Appendix A:
General Methodological Principles of Study Design

When making national coverage determinations, CMS evaluates relevant clinical evidence to determine whether or not the evidence is of sufficient quality to support a finding that an item or service falling within a benefit category is reasonable and necessary for the diagnosis or treatment of illness or injury or to improve the functioning of a malformed body member. The critical appraisal of the evidence enables us to determine whether: 1) the specific assessment questions can be answered conclusively; and 2) the intervention will improve health outcomes for patients. An improved health outcome is one of several considerations in determining whether an item or service is reasonable and necessary.

CMS divides the assessment of clinical evidence into three stages: 1) the quality of the individual studies; 2) the relevance of findings from individual studies to the Medicare population; and 3) overarching conclusions that can be drawn from the body of the evidence on the direction and magnitude of the intervention's risks and benefits.

The issues presented here represent a broad discussion of the issues we consider when reviewing clinical evidence. However, it should be noted that each coverage determination has unique methodological aspects.

1. Assessing Individual Studies

Methodologists have developed criteria to determine weaknesses and strengths of clinical research. Strength of evidence generally refers to: 1) the scientific validity underlying study findings regarding causal relationships between health care interventions and health outcomes; and 2) the reduction of bias. In general, some of the methodological attributes associated with stronger evidence include those listed below:

- Use of randomization (allocation of patients to either intervention or control group) in order to minimize bias.
- Use of contemporaneous control groups (rather than historical controls) in order to ensure comparability between the intervention and control groups.
- Prospective (rather than retrospective) studies to ensure a more thorough and systematical assessment of factors related to outcomes.
- Larger sample sizes in studies to help ensure adequate numbers of patients are enrolled to demonstrate both statistically significant as well as clinically significant outcomes that can be extrapolated to the Medicare population. Sample size should be large enough to make chance an unlikely explanation for what was found.

- Masking (blinding) to ensure patients and investigators do not know to which group patients were assigned (intervention or control). This is important especially in subjective outcomes, such as pain or quality of life, where enthusiasm and psychological factors may lead to an improved perceived outcome by either the patient or assessor.

Regardless of whether the design of a study is a randomized controlled trial, a non-randomized controlled trial, a cohort study or a case-control study, the primary criterion for methodological strength or quality is the extent to which differences between intervention and control groups can be attributed to the intervention studied. This is known as internal validity. Various types of bias can undermine internal validity. These include:

- Different characteristics between patients participating and those theoretically eligible for study but not participating (selection bias)
- Co-interventions or provision of care apart from the intervention under evaluation (confounding)
- Differential assessment of outcome (detection bias)
- Occurrence and reporting of patients who do not complete the study (attrition bias)

In principle, rankings of research design have been based on the ability of each study design category to minimize these biases. A randomized controlled trial minimizes systematic bias (in theory) by selecting a sample of participants from a particular population and allocating them randomly to the intervention and control groups. Thus, randomized controlled studies have been typically assigned the greatest strength, followed by non-randomized clinical trials and controlled observational studies. The following is a representative list of study designs (some of which have alternative names) ranked from most to least methodologically rigorous in their potential ability to minimize systematic bias:

- Randomized controlled trials
- Non-randomized controlled trials
- Prospective cohort studies
- Retrospective case control studies
- Cross-sectional studies
- Surveillance studies (e.g., using registries or surveys)
- Consecutive case series
- Single case reports

When there are merely associations but not causal relationships between a study's variables and outcomes, it is important not to draw causal inferences. Confounding refers to independent variables that systematically vary with the causal variable. This distorts

measurement of the outcome of interest because its effect size is mixed with the effects of other extraneous factors. For observational, and in some cases randomized controlled trials, the method in which confounding factors are handled (either through stratification or appropriate statistical modeling) are of particular concern. For example, in order to interpret and generalize conclusions to our population of Medicare patients, it may be necessary for studies to match or stratify their intervention and control groups by patient age or co-morbidities.

Methodological strength is, therefore, a multidimensional concept that relates to the design, implementation and analysis of a clinical study. In addition, thorough documentation of the conduct of the research, particularly study's selection criteria, rate of attrition and process for data collection, is essential for CMS to adequately assess the evidence.

2. Generalizability of Clinical Evidence to the Medicare Population

The applicability of the results of a study to other populations, settings, treatment regimens, and outcomes assessed is known as external validity. Even well-designed and well-conducted trials may not supply the evidence needed if the results of a study are not applicable to the Medicare population. Evidence that provides accurate information about a population or setting not well represented in the Medicare program would be considered but would suffer from limited generalizability.

The extent to which the results of a trial are applicable to other circumstances is often a matter of judgment that depends on specific study characteristics, primarily the patient population studied (age, sex, severity of disease, and presence of co-morbidities) and the care setting (primary to tertiary level of care, as well as the experience and specialization of the care provider). Additional relevant variables are treatment regimens (dosage, timing, and route of administration), co-interventions or concomitant therapies, and type of outcome and length of follow-up.

The level of care and the experience of the providers in the study are other crucial elements in assessing a study's external validity. Trial participants in an academic medical center may receive more or different attention than is typically available in non-tertiary settings. For example, an investigator's lengthy and detailed explanations of the potential benefits of the intervention and/or the use of new equipment provided to the academic center by the study sponsor may raise doubts about the applicability of study findings to community practice.

Given the evidence available in the research literature, some degree of generalization about an intervention's potential benefits and harms is invariably required in making coverage decisions for the Medicare population. Conditions that assist us in making reasonable generalizations are biologic plausibility, similarities between the populations studied and Medicare patients (age, sex,

ethnicity and clinical presentation), and similarities of the intervention studied to those that would be routinely available in community practice.

A study's selected outcomes are an important consideration in generalizing available clinical evidence to Medicare coverage determinations because one of the goals of our determination process is to assess health outcomes. We are interested in the results of changed patient management not just altered management. These outcomes include resultant risks and benefits such as increased or decreased morbidity and mortality. In order to make this determination, it is often necessary to evaluate whether the strength of the evidence is adequate to draw conclusions about the direction and magnitude of each individual outcome relevant to the intervention under study. In addition, it is important that an intervention's benefits are clinically significant and durable, rather than marginal or short-lived.

If key health outcomes have not been studied or the direction of clinical effect is inconclusive, we may also evaluate the strength and adequacy of indirect evidence linking intermediate or surrogate outcomes to our outcomes of interest.

3. Assessing the Relative Magnitude of Risks and Benefits

Generally, an intervention is not reasonable and necessary if its risks outweigh its benefits. Health outcomes are one of several considerations in determining whether an item or service is reasonable and necessary. For most determinations, CMS evaluates whether reported benefits translate into improved health outcomes. CMS places greater emphasis on health outcomes actually experienced by patients, such as quality of life, functional status, duration of disability, morbidity and mortality, and less emphasis on outcomes that patients do not directly experience, such as intermediate outcomes, surrogate outcomes, and laboratory or radiographic responses. The direction, magnitude, and consistency of the risks and benefits across studies are also important considerations. Based on the analysis of the strength of the evidence, CMS assesses the relative magnitude of an intervention or technology's benefits and risk of harm to Medicare beneficiaries.

Appendix B
CMS Review Table for Reconsideration Lumbar Artificial Disc Replacement

Author/ Year	Study Design	Demographics	Intervention, outcome measures; instruments	Results		Methodological Comments
				Intervention group	Control group	

Author/ Year	Study Design	Demographics	Intervention, outcome measures; instruments	Results		Methodological Comments
				Intervention group	Control group	
	Study, inclusion/exclusion	N, age, sex,				
Bertagnoli/ 2005a	Single level study 24 month case series follow-up. Exclusions included spinal stenosis, osteoporosis, prior fusion surgery, chronic infections, metal allergies, facet arthrosis, more than one level of spondylosis, neuromuscular disease, inadequate vertebral endplate size, worker's comp, spinal litigation, BMI > 35, any isthmic or degenerative spondylolisthesis greater than Grade 1.	N = 104 Median age 47 (range36-60 years), 24 months follow-up Single level implant L3 to S1.	VAS, ODI, medications use, complications, back and radicular pain	**Pre-op** **Post-op** ODI 54 29 VAS 7.5 3. No or occasional leg pain 57% 92% No or occasional back pain 15% 91% Full time work 10% 35% Part time work 3% 24% Post-op 10% on narcotics, no device related complications, several approach related complications.	None	Single surgeon, single center

Author/ Year	Study Design	Demographics	Intervention, outcome measures; instruments	Results		Methodological Comments
				Intervention group	Control group	
Bertagnoli/ 2005b	Multilevel disease Follow-up time 25 – 41 months. Exclusions included circumferential spinal stenosis, osteoporosis, prior fusion surgery, chronic infections, metal allergies, facet joint arthrosis, inadequate vertebral endplate size, neuromuscular disease, worker's compensation, spinal litigation, BMI > 35, isthmic or degenerative spondylolisthesis greater than Grade 1.	N = 25 Mean age: 49.6 (male), 47.7 (female). Age range 34-60. 15 two level implants, 10 three level implants.	VAS pain, ODI, back and leg pain, medication usage, complications	Pre-op Post-op ODI 65.0 21.6 VAS 8.3 2.1 No or episodic leg pain; 48% 100% No or episodic back pain; 8% 92% Post-op 4% on regular narcotics, one case of subsidence, one case of polyethylene extrusion.	None	Single surgeon, single center

Author/ Year	Study Design	Demographics	Intervention, outcome measures; instruments	Results		Methodological Comments
				Intervention group	Control group	
Bertagnoli/ 2006a	Symptomatic adjacent-segment degeneration after remote lumbar fusion. 24 month case series follow-up. Exclusions included circumferential spinal stenosis, osteoporosis, chronic infections, metal allergies, facet joint arthrosis, inadequate vertebral endplate size, workers' compensation, spinal litigation, BMI > 35, isthmic or degenerative spondylolisthesis grater than Grade 1.	N = 20 Age range 18 to 67. 8 cases 2 level, 2 cases 3 level.	VAS, ODI, presence of back and leg pain, patient satisfaction scores, complications	**Pre-op** **Post-op** ODI 65.4 29.9 VAS 7.7 3.4 Presence of leg pain; 50% 0 Presence of back pain; 75% 25% Full time work; 13% 27% Part time work; 23% 38% Post-op 0% on narcotics, no device or approach related complications.	None	Single surgeon, single center

Author/ Year	Study Design	Demographics	Intervention, outcome measures; instruments	Results		Methodological Comments
				Intervention group	Control group	
Bertagnoli/ 2006b	Smokers versus nonsmokers. 24 month case series follow-up. Failed 9 months conservative care. Exclusions included spinal stenosis, osteoporosis, prior fusion surgery, chronic infections, metal allergies, facet joint arthrosis, inadequate vertebral endplate size, more than one level of spondylosis, neuromuscular disease, worker's compensation, spinal litigation, BMI > 35, isthmic or degenerative spondylolisthesis greater than Grade 1.	N= 104 Single level L4 to S1. 70 nonsmokers, 24 smokers. Average age smokers 45.5, average age nonsmokers 49.5.	ODI, VAS pain, % leg pain, work rates	**Nonsmokers** Pre-op Post-op ODI 52 32 VAS 7.5 3.8 Presence of leg pain; 48.6% 9% Full time work 11.6% 30.9% Part time work; 4.3% 36.8% **Smokers** Pre-op Post-op ODI 55 28 VAS 7.5 4.5 Presence of leg pain 50% 16% Full time work 6% 50% Part time work 3% 23.5%	None	Single surgeon, single center

Author/ Year	Study Design	Demographics	Intervention, outcome measures; instruments	Results		Methodological Comments
				Intervention group	Control group	
Bertagnoli/ 2006c	Patients older than 60 years. 24 month case series follow-up. Exclusions included spinal stenosis with neurogenic claudication, history of fusion, T score on DEXA $\leq$ 2.5, chronic infections, metal allergies, facet joint arthrosis, inadequate vertebral endplate size, worker's compensation, spinal litigation, BMI > 35, isthmic or degenerative spondylolisthesis grater than Grade 1.	N = 22 Median age of 63 years (range 61-71 years) 17 single level, 4 - two-level, one - three level.	VAS pain, ODI (raw score)	Pre-op Post-op ODI 27 14 VAS 8 4 Post-op 0% on narcotics, 2 cases of foot drop, one case of loss of proprioception and vibration sensation, two cases of subsidence.	None	Single surgeon, single center Performs prophylactic vertebroplasty

Author/ Year	Study Design	Demographics	Intervention, outcome measures; instruments	Results		Methodological Comments
				Intervention group	Control group	
Chung/ 2006	Mean 3 year follow-up Inclusion: Age between 18 and 60, failed 6 months conservative care, minimum disc height of 4mm, minimum ODI of 40, no more than 2 levels from L3 to S1. Exclusion: scoliosis, spondylolysis, sponsylolithesis, severe facet degeneration, and BMD DEXA less than -2.5. Need confirmatory positive discogram.	N = 36 25 patients single level, 11 patients - 2 levels.	VAS leg and back pain, ODI	Pre-op Post-op ODI 69.2 21 VAS leg 4.7 1.2 VAS back 7.5 3 2 approach related complications (major vein injury).	None	Single surgeon
Marnay/ 2002	7 – 11 year case series follow-up Failed conservative care.	N = 64 1, 2, and 3 level implants.	VAS, ODI, satisfaction, complications, Beaujon score	93% satisfied or entirely satisfied Pre-op Post-op VAS back 8.5 3.0 VAS leg 7.1 1.9 ODI NR 8.3 Beaujon 7/20 16/20 (relative gain 1.69). NR = not reported	None	Abstract

Author/ Year	Study Design	Demographics	Intervention, outcome measures; instruments	Results		Methodological Comments
				Intervention group	Control group	
Mayer/ 2002	12 months case series follow-up Exclusions included spondylolisthesis, spinal stenosis, significant osteoarthritis of facet joints, deformities.	N = 34 Average age 44.	VAS, ODI, operative time, blood loss	Pre-op Post-op VAS 6.3 3.4 ODI 19.1 7.2 blood loss was 117 ml per level; op time 130.9 minutes.	None	
Schroven/ 2006	Nonrandomized ProDisc and fusion. 12 month follow-up Inclusion: 6 months conservative therapy, age 18 to 60, CT or MRI.	24 patients (14 ProDisc), mean age 44 (range 29 to 60).	ODI	**Pro Disc** Pre-op Post-op ODI 38.4 12.5 Complications included subsidence, facet arthritis at 6 months, transient sciatica in 2 patients.	**Fusion** Pre-op Post-op 38 21.4 Complications: intraoperative hemorrhage.	

Author/ Year	Study Design	Demographics	Intervention, outcome measures; instruments	Results		Methodological Comments
				Intervention group	Control group	
Siepe/ 2006	24 month minimum follow-up. Indications: 1. DDD 2. DDD with soft disc herniation 3. osteochondrosis following discectomy 4. DDD with Modic changes. Exclusion: central or lateral spinal stenosis, facet joint problems, spondylolysis/spondylolisthesis, spinal instability, major deformity, metabolic bone disease, previous operation with scarring and radiculopathy, irregular endplate shape, previous infection, metal allergy, spinal tumor, post-traumatic segments.	N = 92 Average age 42.5 (range, 21.9 to 66.1 years). 35.9% female, 64.1% male.	VAS pain, ODI	Pre-op Post-op ODI 39.9 18.9 VAS 7 2.8 Complication rate of 19.6%.	None	

Author/ Year	Study Design	Demographics	Intervention, outcome measures; instruments	Results		Methodological Comments
				Intervention group	Control group	
Tropiano/ 2003	12 months case series follow-up Exclusions included facet degeneration, osteoporosis or osteopenia, structural spinal deformities, absence of posterior elements, known chronic disease of major organ system.	N = 53 Mean age 45 years (range 28 – 67 years).	VAS (back and leg pain), ODI, patient satisfaction, return to work	Pre-op Post-op ODI 56 14 VAS lumbar 7.4 1.3 VAS leg pain 6.7 1.9 100% patient satisfied or entirely satisfied, full resumption of work and ADLs in 72%. Operative time average 104 minutes (range 32 – 250 minutes); mean hospital stay 9 days (range 4 – 31 days); 5 patients with complications.	None	
Tropiano/ 2005	7 – 11 year case series follow-up Minimum of 6 months nonoperative treatment prior to procedure. Exclusions included facet arthrosis, central or lateral recess stenosis, osteoporosis, sagittal or coronal plane deformity, absence of posterior elements.	N = 55 (64 initially) Average age 46 (range 25 to 65).	Stauffer-Coventry score (0-20 points), impairment (0-3 points), lower limb pain (0-3 points), low-back pain (0-3 points)	Pre-op post-op LBP* 2.73 1.35 LLP** 2.42 0.67 Impairment 2.02 0.78 Stauffer-Coventry: 7.04 16.1 5 patients with approach related complications. *Low Back Pain **Lower Limb Pain	None	7 patients had both disc replacement and an adjacent fusion during the same operation

Author/ Year	Study Design	Demographics	Intervention, outcome measures; instruments	Results		Methodological Comments
				Intervention group	Control group	
Zigler/ 2007, Delamarter/ 2005, Zigler/ 2003, Zigler, Burd/ 2003, and Zigler/ 2004	RCT Multiple inclusion and exclusion criteria	N = 286 Mean age: fusion 40.4, ProDisc 38.7. Caucasian: Fusion 78.7% ProDisc 82.6%	Overall success comprised of 10 endpoints	Overall Success ProDisc Sponsor criteria - 63.5%. FDA criteria - 53.4% All adverse events 84%. All device related adverse events 17.9%. Device failures 3.7%.	Overall Success Fusion: Sponsor criteria 45.1% FDA criteria: 40.8% All adverse events 87.5%. All device related adverse events 20%. Device failures 2.7% .	Delamarter/ 2005 Zigler/ 2003, Zigler, Burd/ 2003 and Zigler/ 2004 are early reports of the RCT clinical tria

Medicare National Coverage Determinations Manual

Chapter 1, Part 2 (Sections 90 – 160.26)

Coverage Determinations

Table of Contents

(Rev.)

150.10 - Lumbar Artificial Disc Replacement (LADR) *(Effective August 14, 2007)*

150.10 - Lumbar Artificial Disc Replacement (LADR) (Effective *August 14, 2007*)

(Rev., Issued: , Effective: *August 14, 2007*, Implementation: *October 1, 2007*)

A. General

The LADR is a surgical procedure on the lumbar spine that involves complete removal of the damaged or diseased lumbar intervertebral disc and implantation of an artificial disc. The procedure may be done as an alternative to lumbar spinal fusion and is intended to reduce pain, increase movement at the site of surgery and restore intervertebral disc height. The FDA has approved the use of the lumbar artificial disc for spine arthroplasty in skeletally mature patients with degenerative or discogenic disc disease at one level for L3 to S1.

B. Nationally Covered Indications

N/A

C. Nationally Non-Covered Indications

Effective for services performed *from* May 16, 2006 *through August 13, 2007*, the Centers for Medicare and Medicaid Services *(CMS*) has found that LADR with the Charite™ lumbar artificial disc is not reasonable and necessary for the Medicare population over 60 years of age; therefore, LADR with the Charite™ lumbar artificial disc is non-covered for Medicare beneficiaries over 60 years of age.

Effective for services performed on or after August 14, 2007, CMS has found that LADR is not reasonable and necessary for the Medicare population over 60 years of age; therefore, LADR is non-covered for Medicare beneficiaries over 60 years of age.

D. Other

For Medicare beneficiaries 60 years of age and younger, there is no national coverage determination for LADR, leaving such determinations to continue to be made by the local contractors.

For dates of service May 16, 2006 through August 13, 2007, Medicare coverage under the investigational device exemption (IDE) for LADR with a disc other than the Charite™ lumbar disc in eligible clinical trials is not impacted.

(This NCD last reviewed *August 2007*.)
