Guidance for Industry and/or
FDA Reviewers/Staff

Guidance Document for the Preparation of IDEs for Spinal Systems

Document issued on: January 13, 2000

This document supersedes document the preliminary background document for the future development of a spinal IDE guidance dated 8/26/98.



U.S. Department of Health and Human Services
Food and Drug Administration
Center for Devices and Radiological Health

Orthopedic Devices Branch
Division of General and Restorative Devices
Office of Device Evaluation

GRZAN

Preface

Public Comment:

Comments and suggestions may be submitted at any time for Agency consideration to the Orthopedic Devices Branch, HFZ-410, 9200 Corporate Blvd., Rockville, MD 20850. Comments may not be acted upon by the Agency until the document is next revised or updated. For questions regarding the use or interpretation of this guidance contact, contact Samie Allen at 301-594-3090 or by electronic mail at SXN@cdrh.fda.gov.

Additional Copies:

World Wide Web/CDRH home page at http://www.fda.gov/cdrh/ode/87.pdf or CDRH Facts on Demand at 1-800-899-0381 or 301-827-0111, specify number 2250 when prompted for the document shelf number.

GRZAN

Guidance[1] Document for the Preparation of IDEs for Spinal Systems

INTRODUCTION 3

GENERAL IDE CONTENT INFORMATION 4

REPORT OF PRIOR INVESTIGATIONS 5

1. Clinical Data 5

2. Animal Data 5

3. Mechanical Data 6

4. Biocompatibility 7

5. Bibliography 8

INVESTIGATIONAL PLAN 8

1. Feasibility/Pilot vs. Pivotal Study 8

2. Purpose / Objective Statement 9

3. Study Design 9

3.1 Concurrent Control 10

3.2 Literature Control 10

4. Inclusion Criteria 11

4.1 Lumbar Degenerative Disc Disease (DDD) 12

4.2 Scoliosis 12

4.3 Fractures 12

4.4 Spondylolisthesis 13

4.5 Revision 13

4.6 Cervical DDD 13

5. Exclusion Criteria 13

6. Number of Sites / Investigators / Patients 14

7. Duration / Follow-up Schedule 14

8. Effectiveness Evaluation 15

8.1 Radiographic Success 15

8.1.1 Fusion Status 15

8.1.2 Radiographic Success for Nonfusion Systems 16

8.2 Pain and Function 17

8.3 Height Assessment 18

8.3.1 Disc Height Assessment 18

8.3.2 Vertebral Height Assessment 18

8.4 Health Related Quality of Life 18

9. Safety Evaluation 18

9.1 Subsequent Surgical Interventions 19

9.2 Neurological Status 19

10. Patient / Study Success 20

[1]This document is intended to provide guidance. It represents the Agency's current thinking on this topic. It does not create or confer any rights for or on any person and does not operate to bind FDA or the public. An alternative approach may be used if such approach satisfies the requirements of the applicable statute, regulations, or both.

10.1 Patient Success 20
10.2 Study Success 20

11. Statistical Analyses / Data Presentations 21
11.1 Statistical Analyses 21
11.2 Data Presentations 22

12. Patient Report Forms 23

13. Risk Analysis 23

14. Post-Operative Regimen 23

15. Retrieval Study 24

16. Spinal Tumors 24
16.1 Inclusion Criteria 24
16.2 Duration / Follow-up Schedule 24
16.3 Sample Size / Number of Sites & Investigators 24
16.4 Controls 24
16.5 Clinical Parameters 25
16.6 Safety 25
16.7 Patient / Study Success 25
16.8 Data Presentations / Statistical Analyses 26

DEVICE DESCRIPTION 26

MANUFACTURING 27

MONITORING 27

IRB and INVESTIGATOR INFORMATION 27

LABELING 27

1. Package Labels 28

2. Package Insert 28

3. Surgical Technique Manual 28

INFORMED CONSENT 29

SALES 29

ENVIRONMENTAL IMPACT ASSESSMENT 29

[1]This document is intended to provide guidance. It represents the Agency's current thinking on this topic. It does not create or confer any rights for or on any person and does not operate to bind FDA or the public. An alternative approach may be used if such approach satisfies the requirements of the applicable statute, regulations, or both.

INTRODUCTION

The purpose of this document is to provide guidance on important preclinical and clinical information, which should be presented in an Investigational Device Exemption (IDE) application for spinal systems. This guidance document was prepared in response to discussions and correspondence between the Orthopedic Devices Branch (ORDB) and sponsors of spinal systems.

Additionally, on October 8, 1998, ORDB presented a preliminary background document that involved the ideas pertaining to the development of IDEs for spinal systems to the Orthopedic and Rehabilitation Devices Panel. During the October 8, 1998 Panel meeting, ORDB received input from the Panel members and from the public regarding that preliminary background document and revised it accordingly. Now, this guidance document replaces that preliminary background document.

The suggestions and recommendations written in this document reflect methodologies which have been determined to be acceptable and which, if followed, should help to produce IDE applications with scientifically valid data. This guidance document suggests some important evaluation criteria, test procedures, and endpoints. If the same objectives can be achieved by other means, the sponsor should not refrain from exploring alternative approaches. This guidance document should be viewed as a "living" document. As science changes and scientific techniques are improved, CDRH will periodically revise the document.

This IDE guidance document is applicable to most types of spinal systems including pedicle screw systems, intervertebral body fusion devices (e.g., cages), vertebral body replacement devices, vertebral disc replacements, etc. In this guidance, a spinal "system" is defined here as the complete implant configuration. A "component" is a single element in a system. "Construct" references are typically made when discussing testing. For the purposes of this guidance, "system" and "device" are used interchangeable.

Most systems consist of multiple components (e.g., pedicle screw systems, cages with endcaps, cages with biological components, etc.). However, a system may also consist of a single component (e.g., 1-piece vertebral body replacement device). For systems intended for fusion, which the majority are, the systems are typically used with a bone material (e.g., autograft, allograft, etc.).

FDA regards all implanted spinal systems as significant risk devices. Prior to marketing these systems, a sponsor must obtain clearance of a premarket notification (510(k)) or approval of a premarket approval (PMA) application. In order to support the marketing application of specific types or designs of spinal systems, clinical data may be necessary, and before initiating a clinical trial in the U.S., the sponsor must have an IDE approved by the FDA. Institutional review board (IRB) approval alone is not sufficient to commence a clinical trial on human subjects involving an implanted spinal system. Alternatively, in lieu of an IDE and PMA, a sponsor may submit a product development protocol (PDP). Any sponsor considering this option should refer to *the Guidance for Industry: Contents of a Product Development Protocol* for specific input regarding PDP applications. The suggestions and guidance discussed below are appropriate for evaluating the product in either case.

This guidance document provides a framework to assist in developing clinical studies and generating valid scientific evidence which should provide reasonable assurance of the safety and effectiveness of a spinal system for a stated intended use and indication. This guidance document serves as a supplement to other FDA publications on IDE applications and should not be construed as a replacement for these documents. ORDB would like to be consulted early in the development of preclinical tests and the investigational protocol. Although use of this document to prepare preclinical and clinical protocols will not ensure IDE, PMA, or PDP approval or 510(k) clearance, following this guidance should reduce unnecessary work on the part of sponsors and should allow for a more efficient review by FDA.

Section 205 of the Food and Drug Administration Modernization Act of 1997 (FDAMA) requires FDA, in consultation with the sponsor, to consider the "least burdensome" means that will allow appropriate premarket development and review of a device without unnecessary delays and expense to the sponsor. This requirement to consider the least burdensome means applies to 510(k)s, PMAs, and PDPs. FDA is committed to working with the sponsor to meet this requirement and is open to alternative means of developing an adequate protocol. Please refer to the guidance "Evidence Models for the Least Burdensome Means to Market" which is available at http://www.fda.gov/cdrh/ode/1154.pdf for details.

FDA regulations referred to in this guidance can be found in the Code of Federal Regulations, Title 21, as follows:

- Protection of Human Subjects; Informed Consent (21 CFR 50)
- Standards for Institutional Review Boards for Clinical Investigations (21 CFR 56)
- Good Laboratory Practice (GLP) Regulations (21 CFR 58)
- Investigational Device Exemptions (21 CFR 812)
- Determination of Safety and Effectiveness (defines valid scientific evidence) (21 CFR 860.7)
- Environmental Impact Considerations (21 CFR 25)

All FDA publications referred to in this guidance document can be obtained by contacting the Division of Small Manufacturers Assistance (DSMA) at 800-638-2041 (toll free) or 301-443-6597. Some of the publications can be obtained via DSMA's Internet site at www.fda.gov/cdrh/dsma/dsmamain.html. Specific questions and clarification regarding this guidance document should be directed to ORDB at 301-594-2036.

GENERAL IDE CONTENT INFORMATION

Sponsors of clinical investigations should carefully consider how to adequately support the intended uses/indications for safety and effectiveness of their specific system for a 510(k), PMA, or PDP. Studies should be designed to ensure that the data will provide valid scientific evidence (as defined in 21 CFR 860.7) which will satisfactorily answer all appropriate questions and support the intended uses/indications.

Any IDE should include the following major sections:

(1) Report of Prior Investigations;
(2) Investigational Plan;
(3) Device Description;
(4) Manufacturing;
(5) Monitoring;
(6) IRB and Investigator Information;
(7) Labeling;
(8) Informed Consent;
(9) Sales; and
(10) Environmental Impact Assessment.

What follows is a description of the specific type of information that should be submitted for each of the 10 major sections identified above for spinal systems. However, some of the sections (e.g., Manufacturing, Sales, Environmental Impact Assessment) include generic information that applies to any IDE.

REPORT OF PRIOR INVESTIGATIONS

The following information should be supplied to support the conclusion that patients in the trial will not be placed at undue risk. Complete reports of prior clinical, animal, and laboratory (e.g., mechanical) testing, as well as a comprehensive summary of the testing, should be provided to justify the proposed study.

1. Clinical Data

   A report of any available clinical data regarding the use of the subject system or its components, whether adverse or supportive, that are relevant to the evaluation of the safety or effectiveness of the subject system should be included.

   A brief summary may be sufficient for spinal system types with which there is published experience (e.g., pedicle screw systems, intervertebral body fusion devices with autogenous bone, etc.). However, for new designs of devices with little published experience (e.g., vertebral body replacements, vertebral disc replacements, intervertebral body fusion devices with bone morphogenic proteins, etc.), a more detailed report of any relevant clinical data should be provided. Although it is recognized that this level of detail may not be available for all studies, the type of information that should be included in a more detailed report is as follows (items 3-7 are specific to spinal IDEs):

   (1) a description of the study, including design and patient groups;
   (2) patient selection criteria;
   (3) time course distributions of patients: theoretically due, expected (theoretically due minus deaths and failures), actually evaluated, deaths, removals, revisions, reoperations, and supplemental fixations;
   (4) a description of the scales for the primary evaluation parameters and overall success;
   (5) time course distributions of the primary evaluation parameters, overall success, and complications;
   (6) a description of any patient who underwent a subsequent surgical intervention (e.g., removal, revision, reoperation, and/or supplemental fixation); and
   (7) a description of how the clinical data are relevant and may be used to support the relative safety of the proposed design and/or patient population.

2. Animal Data

   Animal data are sometimes needed to establish the relative safety of the subject system prior to initiating a human clinical trial. Reasons for animal studies include addressing biologic response to particle and substrate materials, dosing studies, functional studies, etc.

   Complete reports of any animal studies conducted on the subject system or components of it, whether adverse or supportive, that are relevant to the evaluation of the safety or effectiveness of the subject system should be included. The animal report should specify the purpose of the study and provide supporting pathological, histological, and radiological evaluations. In addition, the report should describe any differences between the version of the system used in the animal studies and that proposed in the IDE.

   The following are some examples of questions that animal studies might be used to address:

   - For an intervertebral body fusion device used in conjunction with a biological component (e.g., bone morphogenic protein), animal studies may provide information as to whether bony fusion occurred or is progressing and the quantity and quality of bone formed inside the device (a stress protected environment) as compared to an appropriate control group. Animal studies are used to evaluate dilution, dose, and concentration factors for certain systems.

For any system that includes a biological or drug component (e.g., bone morphogenic protein, bovine protein, etc.), the sponsor should follow any applicable DGRD guidances regarding the use of biological components with spinal systems. Sponsors are also encouraged to address any applicable guidances from the Center for Biological Evaluation and Research (CBER) and Center for Drug Evaluation and Research (CDER).

- Animal studies may be used to evaluate a novel system design in earlier development stages (e.g., different design concepts of vertebral disc replacement devices).

- Animal studies may be used to support the biocompatibility of a totally new material, a new material for use in orthopedics, or a new material for spinal exposure. Additionally, these studies may be used to address the probability of generating abrasion or wear particles of the new materials.

Although animal studies may be necessary to address specific questions prior to human use, we recognize that choosing and validating an animal model is difficult because there is no perfect animal model. Many animal studies for spinal systems have involved goats, monkeys, mini-pigs, or kangaroos. In choosing an animal model, the sponsor should consider the morphology, histology, biomechanics, and kinetics as compared to the human situation.

3. Mechanical Data

All spinal systems require some mechanical testing or an acceptable rationale addressing why testing is not necessary in order to establish the relative safety of the subject system. The specific requirements for mechanical testing are influenced by the design, material, method of attachment to the spine, and patient indication.

Complete reports of any mechanical testing conducted on the subject system or its components, whether adverse or supportive, that are relevant to the evaluation of the safety or effectiveness of the subject system should be included. A comprehensive summary of all mechanical testing should be included in addition to complete reports of each test. The following elements, at minimum, should be included as part of each test report: identification of the components that comprised the constructs or subconstructs tested; the set-up; the procedures; rationale that testing involved the worst case design, material, and/or manufacturing-related processing; rationale for the loading modes chosen (axial, bending, torsional, shear, etc.); the results; and a discussion of the results in terms of the expected in-vivo and clinical performance of the system. Unless adequate rationale is provided, all testing should involve the worst case construct design of the total system and not testing of individual components. When there are differences between the proposed system and the system tested, an explanation of how or why the results are relevant in establishing the relative safety of the proposed system should be provided.

- Fatigue - The fatigue testing should involve a minimum of six samples of the worst case construct to generate a stress (load) versus number of cycles (S/N) curve that characterizes the asymptotic endurance limit (e.g., a minimum of two samples per load level with one load level reaching a runout value of five million cycles) compared to a appropriate control device. Rationale for the components chosen as worst case should be provided. The interconnection mechanisms/systems may be tested in the same set of constructs or each in a separate set of constructs. Each interconnection mechanism should be tested or adequate rationale be provided. For spinal systems that are intended to maintain disc mobility and not fuse, the fatigue testing should be performed out to a minimum runout of 10 million cycles to support long term implantation. Additionally, testing should be performed out to a minimum runout of 10 million cycles for intervertebral body replacement devices intended for tumor patients because these

patients may present a great difficulty in achieving fusion, and, therefore, the device is acting more as a stabilizer.

- Static - The testing should involve a minimum of five samples of the worst case construct. As with the fatigue testing, the components tested and the loading mode should be justified.

Examples of the types of construct testing typically performed for a given type of spinal system in order to establish relative safety are as follows:

- For lumbar and thoracic pedicle screw systems that are intended for fusion, static and fatigue bending testing should be provided (e.g., ASTM F1717).

- For cervical, pedicle or lateral mass systems intended for fusion, static and fatigue testing should be provided. The loading mode (torsional or bending) is dependent on the design and material.

- For intervertebral body fusion devices, static, fatigue, and expulsion (push-out testing) testing should be provided. The loading mode (axial, torsional, bending, and/or shear) is dependent on the design, material, and levels of use.

- For vertebral body replacement devices, static and fatigue testing in bending and torsional loading modes should be provided.

- For vertebral disc replacements, static and fatigue tests in multiple loading modes should be provided.

Depending on the design of the system, the sponsor may need to perform different tests in lieu of those identified above, perform additional tests in different testing modes, provide testing on individual components of the subject system, etc.

While there is a voluntary testing standard available for pedicle screw systems (i.e., ASTM F1717), many sponsors used modified versions to address different types of spinal systems. Because there may be testing standards in development, sponsors should contact appropriate standards bodies (e.g., ASTM, ISO) for information regarding test set-ups, parameters, etc. for their specific device type.

A statement as to whether all nonclinical tests comply with the GLP regulations (21 CFR 58) should be provided. Otherwise, a brief statement of the reason for the noncompliance should be provided.

4. Biocompatibility

Biocompatibility testing may be necessary based on the material(s) used to comprise the system. AAMI/ANSI 10993-1 is a recognized standard that a sponsor may refer to for a description of what type of information should be provided to address this issue. Additionally, for any biological or drug component (e.g., bone morphogenic protein, bovine protein, etc.), the sponsor should follow any applicable guidances from DGRD and CBER.

5. Bibliography

A bibliography of all published and unpublished information, whether adverse or supportive, that are relevant to justifying the initiation of an IDE should be provided. In addition, copies of this information and a comprehensive summary of it should be provided. This summary should clearly identify which literature are used to provide supporting information for the proposed indications, study duration, evaluation parameter scales, success definitions, system design, testing results, etc.

Examples of where the literature may be used to provide supporting information or address questions are as follows:

- Can the proposed system dimensionally fit the spinal levels proposed?
- Was the rationale for the unique design features appropriate?
- Was the type of mechanical testing performed appropriate?
- Were the results of the mechanical testing adequate?
- What are appropriate study duration, parameter scales, success criteria, etc. for metastatic tumor patients? For other indications? For specific spinal levels (i.e., cervical vs. lumbar, etc.)?
- Are the inclusion and exclusion criteria appropriate for the proposed indication?
- Are the proposed scales and success definitions validated or appropriate?

INVESTIGATONAL PLAN

The clinical study should be designed and conducted in a manner such that it provides data that will constitute valid scientific evidence within the meaning of 21 CFR 860.7.

1. Feasibility/Pilot vs. Pivotal Study

Unlike IDEs for most orthopedic implants, IDEs for spinal systems often involve the introduction of new device designs and investigational protocols. Therefore, protocols for spinal systems may vary in scope from a feasibility/pilot study to a study used to support the safety and effectiveness of a spinal system (which will be referred to in this guidance as a "pivotal" study). These various types of studies are intended to address different questions and collect different types and amounts of safety and effectiveness information.

The type of study (feasibility/pilot or pivotal) to be performed has an impact on many sections of the IDE (e.g., Report of Prior Investigations, Investigational Plan, Informed Consent, Labeling, etc.). All subsections of the protocol should be discussed specific to the type of protocol being proposed. The sponsor should focus the IDE on one type of protocol at a time. If a feasibility/pilot study is being proposed, then the sponsor should include criteria that should be met prior to expansion to a critical study.

- A feasibility/pilot study is used to collect preliminary safety and/or effectiveness information in a very limited human clinical trial before conducting a study to support the safety and effectiveness of a spinal system. This type of study is typically recommended for spinal device designs and investigational protocols where no data are available or no devices with similar designs are available. If data are already available or a number of similar devices already exist, then no feasibility/pilot study may be necessary. The scope and objective of this type of study are limited. For example, this type of study may be performed to optimize the surgical technique, evaluate a new system design, validate new patient assessment tools, identify clinically meaningful endpoints, or to obtain information on which to base the design of a critical study. The purpose of a feasibility/pilot study may also involve the specialized training of investigators for implantation of a novel device and/or implantation via a novel approach. Depending on the purpose of the feasibility/pilot study, a control may be necessary. Even though it is not possible

to obtain statistical significance from such small studies, a limited evaluation of the safety and effectiveness data should be made. Sponsors should refer to the *Guidance on the Review of Investigational Device Exemptions (IDE) Applications for Feasibility Studies*, IDE Guidance Memorandum No. 89-1.

- A pivotal study for marketing purposes should be proposed when adequate preliminary safety and/or effectiveness information has been provided to justify it. Preliminary information may come from feasibility/pilot studies, animal studies, or other preclinical studies.

2. Purpose / Objective Statement

The clinical protocol should begin with clearly defined objective(s) and hypothesis(es). There should be an overall statement of the purpose/objective of conducting the study (e.g., to evaluate the safety and effectiveness of the system in the treatment of a specific condition as compared to the control system). In addition, the purpose should include a precise, medically accepted definition of the condition to be treated and a scientifically sound rationale for the proposed clinical study. The null and alternative hypotheses for the proposed study should be stated in terms of the specific study endpoints, outcomes, and parameters used to measure the success/failure of the system. The study should then be designed to test these hypotheses.

3. Study Design

A study design should be clearly described so as to collect data from relevant patient populations effectively, to measure and compare appropriate clinical and radiographic assessments, and to reduce the introduction of bias. The choice of study design is guided by, among other things, the following:

- types of alternative treatments or devices for the proposed intended use;
- the adequacy of the information available regarding these alternative treatments or devices;
- the availability of patients for the proposed patient population; and
- the ease of statistical data comparisons.

A historical-based study is one approach. This may involve the use of literature or retrospective data. There are situations that FDA acknowledges that historical-based controls are appropriate, such as in the case of a limited patient cohort based on indication. However, the use of historical-based based controls has many inherent problems and, therefore, many disadvantages. For example, frequently, the cohort of spinal literature does not present data completely in terms of desired study outcomes, relevant time points, description of intended use, parameter scale definitions, success criteria definitions, etc. Additionally, statistically relevant comparisons, which are necessary to support the marketing application, can be difficult to make. Therefore, in some cases, prospectively acquired data may best answer questions of safety and effectiveness.

An alternative study design is a nonrandomized, concurrently controlled study. This study design offers the benefits of prospectively acquired data, and, therefore, allows for tighter control of all parameters. However, the data that are acquired, by design, are inherently more biased than the randomized study design. For example, there may be differences in patient population, surgeon experience, etc. between the investigational and control groups. These biases should be addressed when interpreting the data.

The use of randomized concurrently controlled studies provides many advantages over other types of study designs by offering a tight control of all parameters and by addressing some of the biases introduced by the other study designs.

Randomized concurrently controlled studies, nonrandomized concurrently controlled studies, or historical-based studies may be proposed so long as the study design choice effectively addresses the safety and effectiveness of the spinal system and the inherent biases. Again, FDA is committed to meeting the requirement of considering the "least burdensome" means of protocol development. Regardless of the type of control to be incorporated into the protocol, a complete description of the investigational and control groups should be provided. The sponsor should also provide a rationale for the proposed study design, including how inherent biases are to be addressed.

3.1 Concurrent Control

If selecting to use a concurrent control, the control group selected for study of an investigational device should involve a PMA approved or 510(k) cleared system labeled for this indication or a medically acceptable method of treatment for the indicated disease process. The control group and investigational group should, at a minimum, share the same clinical endpoints. However, the control and investigational groups do not necessarily need to be identical as far as method of treatment.

Several approaches may be employed when selecting an appropriate control:

(1) The control treatment may resemble the investigational system in technique, patient population, expected adverse events, and expected clinical and radiographic outcome. For example, if a sponsor were investigating an intervertebral body fusion device for lumbar DDD, the control group, too, would be treated with interbody fusion.

(2) The control treatment may resemble the investigational system in patient population and expected clinical and radiographic outcome. For example, if lumbar DDD is to be treated by spinal fusion, there may be comparisons of interbody fusion to posterolateral fusion or comparisons of laparoscopic interbody fusion techniques to open interbody fusion surgical techniques. This study design allows for adequate comparisons to be made assuming that anticipated differences in technique and outcome are appropriately accounted for.

(3) The control treatment may resemble the investigational device in patient population and expected clinical outcome only. For example, for the treatment of lumbar DDD, spinal surgery may be randomized against a non-operative treatment of any sort. It is important to bear in mind that treatments that pose greater risk to patients should inherently provide some greater benefit to the patient cohort. The degree of invasiveness and profile of adverse events associated with a treatment are important features to consider when designing a study. A study that finds that conservative, non-operative management of a disease process is as successful as an invasive surgical treatment will probably not support product approval for the surgical treatment but may support approval for the less invasive option. Again, all anticipated differences between treatment groups should be accounted for when determining study success.

The method used to assign patients to the investigational or control group should be specified. Patients may be entered into the study consecutively within a pre-specified time period stated in the protocol, as they present to the investigator, or selectively. If patients are chosen selectively, which, as a method inherently introduces bias, the study should involve a randomized, concurrent control design. If the groups are not to be randomized, selective registration is discouraged.

3.2 Literature Control

If selecting to use a literature control, the literature to support a study should, ideally, provide comprehensive information on the standard of care used to treat a particular condition.

However, references from the literature are frequently incomplete in their reporting of relevant clinical parameters, such as indications for use, surgical techniques, and methods of evaluation. Therefore, comparison to the literature may be problematic with regard to specific endpoints. True comparison to a standard of care is best accomplished prospectively due to the limitations of the literature. However, if the study involves literature controls, the sponsor should provide adequate information describing the proposed literature controls that, at a minimum, should include the following:

(1) The proposed literature controls and copies of the articles should be provided. The methods employed for identifying this body of literature should be clearly stated, e.g., MEDLINE search under keyword X and/or Y. Once the entire cohort of literature is identified, the sponsor should present criteria by which literature was excluded from use as a control. Appropriate keywords for search include indications for use and the specific surgical technique/outcome anticipated, e.g., intervertebral disc replacement for DDD.

(2) A table should be provided which includes the following information for each indication: identification of the reference; system(s)/treatment(s) used; diagnostic criteria; length of follow-up; definitions of the fusion or other surgical outcome; results of appropriate primary and secondary evaluation parameters with definitions of the scales used. If converted scales/scores were presented, a description for the basis for the conversion should be provided. Addition, the number of patients studied should be specified for each element at each follow-up in the table.

An adequate number of appropriate control articles for a specific indication in order to support the safety and effectiveness of the subject device should be provided. FDA recognizes that it may be difficult to extract the information from each reference if multiple indications are involved. If that is the case, then the sponsor should summarize the article, identifying all indications involved and stratifying out any information possible by indication.

(3) A separate table should be provided which includes the complications from each of the articles used to generate the table above. A suggested format is to have columns entitled as "complication, article #1, article #2, ..., literature range." Please provide the number of patients with a particular complication, total number of patients and the corresponding patient rate (%) at each follow-up time. These tables should also include subsequent surgical interventions (i.e., removals, revisions, reoperations, and supplemental fixations) at each follow-up time. FDA recognizes that some of the individual complications may also be reflected in the subsequent surgical interventions (e.g., a dural tear that led to a reoperation); however, this information should still be summarized in one table.

4. Inclusion Criteria

Complete inclusion criteria are essential to adequately define the patient group to be investigated. The criteria for inclusion into any clinical study of a spinal system will differ depending on the disease population targeted for the proposed treatment and the location of the disease process (i.e., cervical, thoracic, lumbar).

Regardless of the indication being investigated for a particular spinal system, the following general inclusion criteria should be considered:

- range of patient ages (skeletally mature, if applicable)
- spinal levels involved (e.g., C2-C7; L2-S1, etc.)
- limit on number of adjacent number of levels to be implanted (e.g., 1-level, 2-level, etc.)
- clinical and radiographic conditions for patient entry (e.g., preoperative pain score, location of pain – back, leg, arm, neck, preoperative function score, preoperative neurological score,

radiographic evidence of …, etc.)
- description with suggested time frame of any unsuccessful, non-operative or conservative treatment (e.g., physical therapy, bracing, traction, medication trials)
- description of any restrictions regarding prior nonfusion surgeries
- patient must understand and sign the informed consent
- patient is able to meet the proposed follow-up schedule
- patient is able to follow the postoperative management program

If a sponsor is designing the inclusion criteria to reflect a minimum level of symptoms secondary to disease, the sponsor should consider that systems that require the greatest degree of intervention/convalescence and have the most morbidity should demonstrate an appropriate degree of clinical and radiographic improvement.

For the purposes of this guidance, patients with spinal tumors who are participating in a clinical trial are assumed to be treated with a vertebral body replacement. The tumor type and life expectancy of these patients are variable and unknown *a priori*. As a result, special conditions apply to the design of clinical trials for vertebral body replacements for the treatment of spinal tumors. Information specific to tumors metastatic to the spine is described in detail in Section 16 below.

4.1 Lumbar Degenerative Disc Disease (DDD)

Many protocols involve the investigation of systems for the treatment of lumbar DDD. DDD should be based on patient history and radiographic studies. FDA suggests that the sponsor consider the following:

DDD should be defined as back and/or radicular pain with degeneration of the disc as confirmed by patient history, physical examination, and radiographic studies with 1 or more of the following factors (as measured radiographically, either by CT, MRI, plain film, myelography, discography, etc.):
- instability as defined by $\geq$3mm translation or $\geq$5° angulation;
- osteophyte formation of facet joints or vertebral endplates;
- decreased disc height, on average by >2mm, but dependent upon the spinal level;
- scarring/thickening of ligamentum flavum, annulus fibrosis, or facet joint capsule;
- herniated nucleus pulposus;
- facet joint degeneration/changes; and/or
- vacuum phenomenon.

4.2 Scoliosis

The sponsor is encouraged to provide a discussion of the scoliotic degrees and nature of curvature at time of surgery for which the device is indicated. Discussion should be provided regarding the long-term goals of treatment and additional therapies envisioned. Special attention should be given to age range of the patient cohort so as to appropriately study all populations that would benefit. For example, pediatric use should be considered for the treatment of scoliosis.

4.3 Fractures secondary to trauma

In designing a study for this indication, special attention should be given to the type of fracture/degree of instability, level of fracture, percent retropulsion into the spinal cord, and the number of levels involved. The sponsor is encouraged to provide rationale for any data pooling that is pertinent to the determination of sample sizes for study.

4.4 Spondylolisthesis

When studying spondylolisthesis, special attention should be given to the grade of disease, that is to say, the amount of slippage. The type of spondylolisthesis, degenerative or lytic, and the location of any defect should be collected. At the time of PMA submission, the sponsor may need to demonstrate the poolability of patients across both types of spondylolisthesis.

4.5 Revision surgery for pseudoarthrosis

It is recognized that patients with failed prior fusion attempts are particularly difficult to treat. The sponsor is encouraged to collect information regarding type of prior surgeries and treatment modalities (e.g., bone growth stimulators). At the time of PMA submission, the sponsor may need to demonstrate the poolability of patients across different numbers and types of prior fusion surgeries.

4.6 Cervical DDD

The aforementioned criteria for lumbar DDD should be modified to reflect the anatomy and symptomatology reflective of the cervical spine. Additionally, the disease should be confined to one or two adjacent levels for study consistency. Rationale for the proposed anatomical sites should be provided given the differences in spinal anatomy within the cervical spine. DDD in the cervical spine should be based on both radiographic and clinical assessments. FDA suggests that the sponsor consider the following:

Radiographically (at least one):
- degenerated disc on MRI;
- decreased disc height on plain film, CT or MRI; and/or
- disc herniation, as demonstrated by CT or MRI.

Clinically: radicular symptoms (at least one):
- arm/shoulder pain;
- decreased reflexes;
- decreased strength; and/or
- decreased sensation.

5. Exclusion Criteria

Complete exclusion criteria are essential in adequately defining the patient group to be investigated. This is because exclusion criteria may address a safety concern associated with a specific type of patient and/or allow for the exclusion of patients who may negatively impact the study results and data analyses. Below is a list of exclusion criteria that are common to many spinal studies. If a sponsor chooses to investigate patients with one or more of these conditions, then they should address the impact on the patient safety, study design, and data analyses.

Exclusion criteria are used to help define the patient population in terms of the following categories: safety concerns; follow-up concerns; those that simplify and clarify the study design, and a combination of any or all of above.

Safety concerns:
- has either a systemic infection or infection at the site of surgery
- has osteopenia, osteoporosis, or osteomalacia to a degree that spinal instrumentation would be contraindicated

- has a disease of bone metabolism
- has an allergy to any component of the investigational device
- in the event that a biologic material is to be implanted, has a history of severe allergy or anaphylaxis
- is pregnant or is interested in becoming pregnant during the duration of the study

Follow-up concerns:
- has a history of substance abuse (recreational drugs, alcohol). The sponsor should supply criteria that define "abuse"
- is a prisoner

Those that simplify/clarify study design:
- has had prior fusion attempts at the involved levels (except if studying pseudoarthrosis patients)
- has a condition or requires postoperative medications that may interfere with bony/soft tissue healing (including tobacco use)
- has an additional spinal condition other than the condition to be studied
- is currently involved in a study of another investigational product for similar purpose
- demonstrates 3 or more "Waddell's Signs of Inorganic Behavior"
- is a worker's compensation case

Combination of any or all of above:
- has a known malignancy
- has a concurrent disease process that would place the patient in excessive risk to surgery. These disease processes should be outlined by the sponsor.
- has a disease process that would preclude accurate evaluation (e.g., neuromuscular disease, psychiatric disease, etc.). The sponsor should specify these based upon the endpoints of this study.
- is involved in spinal litigation
- is obese, as defined by a standard measure (Metropolitan tables, body mass index)

6. Number of Sites / Investigators / Patients

The proposed number of investigators, investigational sites, and patients should be specified. The number of patients per treatment group should be determined using sample size calculations based upon the study's null and alternative hypotheses; see section 11 below. Inadequate and/or unequal patient numbers per investigator may decrease the probability that the patients for a given investigator will be representative of the study population. This may preclude the pooling of data between investigators and/or sites. Therefore, FDA recommends that the sponsor consider addressing inadequate or unequal patient distribution in the study. This may be accomplished by considering a minimum number of patients per site to justify statistically pooling of a multicenter study (e.g., $\geq$25 patients per indication and per treatment group enrolled at each site).

7. Duration / Follow-up Schedule

In order to properly assess all safety and primary effectiveness outcomes, not just fusion, a spinal study should involve a minimum of 2 years of follow-up data. However, a sponsor may propose a shorter study with an adequate rationale.

The evaluation time points should be specified. Evaluations are commonly performed at the following time intervals: preoperatively (within 2 months of surgery) to collect background and demographic information and baseline clinical and radiographic data; operatively to collect surgical and safety data; postoperatively at hospital discharge to collect safety data, and postoperatively at appropriate intervals out to 2 years to collect clinical, radiographic, and safety data. FDA also believes that each patient

should be followed annually until the last patient entered has had his/her 2-year follow-up evaluation. Additionally, for each evaluation time point, an appropriate window should be defined. FDA suggests the following evaluation schedule (e.g., baseline, 6 +/- 2 weeks, 3 months +/- 2 weeks, 6 +/- 1 months, 12 +/- 2 months, and 24 +/- 2 months and then annually). Additional evaluation time points may be necessary based on the design and purpose of the specific spinal system.

8. Effectiveness Evaluation

There are primary and secondary evaluation parameters for all spinal studies that should be measured at each timepoint. The specific parameter scales and methods of interpretation (success/failure criteria) with rationale/validation should be included. As a note, neurological assessments are discussed in the section 9 below.

For lumbar spinal studies, primary evaluation parameters should include: *radiographic success (fusion or nonfusion, depending on spinal system type); back and leg pain; and function*. Secondary evaluation parameters should include: *donor site pain; disc height assessment; and health-related quality of life*. As a note, disc height assessment is not applicable for vertebral body replacement devices. Instead, the evaluation parameter is *vertebral height assessment.*

For cervical spinal studies, primary evaluation parameters should include: *radiographic success (fusion or nonfusion, depending on spinal system type); neck and arm pain; and function*. Secondary evaluation parameters should include*: donor site pain; disc height assessment; and health-related quality of life*. As noted above, disc height assessment is not applicable for vertebral body replacement devices. Instead, the evaluation parameter is *vertebral height assessment.*

The success criteria for each of the individual primary evaluation parameters will differ based upon the design of the system, the patient population, and the goals of the treatment.

8.1 Radiographic Success

8.1.1 Fusion Status

For systems intended for fusion (e.g., pedicle screw systems, intervertebral body fusion devices, etc.), there are numerous radiographic methods that may be used to evaluate fusion status (e.g., A/P, lateral, flexion and extension, etc.).

These radiographic evaluations should be performed at the baseline and at all the postoperative evaluations starting at 3 months at each evaluation time point for the duration of the study. All radiographs used in the determination of fusion should be read by at least two radiologists. At least one of the two radiologists should be an independent radiologist who is masked (when possible) to the treatment to address the potential for observer's bias. For the radiographs for which the two radiologists are in disagreement, a third independent, masked (when possible) radiologist should be brought in to look at those radiographs. The third radiological readings should be the deciding evaluation (i.e., majority rules). Keep in mind that there should be no discussion of the radiograph until all readings are made. In addition, the discrepancies between the first two radiological readings should be quantified statistically, if possible.

For double level involvement, both levels should be fused to be considered a successful fusion. When two intervertebral body fusion devices are implanted per level, both devices should be fused to be considered a successful fusion. For a combination system (e.g., system comprised of intervertebral body fusion device and pedicle screw system), both sites should be fused for a patient to be considered a successful fusion.

For lumbar spinal systems, successful fusion should be based on all of the following radiographic endpoints being demonstrated on roentgenographic examination (A/P, lateral, flexion and extension):

- evidence of bridging trabecular bone between the involved motion segments (e.g., between the involved vertebral endplates for intervertebral body replacement devices, between the facets, pedicle, and/or transverse processes for pedicle screw systems);
- translational motion <3mm; and
- angular motion <5°.

The tolerance for presence of radiolucent lines may be dependent on the type of spinal system. If the sponsor chooses to incorporate radiolucency data as a fusion criterion, then specific information should be provided in order to better assure that adequate radiological assessments are being made to measure the accurate presence or absence of radiolucencies. Even if the sponsor does not choose to incorporate the radiolucency data as a fusion criterion, this data should be collected as confirmatory information.

FDA appreciates that the presence of rigid hardware, such as pedicle screw systems, may confound the results of a flexion/extension series. The definition of fusion, therefore, should include the criteria listed above so as to capture movement information as well as bony growth.

For cervical spinal systems, successful fusion should take into consideration the types of radiographic techniques and criteria suggested for lumbar spinal systems above. The sponsor should provide the rationale for any proposed fusion criteria, particularly for the quantitative translational and angular values considered appropriate for the cervical spine.

FDA acknowledges that the radiographic assessment of many of the novel spinal systems is confounded by the presence of opacifying hardware. Therefore, it is tempting to turn to other imaging modalities to assess the area in question. Regarding the use of other radiographic modalities to demonstrate fusion (CT scan, MRI, etc.), the sponsor should demonstrate the validity and reliability of these modalities prior to using these measures as primary study endpoints. FDA generally expects sponsors to use a more traditional method of assessing the fusion status in addition to any proposed method that has not yet been validated.

8.1.2 Radiographic Success for Nonfusion Systems

For spinal systems intended maintain motion without fusion, the radiographic endpoints should demonstrate maintenance of "normal" types and ranges of motion. For example, the following endpoints may be considered to define radiographic success for the lumbar spine for nonfusion systems:

- no evidence of bridging trabecular bone; and
- evidence of range of motion appropriate for level involved. The amount of motion should be specified as part of the success criteria.

As described in Section 8.1.1 above for fusion systems, a sponsor should consider incorporating radiolucency data as part of the fusion criteria or using it as confirmatory information.

8.2 Pain and Function

Although pain and function are two separate endpoints, there is considerable overlap in their assessments. Pain may not be present at rest, but only with activity, causing enough discomfort to limit a patient's ability to function. The patient may have severe enough discomfort with activity as to avoid it altogether and, therefore, cause a limitation in function. There are assessment instruments that successfully capture the effect of pain on function. There are also assessments that successfully capture pain and function independently. For any assessment instrument used, the method of administration should be identified.

For the independent pain assessment, measurements should address both pain severity and frequency. Pain should be measured by using a validated pain assessment instrument, such as a visual analog scale (VAS), 5-point to 10-point scales, or another appropriate instrument. Back and leg (sciatica) pain should be considered for lumbar disease, and neck and arm (radicular) pain should be elicited for cervical disease. Patients who undergo autologous grafting procedures should also be evaluated for postoperative donor site pain. These areas of pain should be assessed both pre- and postoperatively at specific follow-up times.

For the independent function assessment, the sponsor is advised that VAS, X-point scales, and any such uni-dimensional assessment tools, do not fully capture functional status as a whole. Activity-specific function questions capture this endpoint more comprehensively, and, therefore, the use of single question function assessment is discouraged. FDA encourages the use of pain/function specific evaluation instruments.

For assessments that capture both low back pain and function in lumbar spine studies, there are various instruments such as the Oswestry Disability Questionnaire and the Roland-Morris Disability Scale (see Fairbanks JC, et al., "The Oswestry Low Back Pain Disability Questionnaire" *Physiotherapy*. 1980 Aug; 66(8): 271-271). Other back-related function scales include the Million Questionnaire, the Waddell Disability Questionnaire, and the Quebec Disability Questionnaire.

For assessments that capture both neck pain and function in cervical spine studies, FDA recommends that the sponsor consider the Neck Disability Index and Northwick Park Neck Pain Questionnaire, which are both based upon the Oswestry Questionnaire format (see Vernon H, et al., "The Neck Disability Index: A Study of Reliability and Validity." *J Manipulative Physiol Ther*. 1991; 14(7): 409-415, and see Leak A.M., et al., "The Northwick Park Neck Pain Questionnaire, Devised to Measure Neck Pain and Disability, *British Journal of Rheumatology* 1994; 33:469-474).

Sponsors should keep in mind that if a combination pain/function assessment tool is used as part of a study, then the protocol should also include independent assessments of leg and donor site pain for lumbar studies or arm and donor site pain for cervical studies.

Success criteria for the primary pain parameters (back/leg for lumbar studies and neck/arm for cervical studies) and for function should be clearly defined. The success criteria for pain and function is dependent on the scales used, the indication involved, and the goals of the treatment. With regards to the secondary pain assessment of *donor site* pain, the sponsor should present the data; however, success criteria for this secondary pain parameter are not necessary.

FDA believes that whether combination pain/function scales or independent pain and function scales are used, the success criteria for an individual patient should be based on a clinically meaningful level of improvement. The clinically meaningful level of improvement should be

clearly specified (i.e., X-pts or X% improvement from baseline).

8.3 Height Assessment

8.3.1 Disc Height Assessment

In disease processes where disc height diminishes, and perhaps is the cause of pain/neurological symptoms, this endpoint is relevant in the determination of overall success. Many spinal implants, specifically intervertebral body fusion devices, attempt to restore disc height. Obtaining a measurement of disc height, however, may be somewhat troublesome due to obscuration of the disc space by spinal implants. However, to the extent that this measurement can be made, postoperative disc height should be maintained long term. A baseline disc height measurement may be obtained approximately 6 weeks postoperatively or may be extrapolated from the measurement of adjacent, non-diseased disc spaces. Both anterior and posterior disc heights should be considered in evaluating this endpoint.

8.3.2 Vertebral Height Assessment

In disease processes where vertebral height diminishes, and perhaps is the cause of pain/neurological symptoms, this endpoint is relevant in the determination of overall success. Many spinal implants, specifically vertebral body replacement devices, attempt to restore vertebral height. Obtaining a measurement of vertebral height, however, may be somewhat troublesome due to obscuration by spinal implants. However, to the extent that this measurement can be made, postoperative vertebral height should be maintained long term. A baseline vertebral height measurement may be obtained approximately 6 weeks postoperatively or may be extrapolated from the measurement of adjacent, non-diseased vertebral bodies.

8.4 Health Related Quality of Life

Because the majority of spinal surgeries are performed on an elective basis determined by the limitations on activities and the severity of symptoms as perceived by the patient, an important aspect of ascertaining success is a measure of the patient's overall well-being and satisfaction and, therefore, should be considered in developing a study.

There are a number of assessment instruments geared to this purpose (e.g., SF-36 Short Form Questionnaire). An abbreviated, but still validated form of this survey, is the SF-12. The lesser used Nottingham Questionnaire and EuroQol are also acceptable surveys. Therefore, although the SF-36 is comprised of a physical section, as well as a mental section, the physical section lacks the specificity to capture changes in back function. The SF-36, as well as these other instruments, provides insight into a patient's overall status. The SF-36 and SF-12 incorporate several domains of health into a physical component summary (PCS) and mental component summary (MCS).

9. Safety Evaluation

Safety assessment involves assessment of the complications, including subsequent surgical interventions and neurological complication information. A description of all subsequent surgical interventions, deaths, and neurological complications, including details of their resolution, should be provided.

All preoperative, operative, and postoperative complications, whether device-related or not, should be

recorded. These include "anticipated" complications as well. Pain, neurological, and function symptoms should be considered complications when a patient's complaint for any of these symptoms results in an unscheduled visit or when a patient presents with new or worsening pain, neurological, and/or function symptoms as compared to the previous visit.

9.1 Subsequent Surgical Interventions

Some complications lead to a subsequent surgical intervention. FDA categorizes the subsequent surgical interventions as follows:

- A revision is a procedure that adjusts or in any way modifies or removes *part* of the original implant configuration, with or without replacement of a component. A revision may also include adjusting the position of the original configuration.

- A removal is a procedure where *all* of the original system configuration are removed with or without replacement.

- A reoperation is any surgical procedure at the involved level(s) that does not removal, modification, or addition of any components to the system.

- A supplemental fixation is a procedure in which additional instrumentation not under study in the protocol is implanted (e.g., supplemental placement of a rod/screw system or a plate/screw system).

The sponsor should incorporate the four subsequent surgical interventions above into the protocol as potential risks. It is essential for the sponsor to capture the reason for each subsequent surgical intervention and the action taken (e.g., replacement of a screw, placement of extra bone grafting material, etc.). Along with the presentation of the subsequent surgical interventions pooled into the four categories above, each category should be further stratified. For example, the revision category data presentation may be stratified into revision for translated cage, removal of screws, etc., depending on the reasons identified in a particular study. As another example, the removal category may be stratified into removal for pain at operative site but after fusion, for pseudoarthrosis, etc.

FDA believes that some reasons for performing a removal may constitute a failure; however, this is also very dependent on the device type. The sponsor should clearly identify which reasons for removal constitute a patient failure and provide a rationale. For example, removal of a cage at any time should constitute a failure; however, removal of a pedicle screw system after fusion may not. If removal surgery is recommended in the protocol for a given spinal system, the sponsor should clearly indicate in their protocol how such removals will be interpreted in terms of success and failure of the study. Additionally, the sponsor should also identify any other subsequent surgical intervention(s) that constitute a patient failure.

9.2 Neurological Status

Comprehensive neurological evaluation should be performed due to the potential risks of spinal surgery, in particular, due to the proximity and vulnerability of the spinal cord. While patients may not present with neurological dysfunction preoperatively, spinal surgery introduces the risk of spinal cord injury or other neurological damage. Comprehensive evaluation is necessary. Therefore, assessments should include motor, sensory, and reflex evaluations, and, for studies involving the lumbar spine, straight leg raising evaluations.

Many different assessments exist for evaluating neurological outcome. It is important to select an instrument that is sensitive enough to capture changes that might be anticipated in the population under study. Many of the patients receiving spinal implant surgery are not severely neurologically compromised, and, therefore, not all spinal injury assessment tools can fully capture these changes. The Frankel scale, which discerns between severe degrees of spinal cord injury, is not appropriate assessment tools for a relatively neurologically intact patient cohort. It is important to appreciate that a change that may be classified as minor by the spinal cord injury assessment instruments may be devastating to the previously neurologically intact patient. Such instruments, therefore, should be carefully evaluated and used with caution.

A detailed neurological assessment should capture all degrees of change in the motor, sensory, and reflex portions. Sensory evaluation should be performed by dermatomal levels. The motor evaluation should measure strength in the major muscle groups distal to the level(s) of involvement. For surgeries involving the lumbar spine, a straight leg raising test should be performed and the results recorded in terms of degrees of movement required to reproduce symptoms.

10. Patient / Study Success

Patient success will be used to assess the success rate of an individual patient within a study group. This information will be used as part of the labeling. The study success will be used to determine the safety and effectiveness between the investigational and control study groups.

There may not be sufficiently larger number of patients evaluated at the study endpoint (e.g., 2 years) due to dropouts or censored (incomplete) data, accordingly, the sponsor should perform appropriate statistical analyses, as described in Section 11.

10.1 Patient Success

The success of a patient should be based, at minimum, on success in each of the primary evaluation parameters of fusion (or radiographic success for nonfusion systems), pain, and function as well as no permanent neurological deficit. Depending on the proposed patient population, study design, and study goals, other assessments may be included.

Individual patient success rates should be provided at each postoperative time point, but there will be particular focus on the results at the study endpoint (e.g., 2 years).

10.2 Study Success

Study success takes into consideration the purpose of the treatment and comparison to a control group as well as the study goals (e.g., superiority or equivalency). All primary evaluation parameters, at minimum, as well as safety information, should be accounted for in the definition of study success.

One example of study success criteria of a lumbar equivalency study based on a comparison of the individual patient success rates is as follows:

- *patient success rate of investigational group is no worse than that of control; and*
- *complication rate (including subsequent surgical interventions and neurological complications) of investigational group is no worse than that of control.*

The sponsor should clearly identify and justify the pre-specified allowable difference (delta) used to define “no worse than”.

Depending on the claims that a sponsor may be making (e.g., faster time to fusion, less pain at

earlier time points, etc.), it may be appropriate to incorporate the results of the longitudinal, trend, and survival analyses (Section 11.2 below) as part of the study success criteria.

Study success rates should be provided at each postoperative time point, but there will be particular focus on the results at the study endpoint (e.g., 2 years).

11. Statistical Analyses / Data Presentations

11.1 Statistical Analyses

A complete sample size justification should be provided in the IDE protocol.

- **sample size justification** - A statistical rationale should be provided for the proposed number of patients to be enrolled in the study, including control patients. This number should be based on the ability to detect minimal clinically and statistically significant differences between the investigational and control groups with a given power, pre-specified type I error, expected variability of the outcomes and expected success and complication rates for both the investigational and control groups. Additionally, the expected withdrawal/lost to follow-up rate over time should be taken into account to adjust for the sample size. Consideration should also be given to the number of variables incorporated into the study design (e.g., allograft versus autograft use, one versus two or multi-level spinal fusions, etc.). If pooling of the data cannot be justified at the PMA or 510(k) stage, a subgroup analysis using subgroups which are pre-defined in the IDE protocol may be beneficial.

Additionally, the IDE protocol should include a description of the type of statistical analyses to be performed for the PMA or 510(k) to assess the study results. A rationale and description of each of the statistical analyses to be employed in assessing the effectiveness of the treatment should be provided. This includes an appropriate statistical comparison of the treated patients to the control patients for all evaluation parameters by considering comparable follow-up times. Statistical evidence of device effectiveness is essential. However, this alone is not sufficient for PDP or PMA approval or 510(k) clearance. Sponsors should demonstrate both statistical and clinical significance. FDA believes that the following analyses should be performed for the PMA or 510(k) and, therefore, be described in the IDE protocol:

- **poolability across investigators and sites** – Appropriate statistical and clinical justification, such as homogenity of treatment outcomes across investigators and sites, are recommended.

- **longitudinal data analyses with covariates –** Longitudinal data analysis should be performed for various data types such as continuous, binary, count, or ordinal (ranking) data. Appropriate statistical modeling approaches, including device code (e.g., 0 for control and 1 for investigational device), important patient covariates and demographic characteristics, should be considered in the statistical model in order to evaluate true device effect and adjusting for possible differences in the distributions of important patient covariates and demographic characteristics between control and investigational device groups. Computer software, such as SAS, S-Plus, or others is available to analyze such repeated-measure data.

- **trend analyses** - Clinical outcomes may vary from time to time (time trend), device to device (device effect), and crossover between device and time (device by time interaction). A simple comparison between two arbitrarily selected time points will not reveal true device effect over time. Therefore, FDA suggests that repeated measure analysis of

variance (RMANOVA) for continuous or quantitative data and longitudinal data analysis should be performed.

- **survival analyses –** For success/failure (binary data), statistical survival analysis with a multivariate model (e.g., Cox proportional hazard regression) may be used to estimate the time-specific cumulative fusion success probability and effect of important patient covariates on fusion success/failure while accounting for patients who are dropouts or censored (incomplete).

The sponsor should consider the following statistical issues:

- The number of study patients is likely to decrease over time due to dropouts or censored (incomplete) data. Statistical analysis based on partial data presents a challenge because one cannot be sure that the dropout or censored patients will have the same clinical outcomes with their partially observed clinical outcomes as they would have if they had completed the whole follow-up schedule. Therefore, the sponsor should make every attempt to obtain complete follow-up on all patients entered into the study.

- The reasons for any missing data, whether device, covariate, or clinical outcome related, should be clearly explained. The distribution of missing data between the control and active device groups should be carefully evaluated.

- The sponsor should clearly state the definition of clinical outcomes to be compared between the two treatment groups. If "change from baseline" is used, then the comparability of clinical outcomes and covariates at baseline should be demonstrated between the control and active device groups in order to avoid the confounding effect between device effect and baseline measurements. Please note that, simple randomization does not guarantee comparability of clinical measurements and covariates between investigational and control groups at baseline. If baseline values are quite different between two groups, then appropriate statistical analysis, such as using the baseline as the covariate, should be employed.

11.2 Data Presentations

The types of data presentations should be considered at the time of protocol development in order to better assure that adequate data are collected. Although it is not necessary to provide this information as part of the IDE protocol, FDA believes that the following types of data presentations, stratified by investigational and control groups, should be provided in an IDE annual report, as well as in a PMA or 510(k):

- time course distribution of patient accounting (e.g., theoretically due, deaths, reoperations, revisions, removals, supplemental fixations, expected, actual number followed, and follow-up rate);

- written descriptions of the subsequent surgical interventions, deaths, and neurological complications including any actions taken, resolutions, etc.;

- time course distributions of the following individual parameters – fusion, back or neck pain, leg or arm pain, function (may be combined parameter with back or neck pain), vertebral body height or disc height, health related quality of life, donor site pain, motor strength, sensory, reflexes, and straight leg raisings (for lumbar studies). These time course distributions provide the number of patients at the different levels within a given each scale (e.g., as for pain, function, etc.) or the number of patients that meet each criteria (as for fusion). For example, for fusion, the time course distribution for fusion would show the

number of patients who had translational motion, rotational motion, trabecular bone formation, etc. As another example, the time course distribution for leg pain would show the number of patients with pain at level 1, level 2, etc.;

- time course distributions of all complications, including subsequent surgical interventions and neurological complications for all patients; and

- time course distributions of the individual patient success rates for all patients.

In order to properly evaluate the data, it is imperative that the sponsor clearly identifies the number of patients involved at a give time point in any data presentation.

12. Patient Report Forms

Copies of all patient report forms should be provided. The forms should include all relevant information from the protocol. The following are the type of forms that should be part of a study.

- Enrollment form or Inclusion/Exclusion Criteria form - identifies all inclusion and exclusion criteria in protocol
- Demographic and Preoperative Evaluation - includes demographic information and the clinical and radiographic parameters to be measured
- Operative Data form
- Discharge form
- Postoperative Evaluation form - include the clinical and radiographic parameters to be measured
- Adverse Event form - identifies all potential risks from the protocol and provides for reporting of other adverse events

The sponsor should also provide a table that specifies the time point(s) at which each patient report form is to be completed.

13. Risk Analysis

This section should include adequate information to determine that the benefits and knowledge to be gained outweigh the risks to the patients. This information includes:

- a description and analysis of all increased potential risks to which patients will be exposed by investigation;
- the manner in which the potential risks will be minimized; and
- rationale for the investigation.

When listing all of the potential risks, they should be stratified by those general to spinal surgery and those specific to the system. The sponsor should also provide a statement that all adverse events/complications, device related or not, will be reported and recorded.

The list of potential risks should be the same in the Risk Analysis section, Adverse Event patient report form, package insert, and informed consent document.

14. Post-Operative Regimen

Any additional patient care procedures to be employed during the treatment period (e.g., surgery, rehabilitation, immobilization, weight bearing, etc.) should be detailed.

15. Retrieval Study

Because of the unknown long term device performance of some types of spinal systems, particularly the resulting bony fusion characteristics, a sponsor should incorporate a plan that focuses on the retrieval analyses of a spinal system that is implanted and subsequently removed. Histological information (e.g., bony ingrowth quality, bone quantity, response to potential wear debris, etc.) and metallurgical information (e.g., metal wear, deformation, cracking, corrosion, etc.) should be collected and reported in the annual reports.

16. Spinal Tumors

Studies to analyze the safety and effectiveness of vertebral body replacement devices for the treatment of spinal tumors should take into account several factors which differentiate this population from other vertebral body replacement patients in particular, and spinal disease patients in general. These factors include such things as the life expectancy of the patient, the total number of patients eligible to participate, the specific purpose of the device, the duration of the clinical trial and the appropriate clinical outcome parameters.

16.1 Inclusion Criteria

For inclusion into a tumor study, the sponsor should consider the previous treatments (e.g., radiation, chemotherapy), type of tumor (primary benign, primary aggressive/malignant, or metastatic), and the goals of the treatment. FDA recognizes that a tumor study may involve a broad range of disease entities and life expectancies. Patients are expected to present with a primary complaint of pain, neurological deficit, and/or functional deficit or with the intent to prevent a pathological fracture.

16.2 Duration / Follow-up Schedule

It is expected that long-term follow-up for many of the patients enrolled in a trial for this type of device may not be possible and that there will be a large number of patients who would be categorized as lost-to-follow-up. As a result, from the standpoint of data collection, the goal of the trial is to gather as much data as possible. The sponsor should describe the follow-up schedule and duration of the study based on the patient population at hand and the study goals. All patients should be followed until the last enrolled patient has reached the justified study endpoint or has withdrawn from the trial due to death.

16.3 Sample Size / Number of Sites & Investigators

The total number of patients that would meet the inclusion/exclusion criteria is relatively small and the rate at which these patients would be available for enrollment is very low. Therefore, no limits will be placed on the number of sites/investigators that may be enrolled for a particular clinical trial. IRB approval will still be required for each enrolled site, however. In addition, monitoring in accordance with the IDE regulations will also be required.

16.4 Controls

Controls should be incorporated. A concurrent control population is preferred. The concurrent controls may be used at separate sites than the investigational patients and may involve use of any currently acceptable standard of care for spinal tumors, *e.g.*, anterior plates with bone graft, PMMA, etc.

Other potential sources of control data are retrospective, consecutive data (case series) from the participating investigators or literature. Because of the large variability in treatments and evaluated clinical outcome parameters, these possible sets of control data present potential difficulties during data analysis. In order to allow for proper comparison, the same clinical outcome parameters from these populations should have been collected at similar time intervals. Consequently, these alternate sources of control data should be discussed in detail with FDA prior to initiation of trial.

16.5 Clinical Parameters

Patients should be evaluated for the following:

- pain (leg and/or back)
  Attempts need to be made to distinguish pain due to problems with the spine, *e.g.*, nerve root impingement, from the general pain that these patients might experience as the result of their medical condition.

- neurological status
  The same assessments used for evaluation of neurological status for patients with other spinal devices are applicable to these patients.

- patient function (functional independence)
  Evaluation of function should focus on the ability of patients to function independently, *e.g.*, how does their ability to better move around the house/neighborhood, dress themselves, etc. compare to their pre-op status.

- radiological assessments
  The radiological assessments depend on the patient population and study goals. These assessments may include integrity of implant, maintenance of correction, lack of migration, and/or fusion.
- health-related quality of life
  It is anticipated that an improvement in a patient's "quality of life" will occur as a result of receiving the implant. Therefore, the sponsor should identify an instrument that appropriately and adequately captures this aspect of outcome for the population under study.

16.6 Safety

Safety should be evaluated similar to that described in Section 9 above.

16.7 Patient / Study Success

Individual patient success may be based on a clinically meaningful level of improvement for a fixed number of clinical parameters. The sponsor may also consider basing individual patient success on a clinically meaningful level of improvement in the patient's primary complaint or reason for surgery (e.g., to relieve pain, prevent pathological fracture, etc.). For example, a patient who presents with a primary complaint of pain should improve in pain to be considered successful for that parameter, but only maintenance or improvement in neurological status and patient function are necessary to be considered successful for those parameters. Patients presenting with multiple primary complaints should improve in each one of them in order to be considered successful. A device that fails from the mechanical/positional standpoint (device function) and requires that a patient undergo a second, corrective surgery will categorize that subject as a failure.

Study success criteria based on a comparison of the individual patient success rates between the investigational and control groups or a comparison of the individual parameters between the two groups.

16.8 Data Presentations / Statistical Analyses

Data presentations similar to those described in Section 11.2 apply, with all available data being presented. Because the life expectancy of the enrolled patients will be variable, there is the potential to produce a dataset that is very "uneven" and predisposed to large losses of information at later evaluation timepoints. As a result, patient clinical behavior should be assessed as a function of time using survivorship and longitudinal analyses. These analyses will show the effect of important patient covariates on success/failure while accounting for patients who are dropouts or censored (incomplete). Again, these analyses should involve all available data.

DEVICE DESCRIPTION

In accordance with 812.25, adequate device description information should be submitted. More specifically, the following information, at minimum, should be provided for any spinal system:

(1) a table that includes each component name and corresponding part number;

(2) a complete written description of the individual components and how any components interconnect;

(3) complete mechanical drawings with all dimensions and tolerances of each individual component and, if applicable, of the total system;

(4) supporting magnified sketches or photographs of the major interconnection mechanisms and of the entire system attached to a spinal model;

(5) a description of the material(s) and any voluntary material standard(s) to which the material(s) conform. Depending on the material, biocompatibility issues may need to be addressed. Biocompatibility information should be included in the Report of Prior Investigations section of the IDE;

(6) a description of anticipated changes to the system; and

(7) a list of all instruments *unique* to the implantation of the subject system, the material and voluntary material standard to which they conform, and supporting magnified sketches or photographs of them.

For any concurrent control system, a written description, any available mechanical drawings and photographs, and material information should be provided.

MANUFACTURING

As per 21 CFR 812.20(b)(3), a description of the methods, facilities and controls used for the manufacture, processing, packaging, and storage of the device in sufficient detail so that a person generally familiar with good manufacturing practices can make a knowledgeable judgment about the quality control used in the manufacture of the device is to be provided.

As part of that information, the following should be provided for any spinal system:

(1) basic manufacturing information that deals with device design issues (e.g., application of coatings, unique material processing, processing of the bone morphogenic protein, etc.); and

(2) sterilization information. For systems provided sterile, the sterilization method(s) used, all parameters of the sterilization cycle, the validation method, the resulting sterility assurance level (SAL), a description of the packaging, and data supporting any pyrogenicity claims should be provided. For systems provided nonsterile which must be sterilized prior to use or for which is labeled as to be resterilized, the recommended sterilization method(s), all parameters of the sterilization cycle, the validation method, the resulting SAL, and a description of the packaging should be provided.

MONITORING

The following is not information specific to a spinal IDE but is necessary to complete the IDE submission:

(1) a copy of the written procedures for monitoring the investigation. These procedures should meet the minimum requirements described in 21 CFR 812.46; and

(2) the name and address of the monitor.

IRB and INVESTIGATOR INFORMATION

The following is not information specific to a spinal IDE but is necessary to complete the IDE submission:

(1) a sample of the investigator agreement. At minimum, the agreement should include the information described in 812.43(c);

(2) certification that all investigators participating in the investigation will sign the investigator agreement prior to entrance into the study;

(3) the name and site address of each investigator who has signed the investigator agreement; and

(4) the name, address, and chairperson of each IRB. In addition, any action taken by each IRB (e.g., approval) and how many IRBs are currently reviewing the investigation or will review it in the future.

LABELING

In accordance with 21 CFR 812.20(b)(10), copies of all labeling for the system should be provided. The Labeling section for the spinal system IDE should, at minimum, include sample package labels, a package insert, and a surgical technique manual. Although the information requested for each is generic to any orthopedic device, the level of detail is typically greater for spinal systems (e.g., the listing of all inclusion criteria, implantation procedures, system directions, etc.).

1. Package Labels

The draft package labels should include, at minimum, the following information:

- system and device names;
- quantity;
- material and voluntary material standard;
- sterile or nonsterile notation;
- name and address of business of manufacturer, packer, or distributor (in accordance with 21 CFR 801.1); and
- a statement that the use of the device is investigational for the proposed indication as follows: "CAUTION - Investigational Device. Limited by Federal (or United States) law to investigational use."

If the components of a system are packaged individually, then please provide draft package labels for each type of component (e.g., screw, hook, rod, biologic material, etc.).

2. Package Insert

The draft package insert should include, at minimum, the following information:

- system name;
- brief device description with material information and any system directions;
- inclusion criteria;
- list of any pertinent contraindications, warnings, and precautions;
- list of potential adverse events/complications;
- sterile or nonsterile notation;
- recommended sterilization parameters if provided nonsterile or if resterilization is allowed; and
- a statement that the use of the device is investigational for the proposed indication as follows: "CAUTION - Investigational Device. Limited by Federal (or United States) law to investigational use."

The information provided in the package insert should be consistent with that provided in the rest of the submission, including any inclusion criteria, potential risks, device description information, etc.

3. Surgical Technique Manual

The draft surgical technique should include, at minimum, the following information:

- system name;
- written procedures of how to implant the subject system along with supporting magnified sketches or photographs of each major step; and
- a statement that the use of the device is investigational for the proposed indication as follows: "CAUTION - Investigational Device. Limited by Federal (or United States) law to investigational use."

If the protocol involves a concurrent control system, then the surgical technique manual should include adequate instructions for implantation of the control system, or a separate surgical technique manual for the control system should be supplied.

INFORMED CONSENT

In accordance with 21 CFR 812.20(b)(11), copies of all forms and informational materials to be provided to the subjects in order to obtain informed consent are to be provided. In addition, a statement that the subject must sign the informed consent document prior to entrance into the study should be provided.

The informed consent document is to meet the general requirements as described in 21 CFR 50.20. In addition, according to 21 CFR 50.25(a), the informed consent should include the following items:

(1) a statement that the study involves research;
(2) an explanation of the purposes of the research;
(3) the expected duration of the subject's participation;
(4) a description of the procedures to be followed (e.g., evaluation time points, general types of evaluations, any special postoperative regimen, etc.);
(5) identification of any procedures which are experimental (i.e., clearly identify the investigational system of the study);
(6) a description of any reasonably foreseeable risks or discomforts to the subject;
(7) a description of any benefits to the subject or others;

(8) a disclosure of appropriate alternative procedures or courses of treatment that might be advantageous to the subject;
(9) a statement describing the extent to which confidentiality of the subject's records will be maintained and that notes that FDA may inspect the records;
(10) an explanation as to whether any compensation and/or medical treatments are available if injury occurs and, if so, what they consist of or sources of further information;
(11) an explanation of whom to contact for answers to questions about the study and the subject's rights and whom to contact in the event of a research-related injury; and
(12) a statement that participation is voluntary and that subjects may refuse to participate or discontinue participation at any time without penalty or loss of benefits.

If necessary, additional elements to the informed consent document as per 21 CFR 50.25(b) may need to be addressed.

While the information requested above is not specific to a spinal IDE, it is important that the informed consent document be consistent with the information provided in the Device Description section and the Investigational Plan section, including the proposed patient population, purpose, duration, follow-up time points, evaluation parameters, potential risks, etc. This is because there are some variations between spinal IDEs in terms of the designs, surgical approaches, patient groups, and parameters.

SALES

The following is not information specific to a spinal IDE but is necessary to complete the IDE submission. In accordance with 812.20(b)(8), if the system is to be sold, the amount to be charged and an explanation of why the sale does not constitute commercialization of the system are to be provided.

ENVIRONMENTAL IMPACT ASSESSMENT

The following is not information specific to a spinal IDE but is necessary to complete the IDE submission. An environmental assessment as described in 21 CFR 25.31(a) or a statement of categorical exclusion from this requirement under 21 CFR 25.24(e)(7) is to be provided.

Note: *FDA acknowledges that the spine area may be considered complicated to sponsors. ORDB is eager to provide input prior or during the development of an IDE for spinal systems, especially in the areas of:*

- *appropriate mechanical testing for a specific system design;*
- *appropriate animal study design;*
- *presentation of existing clinical data; and*
- *specific and general protocol questions (appropriateness of inclusion criteria, evaluation parameters, success criteria, control group, etc.).*