nonideal anterior positioning of the prosthesis. Symptoms were developed in 36% of those patients. Prosthesis placement lateral to the ideal midline position was associated with develop-

Table 5. Effect of Implant Positioning on ROM and Clinical Outcome

Author	Level	Study Design	FU	No. Pts	Effect of Placement
CHARITÉ					
McAfee, 2005[10]	I	Prosp	2 yr	276	Affects both outcomes and ROM
David, 2007[12]	IV	Retro	13.2 yr	106	Anterior placement is correlated with the development of symptomatic facet arthrosis
Lemaire, 2005[4]	IV	Retro	10 yr	100	All patients that developed facet arthrosis had nonideal placement
Zeegers, 1999[17]	IV	Prosp	2 yr	50	No effect
Cinotti, 1996[19]	IV	Retro	3.2 yr	46	Affects ROM
ProDisc					
Patel, 2006[20]	IV	Prosp	2 yr	52	No effect
Chung, 2006[25]	IV	Prosp	>2 yr	36	No effect
Maverick					
Le Huec, 2005[33]	IV	Prosp	2 yr	64	No effect if implant was between 0 and 7 mm from the posterior wall

Table 6. Effect of Range of Motion After Implantation on Clinical Outcome

Study	Level	Design	FU	No. Pts	ROM	Effect of ROM on Outcome
CHARITÉ						
Putzier, 2006[13]	IV Therapeutic	Retro	17 yr	53	Functional–mobile implants: 17%	Patients with functional implants were less satisfied
ProDisc						
Huang, 2005[22]	III Prognostic	Retro	8.6 yr	39	ROM >5°: 28%	Better outcomes with ROM >5°
Chung, 2006[25]	IV Therapeutic	Prosp	3 yr	36		Better outcomes with higher ROM

ment of symptoms. David[12] reported that 4.7% of patients required posterior fusion for symptomatic facet arthrosis within 3 to 12 years after implantation. Symptomatic facet arthrosis accounted for 45.4% of index level reoperation. That study also correlates the development of symptomatic facet arthrosis with anterior placement of the prosthesis. Similarly, Van Ooij *et al*,[15] in a series of 27 patients with unsatisfactory results after Charité disc replacement reported a 40.7% incidence of symptomatic facet arthrosis. The mean interval from surgery to facet arthrosis was 4.4 years. Shim *et al*,[11] in a level III comparative study reported no statistical difference of the facet degeneration between patients implanted with Charité and ProDisc.

What Is the Rate of Heterotopic Ossification, and What Is Its Effect on ROM and Clinical Outcome? One level I and 4 level IV studies were found (Table 8). In a prognostic level I study, Tortolani *et al*[8] reported a 4.3% incidence of heterotopic ossification at 2-year follow-up. The presence of heterotopic ossification did not significantly affect ROM or clinical outcome. Five level IV studies were also found. Cinotti *et al*[19] reported a 15.2% incidence of perianular ossifications; and 57% of patients with ossifications had spontaneous interbody fusion. However, perianular ossifications did not affect clinical outcome. David[12] reported partial ossification in 3.8% of patients and complete ossification with spontaneous fusion in 2.8% of patients. Ossifications occurred only in patients treated with postoperative brace and activities restriction, whereas was not noted in patients who had early active physiotherapy.[12] Putzier *et al*[13] reported that 60% of patients had spontaneous fusion and another 13% had signs of possible or likely motion impairment. Patients with functional implants without signs of heterotopic ossification were less satisfied than those with spontaneous ankylosis. Lemaire *et al*[4] reported a 3% incidence of heterotopic ossification, without any cases of spontaneous arthrodesis. However, 9% of patients in that study had ROM <2°, which is beyond measurement error accepted by the FDA.

What Is the Incidence of Adjacent Level Degeneration After ADR? One level III and 6 level IV studies were found (Table 9). Cinotti et al[19] report 0% incidence in 3.2 years of follow-up based on MRIs performed on 10 patients out of the 46 included in the authors' series. However, no selection criteria for the 10 patients were provided. Other studies with more than 3 years of follow-up, report that the incidence of ALD ranges between 17%[13] and 28.6%.[11] Additional surgery was required in 2% to 3% of patients in 2 series[4,12]

What Is the Effect of Motion Preservation on Adjacent Level Degeneration? Only 2 level IV studies were found (Table 10). Data suggest that preservation of motion after ADR may reduce the risk for adjacent level degeneration.[13,21]

■ Discussion

In the ideal EBM model, the best available evidence from the literature is combined with clinical experience and patients' values. When dealing with new technology, there is, however, a lack of physician experience. Patient values may be artificially manipulated and overly optimistic due to marketing and advertising, leading to the misconception that "newest means best." Under such

Table 7. Incidence of Radiographic and Symptomatic Facet Degeneration

Author	Level	Study Design	FU	No. Pts	Radiographic	Symptomatic
CHARITÉ						
Trouillier, 2006[9]	I	Prosp	6 mo	13	No evidence of sclerosis of facet joints measured by CT osteoabsorptiometry	
Shim, 2007[11]	III	Retro	3 yr		36.6%	
David, 2007[12]	IV	Retro	13.2 yr	106		4.7%
Lemaire, 2005[4]	IV	Retro	10 yr	100	11%	4%
Van Ooij, 2003[15]	IV	Retro	7.5 yr	27		40.7% incidence of facet joint arthrosis among patients with unsatisfactory results
ProDisc						
Patel, 2006[20]	I	Prosp	6–24 mo	52	0%	
Shim, 2007[11]	III	Retro	3 yr		32%	

Table 8. Incidence of Heterotopic Ossification (HO) and Its Effect on Range of Motion (ROM) and Clinical Outcome

Study	Level	Design	FU	No. Pts	HO (%)	Effect on ROM	Effect on Outcome
CHARITÉ							
Tortolani, 2007[8]	I	Prosp	2 yr	276	4.3	No effect	No effect
David, 2007[12]	IV	Retro	13.2 yr	106	6.6	Negative	
Putzier, 2006[13]	IV	Retro	17 yr	53	73	Negative	Negative
Lemaire, 2005[4]	IV	Retro	11.3 yr	100	3		
Cinotti, 1996[19]	IV	Retro	3.2 yr	46	15.2	Negative	No effect

circumstances, it becomes even more imperative for a clinician embarking on the use of new technology, to fully understand what "best evidence" exists for newer techniques. This distortion of the related values of the evidence based medicine tripod; physician's experience, best evidence, and patient values are obvious in the early introduction of motion technology. Few physicians have little if any experience with these techniques or devices. Patients have been bombarded by the lay press and manufacturers representatives that artificial discs and other parts replacements of the spine will be the answer to their misery and disability. In attempt to determine what is the level of the best existing evidence for several factors that may affect outcomes, the authors undertook this study. Understanding the strengths and weakness of the available literature can better allow the medical practitioner and the patient make well informed decisions regarding treatment options.

Sorting through medical literature to obtain answers can oftentimes be difficult. Research studies are susceptible to invalid conclusions resulting from bias, confounding, or chance. With the introduction of EBM techniques, however, the medical literature can be sorted into levels of evidence based on scientific merit. This technique offers a useful tool for clinicians trying to find answers regarding patient care. Higher level studies minimize bias, confounding and chance making their conclusions more likely reliable. By the very nature of their design, lower level studies do not address bias, confounding and chance making their results more prone to error. Higher level studies, however, may also have unavoidable methodologic flaws. In the ADR literature, the FDA IDE studies represent the highest quality evidence available.[34,35] Those studies are randomized, controlled, and use validated outcome measures with a minimum of 2 year follow-up. Entry criteria and patient randomization for the studies is generally good. Lacking in all studies, however, is blinding. The reason for lack of blinding can be easily understood; nonetheless, this exerts a bias on outcome and should be considered when weighing their conclusions.

Not surprisingly, the majority of the experimental studies were level IV, with only limited higher level studies. This reflects the difficulties in performing a randomized controlled trial, as well as the reluctance among clinicians and patients to deviate from their concepts of what the optimum treatment should be. In the absence of higher level studies, most of the best evidence concerning ADR comes from level IV studies (Table 1).

Patient Selection Issues

The existing evidence does not provide definite conclusions to most of the questions asked in this study. All data addressing the patient selection issues comes from level IV studies. The available studies evaluating the question of single *versus* multilevel surgery provide conflicting results. Concerning the outcomes of ADR at L4–L5 *versus* L5–S1, 2 studies did not show statistically significant difference; however, 1 reported a trend for better outcomes at L4–L5. The role of patient's age remains unclear, and the results of the available studies are conflicting. Although some studies report similar results in older patients, the possibility of higher complications and the morbidity of additional surgical interventions should also be considered in decision-making. Similar conflicts were found about the effect of prior surgery. The role of preexisting facet arthrosis is still obscure. Preexisting facet arthrosis is currently a contraindication to ADR; however, one study suggests that mild to moderate facet degeneration did not influence ADR outcomes.[33] Clinically significant facet arthrosis is reported to

Table 9. Reported Rate of Adjacent Level Degeneration (ALD) After ADR

Author	Level	Study Design	FU	No. Pts	Radiographic ALD (%)	Surgery for ALD (%)
CHARITÉ						
Shim, 2007[11]	III	Retro	3 yr	33	19.4	
David, 2007[12]	IV	Retro	13.2 yr	106		2.8
Lemaire, 2005[4]	IV	Retro	11.3 yr	100		2
Putzier, 2006[13]	IV	Retro	17 yr	53	17	
Cinotti, 1996[19]	IV	Retro	3.2 yr	10/46	0	
ProDisc						
Shim, 2007[11]	III	Retro	3 yr	24	28.6	
Huang, 2006[21]	IV	Retro	8.7 yr	42	24	
Bertagnoli, 2002[31]	IV	Prosp	3 mo–2 yr	108	9.2	

Table 10. Effect of Motion Preservation on the Incidence of Adjacent Level Degeneration (ALD)

Author	Level	Study Design	FU	No. Pts	Radiographic ALD
CHARITÉ					
Putzier, 2006[13]	IV	Retro	17 yr	53	20% in patients with spontaneous fusion 0% in patients with ROM >3°
ProDisc					
Huang, 2006[21]	IV	Retro	8.7 yr	42	34% in patients with ROM <5° 0% in patients with ROM >5°

be present in 66% of patients undergoing fusion surgery.[35] Nevertheless, the extent of facet degeneration that can accepted in motion preservation surgery remains to be evaluated, as most of the candidates for this surgery are expected to have some degree of facet arthrosis.

Surgical Precision

Higher level studies seem to support the importance of surgical precision on clinical outcome. A level I study reports that the accuracy of placement in the sagittal plane affects both ROM and clinical outcomes. Data from level IV give mixed results. Some studies support that less than ideal placement of the prosthesis is correlated with the development of symptomatic facet arthrosis[4,12] or with decreased ROM,[17] whereas other report that accuracy of placement had no effect.[17,20,25,33]

Motion Preservation Validity

Considering motion technology concept issues, data from a level III prognostic study suggest that higher ROM of the implanted segment may be related with better outcomes. This is supported by a prospective level IV study,[25] whereas contradicted by a retrospective level IV study.[13] Several IV studies report degradation of facet degeneration after the implantation[4,11] Furthermore, the commonest reason for conversion to fusion in long-term follow-up is the development of symptomatic facet arthrosis.[12,15] Although it is theoretically postulated that prosthesis design and constrain may have a significant role in development of facet arthrosis, data from a level III comparative study shows similar rates of facet degradation in a constrained *versus* a semiconstrained device.[11]

One of the main theoretical advantages of disc arthroplasty over spinal fusion is the prevention of the accelerated degeneration of the adjacent segments. However, the incidence of adjacent level degeneration in level IV studies with follow-up more than 3 years ranges between 17%[13] and 28.6%,[11] and can require additional surgery in 2% to 3% of patients.[4,12] In these studies, the surprisingly high incidence of adjacent level degeneration suggests that disc arthroplasty may not have a protective effect on the adjacent segments as initially thought. In contrast, 2 level IV studies with long follow-up suggest that preservation of motion may have a prophylactic effect on adjacent discs.[13,22] Huang *et al*[22] suggested that ROM $\geq 5°$ is a plausible crucial threshold to prevent adjacent level degeneration. The motion data provided to FDA from the IDE of Charité shows that at 24 months after impanation 33% patients had less than 5° of ROM.[36] Since it may take more than a decade for symptomatic junctional degeneration to develop, longer follow-up are necessary to shed more light on the effect of ADR *versus* fusion in randomized prospective trials.

A limitation of the present study is that it includes only published data. The published medical literature may not reflect the most current information such as that presented at medical meetings, it is, however, peer reviewed and thoroughly dissectible by EBM standards. An additional problem is the complexity of patent population that results from trying to fuse different level studies to draw a "bottom line" conclusion. In lower level studies, there are minimal control of study bias, chance, confounding, and thus significant potential effect on the results by these factors. The effects of factors are minimized in level I and level II studies and as hopefully more of these level trails are published increasingly reliable conclusions can be drawn.

■ Conclusion

Existing evidence does not allow drawing definite conclusions in the majority of the clinical questions regarding indications and factors that may affect outcomes. Where feasible, conclusions are mainly drawn from lower level least reliable evidence. Highest quality data are short-term and longer term data are of lower quality and in many instances conflicting. This lower level data, however, is plentiful and often quoted. The clinician must understand when basing important clinical decisions that the scientific ground on which he/she is treating may not be as solid, as one would wish. There exist no long-term studies of high level scientific merit that demonstrate long-term efficacy of motion preservation technology over traditional techniques. Additionally, there exists limited data to suggest or support that junctional breakdown above fusions is clinically altered or different than the normal degenerative process expected over ensuing period of time. There is limited data to suggest that motion technologies prevent the natural progression of degeneration, either at the index level or adjacent segments at this time. However, it is important to clarify that lack of evidence is not synonymous to lack of benefit. High-level studies with long-term follow-up

are necessary to shed more light to important clinical issues.

■ Key Points

- The majority of studies addressing questions regarding indications and factors that may affect outcomes are level IV, with only limited numbers of higher level studies.
- Existing evidence does not provide definite conclusions in many important clinical questions regarding factors that may affect outcomes.
- Where feasible, conclusions are mainly drawn from lower level, least reliable evidence.
- Highest quality data are short-term while longer term data are of lower quality and in many instances conflicting.

References

1. Straus S, Richardson W, Glasziou P, Haynes B. *Evidence Based Medicine*. 3rd ed. London; Elsevier Churchill Livingstone, 2005.
2. Sackett DL, Straus S, Richardson W, et al. *Evidence-Based Medicine. How to Practice and Teach EBM*. New York: Churchill Livingstone, 2000.
3. Dawes M, Summerskill W, Glasziou P, et al. Sicily statement on evidence-based practice. *BMC Med Educ* 2005;5:1.
4. Lemaire JP, Carrier H, Ali el-HS, et al. Clinical and radiological outcomes with the Charité trade mark artificial disc: a 10-year minimum follow-up. *J Spinal Disord Tech* 2005;18:353–9.
5. Huang RC, Lim MR, Girardi FP, Cammisa FP Jr. The prevalence of contraindications to total disc replacement in a cohort of lumbar surgical patients. *Spine* 2004;29:2538–41.
6. Wong DA, Annesser B, Birney T, et al. Incidence of contraindications to total disc arthroplasty: a retrospective review of 100 consecutive fusion patients with a specific analysis of facet arthrosis. *Spine J* 2007;7:5–11.
7. Wright JG, Swiontkowski MF, Heckman JD. Introducing levels of evidence to the Journal. *J Bone Joint Surg Am* 2003;85:1–3.
8. Tortolani PJ, Cunningham BW, Eng M, et al. Prevalence of heterotopic ossification following total disc replacement. A prospective, randomized study of two hundred and seventy-six patients. *J Bone Joint Surg Am* 2007; 89:82.
9. Trouillier H, Kern P, Refior HJ, Muller-Gerbl M. A prospective morphological study of facet joint integrity following intervertebral disc replacement with the CHARITÉ Artificial Disc. *Eur Spine J* 2006;15:174–82.
10. McAfee PC, Cunningham B, Holsapple G, Adams K, et al. A prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of lumbar total disc replacement with the CHARITÉ artificial disc versus lumbar fusion: part II: evaluation of radiographic outcomes and correlation of surgical technique accuracy with clinical outcomes. *Spine* 2005;30:1576–83.
11. Shim CS, Lee SH, Shin HD, et al. Charité versus ProDisc: a comparative study of a minimum 3-year follow-up. *Spine* 2007;32:1012–8.
12. David T. Long-term results of one-level lumbar arthroplasty: minimum 10-year follow-up of the CHARITÉ artificial disc in 106 patients. *Spine* 2007; 32:661–6.
13. Putzier M, Funk JF, Schneider SV, Gross C, et al. Charité total disc replacement-clinical and radiographical results after an average follow-up of 17 years. *Eur Spine J* 2006;15:183–95.
14. Regan JJ. Clinical results of Charité lumbar total disc replacement. *Orthop Clin North Am* 2005;36:323–40.
15. van Ooij A, Oner FC, Verbout AJ. Complications of artificial disc replacement: a report of 27 patients with the SB Charité disc. *J Spinal Disord Tech* 2003;16:369–83.
16. Sott AH, Harrison DJ. Increasing age does not affect good outcome after lumbar disc replacement. *Int Orthop* 2000;24:50–3.
17. Zeegers WS, Bohnen LM, Laaper M, Verhaegen MJ. Artificial disc replacement with the modular type SB Charité III: 2-year results in 50 prospectively studied patients. *Eur Spine J* 1999;8:210–7.
18. Lemaire JP, Skalli W, Lavaste F, et al. Intervertebral disc prosthesis. Results and prospects for the year 2000. *Clin Orthop Relat Res* 1997;337:64–76.
19. Cinotti G, David T, Postacchini F. Results of disc prosthesis after a minimum follow-up period of 2 years. *Spine* 1996;21:995–1000.
20. Patel VV, Andrews C, Pradhan BB, et al. Computed tomography assessment of the accuracy of in vivo placement of artificial discs in the lumbar spine including radiographic and clinical consequences. *Spine* 2006;31:948–53.
21. Huang RC, Tropiano P, Marnay T, et al. Range of motion and adjacent level degeneration after lumbar total disc replacement. *Spine J* 2006;6:242–7.
22. Huang RC, Girardi FP, Cammisa FP Jr, et al. Correlation between range of motion and outcome after lumbar total disc replacement: 8.6-year follow-up. *Spine* 2005;30:1407–11.
23. Siepe CJ, Mayer HM, Heinz-Leisenheimer M, Korge A. Total lumbar disc replacement: different results for different levels. *Spine* 2007;32:782–90.
24. Siepe CJ, Mayer HM, Wiechert K, Korge A. Clinical results of total lumbar disc replacement with ProDisc II: three-year results for different indications. *Spine* 2006;31:1923–32.
25. Chung SS, Lee CS, Kang CS. Lumbar total disc replacement using ProDisc II: a prospective study with a 2-year minimum follow-up. *J Spinal Disord Tech* 2006;19:411–5.
26. Bertagnoli R, Yue JJ, Nanieva R, et al. Lumbar total disc arthroplasty in patients older than 60 years of age: a prospective study of the ProDisc prosthesis with 2-year minimum follow-up period. *J Neurosurg Spine* 2006;4: 85–90.
27. Bertagnoli R, Yue JJ, Shah RV, et al. The treatment of disabling single-level lumbar discogenic low back pain with total disc arthroplasty utilizing the Prodisc prosthesis: a prospective study with 2-year minimum follow-up. *Spine* 2005;30:2230–6.
28. Bertagnoli R, Yue JJ, Shah RV, et al. The treatment of disabling multilevel lumbar discogenic low back pain with total disc arthroplasty utilizing the ProDisc prosthesis: a prospective study with 2-year minimum follow-up. *Spine* 2005;30:2192–9.
29. Tropiano P, Huang RC, Girardi FP, et al. Lumbar total disc replacement. Seven to eleven-year follow-up. *J Bone Joint Surg Am* 2005;87:490–6.
30. Tropiano P, Huang RC, Girardi FP, Marnay T. Lumbar disc replacement: preliminary results with ProDisc II after a minimum follow-up period of 1 year. *J Spinal Disord Tech* 2003;16:362–8.
31. Bertagnoli R, Kumar S. Indications for full prosthetic disc arthroplasty: a correlation of clinical outcome against a variety of indications. *Eur Spine J* 2002;11(Suppl 2):S131–6.
32. Mayer HM, Wiechert K, Korge A, Qose I. Minimally invasive total disc replacement: surgical technique and preliminary clinical results. *Eur Spine J* 2002;11(Suppl 2):S124–30.
33. Le Huec JC, Basso Y, Aunoble S, et al. Influence of facet and posterior muscle degeneration on clinical results of lumbar total disc replacement: two-year follow-up. *J Spinal Disord Tech* 2005;18:219–23.
34. Blumenthal S, McAfee PC, Guyer RD, et al. 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. *Spine* 2005;30:1565–75.
35. Zigler J, Delamarter R, Spivak JM, et al. Results of the prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of the ProDisc-L total disc replacement versus circumferential fusion for the treatment of 1-level degenerative disc disease. *Spine* 2007;32: 1155–63.
36. Mirza SK. Point of view: commentary on the research reports that led to food and drug administration approval of an artificial disc. *Spine* 2005;30: 1561–4.