DEPARTMENT OF HEALTH AND HUMAN SERVICES
PUBLIC HEALTH SERVICE
FOOD AND DRUG ADMINISTRATION

ORTHOPAEDICS AND REHABILITATION

DEVICES ADVISORY PANEL MEETING

December 11, 1997

Bethesda Holiday Inn
Versailles Ballrooms III and IV
8120 Wisconsin Avenue
Bethesda, Maryland

Proceedings By:

CASET Associates, Ltd.
10201 Lee Highway #160
Fairfax, Virginia 22030
(703) 352-0091

PARTICIPANTS:

Panel Members:
Barbara D. Boyan, Ph.D.
Doris M. Holeman, Ph.D.
Cato T. Laurencin, M.D., Ph.D.
Michael J. Yaszemski, M.D., Ph.D.
Thomas W. Bauer, M.D., Ph.D.
Edward Y. Cheng, M.D.
Cindy Domescus, Ph.D.
David Hackney, M.D.
Joseph E. Hale, Ph.D.
Janine Janosky, Ph.D.
Casey Kerrigan, M.D.
Harold Wilkinson, M.D.
Celia Witten, Ph.D., M.D.
Leon Grobler, M.D.
Jodi Nashman
Raymond Silkaitis, Ph.D.

TABLE OF CONTENTS

	Page
Open Public Hearing	
Read Into the Record by Ms. Nashman	
William J. Kane, M.D., Ph.D.	7
John J. Regan, M.D.	10
Harry B. Skinner	11
Hanson A. Yuan	12
Presentations	
Edgar Dawson, M.D.	16
Neil Kahanovitz, M.D.	20
John Dichiara	23
Sofamor Danek Presentations	
Dvorah A. Richman, Esquire	25
Robert T. Rylee, II	28
Stephen E. Heim, M.D.	31
United States Surgical Corporation	
Louis J. Mazzarese	34
Scott Blumenthal, M.D.	37
Tushar C. Patel, M.D.	40

SpineTech, Inc.

John A. Dowdle, M.D. 42

Richard C. Jansen 45

Discussion of Clinical Endpoints 54

TABLE OF CONTENTS (Cont.)

Page

Open Session -- Sofamor Danek Novus LC Interbody Fusion System

Manufacturer Presentations 98

FDA Presentations 138

Panel Discussion 160

Questions and Voting 224

Open Session -- Acromed Brantigan Cage with VSP Fixation:

Manufacturer Presentations 285

FDA Presentations 334

Panel Discussion 363

Questions and Voting 376

P R O C E E D I N G S [9:09 a.m.]

MS. NASHMAN: We are ready to begin this meeting of the Orthopaedics and Rehabilitation Devices Panel, running a little bit over for starters.

My name is Jodi Nashman. I am the executive secretary of this panel, biomedical engineer and a reviewer in the Orthopaedic Devices Branch.

I would like to remind everyone that you are requested to sign in on the attendance sheets, which are available at the table by the door. You may also pick up an agenda and information about today's meeting, including how to find out about future meeting dates from the advisory panel phone line and how to obtain meeting minutes or transcripts.

I am going to read two statements that are required to be read into the record: the deputization of temporary voting member statement and the conflict of interest statement.

First, the conflict of interest statement: The following announcement addresses conflict of interest issues associated with this meeting and is made part of the record to preclude even the appearance of an impropriety. To determine if any conflict existed, the Agency reviewed the

submitted agenda and all financial interest reported by the committee participants.

The conflict of interest statutes prohibit special government employees from participating in matters that could affect their or their employers financial interests. However, the Agency has determined that participation of certain members and consultants, the need for whose services outweigh the potential conflict of interest involved is in the best interest of the government.

A full waiver has been granted to Dr. Cato Laurencin for his interest in firms, which could potentially be affected by the Panel's decisions. A copy of this waiver may be obtained from the Agency's Freedom of Information Office, Room 12A-15 of the Parklawn Building.

We would like to note for the record that the Agency took into consideration other matters regarding Dr. Leon Grobler and Dr. Barbara Boyan. Dr. Grobler reported that his institution has a financial interest in a firm at issue for matters related to today's discussion. Since Dr. Grobler has no personal interest in this matter, the Agency has determined that he may participate fully in today's discussions.

Dr. Boyan reported interests in orthopaedic firms

on matters not related to issues before the Panel. Since these matters are not related to the agenda of this meeting, the Agency has determined that Dr. Boyan may participate fully in today's discussions.

In the event that the discussions involve any other products or firms not already on the agenda for which an FDA participant has a financial interest, the participants should excuse themselves from such involvement and their exclusions will be noted for the record.

With respect to all other participants, we ask in the interest of fairness that all persons making statements or presentations disclose any current or previous financial involvement with any firm whose products they may wish to comment upon.

Next is the appointment to temporary voting status.

Pursuant to the authority granted under the Medical Devices Advisory Committee Charter, dated October 27, 1990, as amended April 20, 1995, I appoint the following people as voting members of the Orthopaedics and Rehabilitation Devices Panel for the duration of the meeting on December 11 and 12, 1997: Dr. Cato T. Laurencin, Dr. Michael J. Yaszemski, Dr. Edward Y. Cheng, Dr. Leon J.

Grobler, Dr. Joseph E. Hale, Dr. Janine Janosky, Dr. D. Casey Kerrigan and Dr. Harold Wilkinson.

For the record, these people are special government employees and are either a consultant to this panel or a consultant or voting member of another panel under the Medical Devices Advisory Committee. They have undergone the customary conflict of interest review. They have reviewed the material to be considered at this meeting.

Also, because the position of panel chairperson for the Orthopaedics and Rehabilitation Devices Panel is currently vacant, I appoint Dr. Barbara D. Boyan to act as temporary chairman for the duration of the Orthopaedics and Rehabilitation Devices Panel meeting on December 11 and 12, 1997.

For the record, Dr. Boyan is a special government employee and is a voting member of the Orthopaedics and Rehabilitation Devices Panel. Dr. Boyan has undergone the customary conflict of interest review and she has reviewed the material to be considered at this meeting.

Before I turn the meeting over to Dr. Boyan, I would like to introduce the distinguished Panel members, who are giving their time to help the FDA in the matters that are being discussed and also introduce the other FDA staff

seated at this table.

For my own ease, I am going to read the list in alphabetical order instead of trying to read going around the table.

The chairman of the meeting is Dr. Barbara D. Boyan from the University of Texas Health Sciences Center.

Dr. Doris Holeman is the consumer representative for the meeting from Albany State College

Dr. Cato Laurencin is a consultant, deputized to vote, and he is professor of orthopaedic surgery.

Dr. Raymond Silkaitis is the industry rep for this meeting. He is representing Gliatech. He will only be serving for today's portion of the meeting.

Dr. Michael Yaszemski is a consultant deputized to vote, with the Mayo Clinic.

Dr. Edward Cheng is a consultant, deputized to vote, from the University of Minnesota.

Dr. Leon Grobler is a consultant, deputized to vote from the University of Iowa Hospitals.

Dr. Joseph Hale is a consultant, deputized to vote, from the Minnesota Sports Medicine Center.

Dr. Janine Janosky is a biostatistician, consultant, deputized to vote, from the University of

Pittsburgh.

Dr. Casey Kerrigan is a consultant, deputized to vote, from the Harvard Medical School.

And Dr. Harold Wilkinson is a consultant, deputized to vote, from the University of Massachusetts.

Also sitting at the table is Dr. Celia Witten, the division director of the Division of General and Restorative Devices.

At this time, I would like to turn the meeting over to our chairperson, Dr. Barbara Boyan.

DR. BOYAN: Good morning. I am Dr. Barbara Boyan and I am the chairperson for the meeting.

Today, we will be making recommendations to the Food and Drug Administration on two premarket approval applications. I would like to note for the record that the voting members present constitute a quorum as required by 21(CFR)Part 14.

Agenda Item: Open Public Hearing

We will now proceed to the open public hearing session of this meeting. I would like to ask at this time that all persons addressing the Panel come forward and speak clearly into the microphone, as the transcriptionist is dependent on this means of providing an accurate record of

this meeting.

We are requesting that all persons making statements during the open public hearing of the meeting disclose whether they have financial interests in any medical device company before making your presentation to the Panel. In addition to stating your name and affiliation, please state the nature of your financial interests, if any.

MS. NASHMAN: Okay. Just to give you all an idea of how this is going to be organized, there are four pieces that I need to read into the record that were mailed to me. Then I will turn the meeting over to Dr. Boyan. She will read through the names of individuals who contacted me and request that they come to the podium and make their presentation of the length specified.

I am going to start with correspondence from a Dr. William Kane, M.D., Ph.D. He is associated with the Orthopaedic and Arthritis Institute.

"At the December 11 and 12 meeting of the FDA Orthopaedic Advisory Panel, I am informed that there will be a discussion on the minimum acceptable length of patient follow-up for marketing approval of spinal implant devices.

"As a spine surgeon for nearly four decades, as

well as a researcher and teacher both at the University of Minnesota Medical School and at Northwestern University Medical School over the same length of time, I have been privileged to see the transition from the pre-instrumentation era of spine surgery to the present state of sophistication. I was the orthopaedic resident assigned to the first patient in the State of Minnesota to receive a Harrington rod. In August 1960, Drs. Harrington and Moe operated together and, today, Mary S., that scoliosis patient is married with children.

"I also worked closely with Allan F. Dwyer after he had introduced anterior spinal instrumentation for scoliosis correction and after he later introduced electrical bone growth stimulation for augmentation of lumbosacral fusion procedures.

"More recently, I have weighed the advantages and disadvantages of pedicle screws and have participated in a number of educational debates, as well as published on this particular issue. I have also served a term on the FDA's Advisory Panel on Orthopaedic Devices while pedicle screw approval was weighed, and was asked to speak to this controversial issue by the FDA in a formal presentation in August of 1993, at the Pedicle Screw Spinal Fixation Device

Systems Symposium of the Orthopaedic and Rehabilitation Devices Panel of the Medical Devices Advisory Committee, Food and Drug Administration.

"I have been able to be at the forefront of the important issues dealing with spine surgery and its innovations over these years.

"It is from this background that I wish to address the necessity to maintain at least a two-year patient follow-up for assessing the utility of a device or implant in the spinal column. It may even take longer for devices which are mobile, that is, which move in conjunction with the spinal element, but I can think of no instance where it would not take a minimum of two years to assess a non-moving implant, which was designed to assist in achieving a fusion or arthrodesis of vertebral elements, whether from the anterior or the posterior approach. Even at one year post-surgical implantation, it is not easy and frequently impossible to assess the status of the fusion. This is especially so, in my experience, for interbody implants which, in themselves, obscure the detection of either a solid fusion or a definite pseudarthrosis.

"I am aware that the FDA has been directed to speed up the assessment process, but, quite frankly, I do

not believe that the directive should be applied to every phase of the assessment process. You simply can't legislate the speed with which natural healing of the body, such as spinal arthrodesis via osteogenesis, can occur. Speed up the other phases, if need be, but don't shortcut the most critical element in the approval process; namely, how a device works in a patient's spine.

"My viewpoint is not unique. If the editorial boards of the two most important scientific journals dealing with spinal implants, namely, the Journal of Bone and Joint Surgery and Spine, insist on two-year patient follow-ups before publishing an article in their journals, how can the FDA, which has an even graver responsibility, namely, putting its imprimatur on an implant for marketing and general usage in this country and by extension worldwide, do so with less than two year patient follow-up?

"Thank you for your attention. Yours Respectfully, William J. Kane, M.D., Ph.D."

The second letter is from a Dr. John J. Regan, from the Texas Back Institute.

"This letter is regarding your evaluation of spinal implant devices. I feel it is essential to have a minimum two year follow-up to evaluate the safety of these

devices. With each new technology, there is a risk of problems with the device, which cannot be adequately evaluated in a one year period of time. If we are evaluating a fusion device, the implant may be satisfactory for a one year period of time; however, between the one year and the two year period of time, the implant may fail and this may result in a complication. The panel should be aware of the public health concerns inherent in premature approval of a spinal implant device before this minimum two year period of time. I am sure the panel is aware that in peer review literature for follow-up studies, a two year follow-up is the minimum requirement for any article to be accepted for publication. i think this principle should be followed when evaluating new spinal implant devices.

"The other concern is that with any variation in any new technology, there may be differences that are not perceived, even in biomechanical testing. These problems may not be apparent for this twelve month period of time. Please consider these concerns about the safety and efficacy of this new technology and the public health concerns inherent in premature approval of a device prior to the standard two year interval when you are considering new spinal implant devices. Sincerely, John J. Regan, M.D."

The next correspondence is from Harry Skinner. He is a special government employee, SGE, determined to be conflicted at this meeting. He, however, wished to have the following information presented.

He writes, "If a study is designed to go 24 months and is stopped in less than 24 months, one or more of the following three criteria must be met:

"1 -- The study must have been designed to have an interim analysis, usually with a higher P value required for significance. This should be pre-approved by the FDA.

"2 -- The study must have been demonstrating obviously terrible results, presumably from an unexpected complication, in relation to the control group. In this case, the FDA should be notified and permission to do an interim analysis obtained. In this case, the interim analysis would be done on an ethical basis.

"The study must have been demonstrating obviously wonderful results in relation to the control group. An argument could be made to terminate the study on ethical grounds, but this argument is Much, MUCH, weaker because potential two year complications are lost and the control group is STILL getting accepted treatment.

"I think it is appropriate for the FDA to accept

shorter term data, when the perceived risk is low and then ask for postmarket surveillance. This, however, MUST be in the study design. It is bad statistics and bad science to do monthly analysis of the data to look for statistical significance. I think that in approving devices for indications, such as new spine cages where there are other ones already approved, the use of shorter study times where the risks have already been identified, is good science. The other side of the coin is that it is bad business because someone had to pay for the earlier device's longer study. That is a conundrum I don't have an answer for!"

The fourth and last correspondence that I am going to read is from Dr. Hansen Yuan, who is also a special government employee, who would have been determined conflicted if he were to attend this meeting.

"I am writing this letter after reading the Federal Register of the upcoming Advisory Panel Meeting on December 11th and 12th. This is in relationship to revisiting the topic of minimum acceptable length of patient follow-up for marketing approval of spinal implant devices.

"I will be out of the country at the time of the meeting; otherwise, I would have been present to make comment on this very important topic. I have been, and

currently still am, a non-voting panel member of the FDA Device Panel.

"I was present over the last three years on all discussion concerning this topic,where at one time there was a discussion concerning lengthening minimum acceptable follow-up for marketing approval for spinal implant devices to three years. At the June 12, 1995 Panel Meeting, there was approximately four hours of testimony and discussion on this issue. The majority of the panel agreed that the current accepted standard by the spine community including the editors and associate editors of the Journal of Spine and Spinal Disorder and the Journal of Bone and Joint, that all follow-up should be at least two years as a minimum; that is, where the majority of cases should stand. At that meeting, we did conclude that on new devices, the panel may suggest to the FDA that postapproval studies could continue in special cases. In the case of new technologies, such as interbody fusion cages, that in addition to the two year follow up, we should follow these patients up for five or six years. This data would be reported back to the FDA.

"The above, I am sure, is information that you know. However, I understand there will be a presentation to the FDA Advisory Panel concerning a new cage where a

substantial number of investigational subjects have not yet reached a two year follow-up. I think the FDA Panel and the FDA should maintain the view that a substantial portion of study subjects with spinal implants reach the two year mark for follow-up as a basis of assessing safety and effectiveness for approval. The standard for protection of the general public, a minimum of two years, at this time, should not be altered since there is no science to prove that it can be shortened and similarly, there is no science to prove that it should be lengthened. I strongly urge that we maintain the predominant cases to reach the two year mark in assessing safety, efficacy, and outcome before approving any of these new devices.

"Thank you for the opportunity to present this point of view. Sincerely, Hansen A. Yuan, M.D."

Please note that the FDA does not have any financial disclosure on the writers, with the exception of the two special government employees mentioned.

At this time, I would turn the meeting back over to Barbara Boyan.

DR. BOYAN: At this point, the system that I would like to use is that as the speakers come forward to address the panel, that they state their name, their affiliation,

their financial -- any financial issues that need to be addressed and that they keep their time such that there is an opportunity for the Panel to ask any questions they might want to ask of the speakers.

Is there anybody in the audience that is wishing to make a statement? Oh, I need to read the list. Jodi says I have to read the list. Okay.

There are going to be presentations -- each of these presentations will be five minutes apiece: Edgar Dawson from UCLA Comprehensive Spine Center; Neil Kahanovitz from the North American Spine Society and John -- I hope I pronounce this correctly -- Dichiara from the Orthopaedic Surgical Manufacturers Association.

After that, there will be three presentations from three different companies. Each company will have an opportunity to speak for 15 minutes total. Speaking for Sofamor Danek will be Dvorah A. Richman, Esquire, from King and Spaulding; Robert Rylee from Sofamor Danek, Edgar Dawson and Stephen Heim.

The U.S. Surgical Corporation will be represented by Louis Mazzarese from U.S. Surgical; Scott Blumenthal, Dr. Scott Blumenthal and Dr. Tushar Patel.

And SpineTech will have 15 minutes and they will

be represented by Richard Jansen from SpineTech and Dr. John Dowdle.

In these presentations I ask that you keep the discussion general. We are not here to discuss any specific product. We are here to have a general discussion on the topic.

Would Dr. Dawson please come forward.

MS. NASHMAN: I am sorry to say that I don't believe we have a timer for you to look at. I am just going to go with my watch, which is timed for five minutes.

DR. DAWSON: I will be honest.

Good morning, Panel chairperson and Panel members. My name is Edgar Dawson. I am a spinal surgeon from Los Angeles, California. For over 25 years, I have been a faculty member in the Department of Orthopaedic Surgery at the UCLA Medical Center in Los Angeles. And I am currently clinical professor of orthopaedics.

I am a consultant to Sofamor Danek for education in practice management. I have been chairman of the American Academy of Orthopaedic Surgeons Committee on the Spine and I am a past president of the Scoliosis Research Society.

Over the past five years, I have had the

opportunity to use spinal interbody fusion devices in the treatment of my patients. I continue to use them today since I believe they are very effective in promoting the fusion of the lumbar spine and they have a good safety profile as compared to other spinal implant devices.

I am here today to prove a clinical perspective on the validity of assessing clinical outcome and, therefore, safety and effectiveness of anterior interbody fusion devices at 12 months following surgery.

The scientific and medical question is: Are 24 month follow-up data necessary to determine whether these interbody fusion devices work? Based on my experience, the answer is an unequivocal "no."

Fusion is the primary effectiveness endpoint of these types of devices and is assessed by radiographic means. Evidence of fusion may include bridging of trabecular bone, stability and lucent line criteria. Because distraction of the disc space occurs during the implantation of these interbody fusion devices, the affected vertebral bodies are stable immediately postoperatively.

The threaded design of the implants, as well as the pressure exerted by the vertebral bodies, work to maintain stability of that motion segment. This stability

generally results in early postoperative pain relief. A stable environment, as well as the autogenous bone graft in the cages, promotes the formation of bone, which bridges the vertebral bodies and results in permanent fusion.

In terms of the postoperative radiographic findings and the clinical course of treated patients, generally at six months following lumbar interbody fusion surgery, many patients demonstrate spinal stability. There is absence of lucent lines around the implants and in many cases evidence of bridging trabecular bone. These patients will usually be well clinically and are considered fused.

By the 12 month postoperative visit, treated patients typically fall into three categories. Firstly, there are those that are determined to be fused at six months and are still fused. There are those that were not fused at six months but now at 12 months have progressed to fusion. Finally, there are those that were not fused at six or 12 months and probably will never fuse. Patients in this category will likely have pain and even some disability. When presented with this discomfort and supporting radiographic data, a surgeon will not wait an additional 12 months to see whether or not patients in this category will improve, but will likely perform a posterior operation to

augment the fusion. In fact, waiting an additional 12 months to see whether a symptomatic patient will eventually fuse is not sound medical practice.

In short, patients having lumbar interbody fusion procedures with cage devices will either be fused by 12 months following surgery or will probably never be fused as a result of that particular operation. Patients who are fused say fused and those who are not may require additional surgery. Clinical outcomes between 12 and 24 months typically do not change.

My comments are supported by the testimony of Dr. Hansen Yuan, whose letter you just heard. He is a noted orthopaedic spine surgeon. Dr. Yuan is a past president of the North American Spine Society and at a presentation to this advisory panel meeting in 1995 said, and I will quote, "In our scientific literature, fusions occur mostly between three and six months. A significant percentage continue to fuse up to one year. If that data is good for fusions up to one year follow-up, it is a very positive thing that the study is demonstrating efficacy."

In summary, based on my clinical experience and on the literature, 12 month follow-up results are indicative of longer term clinical outcomes. Follow-up times beyond 12

months are not necessary to characterize the safety and effectiveness of these devices.

Thank you very much.

DR. BOYAN: Thank you.

Are there any questions from the Panel that they would like to address specifically?

Dr. Yaszemski.

DR. YASZEMSKI: Dr. Dawson, may I ask, is your feeling that the 12 month follow-up is an appropriate follow-up specific to anterior fusion devices? Since this is a general discussion about 12 versus 24, do you feel that way also for posterolateral fusion devices? Or is this specific to those that are anterior?

DR. DAWSON: Yes. I think it is the whole biological process of spinal arthrodesis. At 12 months, they are either fused or they never will without further intervention.

DR. YASZEMSKI: Thank you very much.

DR. BOYAN: I will call the next speaker up here in a second. We just had a change in plans. We have decided because of the number of speakers, that we will, instead of having each speaker have questions addressed to them at the time that they speak, will hold all questions

until the end of all the speakers and then we will have a general discussion, plenty of opportunity for open discussion.

So, our next speaker is Dr. Kahanovitz.

DR. KAHANOVITZ: Good morning.

My name is Dr. Neil Kahanovitz and I am the second vice president of the North American Spine Society. I am a research consultant in the area of electrical stimulation for ElectroBiology Incorporated.

The North American Spine Society represents over 1,700 physicians involved in the care of spinal disorders and is the largest such multidisciplinary organization in the world. On behalf of the North American Spine Society, I would like to address two issues facing the expert panel.

The first is the review of the medical devices by the expert medical panel. I would like to make it very clear that I am not here today to support or detract from any of the medical devices being reviewed. The North American Spine Society relies upon, as does the entire medical community, the work of the expert medical panels to advise the FDA on the safety and effectiveness of new medical devices.

In the past, the North American Spine Society has

worked closely with the FDA and would welcome further efforts in the future, but it is solely up to the judgment of these expert medical panels and the FDA to ascertain when a medical device is save and effective for the American patient.

However, the other issue, which is not as straightforward is the matter of clinical follow-up times. On June 12, 1995, I was present at the Orthopaedics and Rehabilitation Devices Advisory Committee Panel meeting. In response to a previously submitted request to clarify the duration of clinical follow-up times and whether a mandatory follow-up time was required, Paula Wilkerson, the executive secretary, made it clear that there was none and went on to state that follow-up times are determined on a device-by-device basis.

Unfortunately, the perception that a mandatory minimum follow-up time is required for new medical devices is still prevalent. It is the position of the North American Spine Society that follow-up times should and must be determined on an individual basis, not arbitrarily set without regard to the uniqueness of a particular medical device nor the clinical setting in which that device is to be used.

Follow-up time should be based on the expert medical panel's requirements to adequately assess the safety and effectiveness of a particular medical device regardless if that time frame is determined to be six months, one year, two years or even three years.

The intent of the recently passed FDA Modernization Act is to bring medical devices to the American public in a more timely manner while still assuring their safety and effectiveness. Abolishing any notion of arbitrary follow-up times is clearly one way to implement the spirit of this legislation. It is our hope that the FDA will continue to work closely with the investigating physicians early in the development of IDE and PMA trials so that appropriate follow-up times can be determined for each specific medical device, thereby avoiding unnecessary delays in the approval process.

The North American Spine Society would hope that the FDA will clarify any confusion that currently exists over mandatory clinical follow-up times. Each medical device must be evaluated on the specific safety and effectiveness merits of that individual device. Clinical follow-up times must ensure that the expert medical panel is able to adequately guarantee that medical devices are safe

and effective, but the use of mandatory and arbitrary follow-up times can only serve to delay the approval process and ultimately the access to these medical devices by the American patient.

Thank you.

DR. BOYAN: Thank you.

Our next speaker, please. John Dichiara.

DR. DICHIARA: Good morning. My name is John Dichiara. I am a representative of OSMA, the Orthopaedic Surgical Manufacturers Association. It was formed in the mid-1950s, consists of manufacturers of medical devices used in orthopaedic procedures.

OSMA's primary concern is the patient. OSMA supports the development of performance standards when needed to help assure reasonable safety and effectiveness through participation in national and international standards development organizations.

Another goal of OSMA is to make orthopaedic devices available for patients' needs with minimal delay or cost impact because of government regulation and has been involved in the regulatory process since the early 1970s.

OSMA also works closely with certain health care professionals to ensure continued improvement in orthopaedic

devices. Currently, there are 23 members in this organization.

With regard to clinical follow-up, which is one of the subjects before the Panel, OSMA believes that no hard and fast rules can be applied that would be appropriate for all devices. For example, devices that are intended to act as internal splints until bony union occurs are not intended to function for two years. Bony union usually occurs in less than six months. Once it has occurred, the device is redundant.

Most orthopaedic textbooks and manufacturers labeling for these types of devices recommend removal of the device once bony union has been accomplished. Additionally, we believe that in the past the FDA didn't always require two year follow-up. For example, we believe that bone growth stimulators used to facilitate bony union were not required to be followed for two years because if they had to be used for that long on the patient, they would not be considered to be effective.

OSMA recommends that the FDA seek guidance from those orthopaedic professionals experienced in the use of the devices as to what is a reasonable length of clinical follow-up to assess the safety and effectiveness of these

devices.

Thank you.

DR. BOYAN: Thank you.

Our next series of speakers will represent Sofamor Danek. The first speaker will be Dvorah Richman.

MS. RICHMAN: Good morning. I appreciate the opportunity to address the Panel today. My name is Dvorah Richman. I am with King and Spaulding. I am a paid consultant to Sofamor Danek. I do not have any financial interest in the company or any device company.

The determination of appropriate clinical study follow-up times involves complex medical considerations. Accordingly, I don't presume to suggest to the Panel how such determinations should be made. Rather my role today is to make clear that there are no legal restrictions on the Panel or FDA in this regard.

There have been ongoing discussions for some time regarding study follow-up requirements. In 1995, at a meeting of this panel, FDA articulated its position for the record. Quoting from the transcript, "The duration of study follow-up time needed to evaluate the safety and effectiveness of the multitude of orthopaedic devices regulated by FDA vary substantially.

"The device's application, its material and design properties and the endpoint of the clinical trial all affect the duration of the study. Therefore, the length of a clinical study is device specific with the general intent of requiring the minimum duration of follow-up needed to recognize critical safety and effectiveness factors."

This case-by-case approach is evident in PMA approvals for implantable devices having less than two years of clinical follow-up. This approach is also fully consistent with FDA's PMA regulations, which do not establish a minimum clinical study follow-up period and the Agency's PMA manual, which clarifies that, "Patients must be followed up for an adequate period of time to assure both lasting effectiveness, if that is claimed, and continued safety."

This flexibility is appropriate because PMA applications, even for similar devices, invariably contain different kinds of safety and effectiveness data and information, including clinical data from studies with different study designs.

You will probably hear today that FDA is obliged, based on past Agency practice, to apply a two year clinical study follow-up requirement to all PMAs for implantable

orthopaedic devices. No such obligation exists. There are no Agency regulations, guidelines or official FDA publications or authority citing to a two year requirement. Indeed, as noted earlier, FDA has told this panel just the opposite, that study length is device specific, based on application of the device, material and design properties and clinical trial endpoint.

Any expressions by individual FDA staff members concerning two year follow-up are not binding on the advisory panel or the Agency. FDA thus retains the discretion to approach each product application on a case-by-case basis. Accordingly, any industry expectation that FDA will for the foreseeable future require a two year follow-up is simply not warranted.

Moreover, claims of dissimilar treatment between companies should be seen for what they are, attempts to preserve a competitive advantage. The only real question for this panel is whether a PMA demonstrates per the Food, Drug and Cosmetic Act's PMA provisions, that there is a reasonable assurance of safety and effectiveness for the device.

In considering any PMA, the Panel and FDA should also be mindful of strong language in the new FDA law signed

by President Clinton on November 21st. This law, most of which becomes effective in 90 days, requires FDA to consider the least burdensome, appropriate means of evaluating device effectiveness and to determine whether data requirements can be reduced through reliance on postmarket controls.

All told, the Panel must be concerned only with scientific merits of the PMA being considered and whether the data included therein are sufficient to demonstrate a reasonable assurance of safety and effectiveness.

Thank you very much.

DR. BOYAN: Thank you.

I want to remind Sofamor Danek that you have ten minutes remaining. Our next speaker is Robert Rylee.

MR. RYLEE: Good morning. I am Bob Rylee. I am an employee of Sofamor Danek Group. I appreciate the opportunity to make a few brief remarks this morning on the subject of clinical endpoints for devices.

The subject has been raised as a result of two citizens' petitions filed with the Food and Drug Administration. I think there are two broad issues involved here. First, should a competitor be allowed to use this process in an effort to gain a marketplace advantage by raising a red herring issue.

In its written response to the petition, Sofamor Danek has dealt with both the legal and factual issues to reach the compelling answer of "no." In the interest of time, I will not be redundant by repeating those issues and those answers. However, I think it is important to recognize what is at play here.

When you strip away the verbiage from the petitions, they are simply an attempt to use the Agency and this panel in order to maintain a time competitive advantage in the marketplace. This is not the function of the Agency or this panel. The focus by the Agency and this panel should be on medical science, not competition. Competitive forces should work only in the marketplace not at this hearing.

To the Agency and to this panel, I would suggest you do not have a dog in this fight. So, don't pay attention to the barking.

The second issue, which I think is appropriate for this panel is to determine what is the appropriate time line or endpoint for the approval or clearance of a device. I suggest that the correct answer to that is "that depends." It depends on the device and its intended function. This is a medical judgment issue. The appropriate endpoint should

be when there can be objective determination that the device has reasonably performed its function.

It is up to medical experts to make that particular decision. The urging of a so-called rule of two years or 24 months is totally without medical or scientific merit. In the early 1980s, ASTM and later the International Standards Organization with the participation of a broad range of people, including Dr. Joe Davis of the FDA, Dr. Manuel Horowitz of the National Bureau of Standards, a host of orthopaedic surgeons, like Pat Lang, Vick Frankel(?) and on and on, worked on these issues.

The consensus then and still now is that there is no basis for a set, arbitrary time limit to evaluate the function of the device. Due to the vastly different nature and function of devices, the time line should be set according to the device and its function.

There have been standards published by ASTM and ISO relating to time lines for the studies of biocompatibility, but in all of the standards relating to devices, there is no identification of a set time line to establish the function. For example, we all would recognize there is quite a difference in the endpoint in determining the functional characteristics of a bone fixation device

versus a prosthetic device.

Typically, if the function is to aid in the healing of bone, fixation of bone, union of bone, that can be observed in a time period of from two to six months. If it goes beyond two months, it may go to union within 12, but as has been previously stated here, at 12 months, if union has not been obtained, the great probability is that it will never be obtained without additional intervention.

In conclusion, I suggest there should be no arbitrary time limits. Instead, the time line should be set for each particular device as a part of its protocol and that time line should be when there can be objective determination that the device has performed the function as intended.

Thank you.

DR. BOYAN: Thank you.

Our next speaker is Dr. Dawson, five minutes total.

DR. HEIM: Good morning. My name is Steve Heim. I am an orthopaedic spinal surgeon in the Chicago area. I do have a financial interest in Sofamor Danek, although I have served as neither a clinical investigator nor do I have any direct financial interest in any of the products

presented to the Panel over the next two days.

I do appreciate the opportunity to present an important concept regarding interbody cage devices in general to this panel. This concept is the ability to determine fusion in constructs employing interbody fixation devices.

I have had the opportunity to serve as a clinician in the development of three different interbody cage devices for different manufacturers over the past four years. I have used this type of device in over 75 patients in this past four year time period. As an orthopaedic surgeon in a practice limited to care of spinal disorders, I have over 12 years of personal experience following fusion patients and attempting to evaluate fusion success.

In the past, I have worked clinically in the development of titanium spinal implants and imaging techniques for posterorlateral fusion techniques. Historically, two year follow-up has been arbitrarily adopted due to difficulty in imaging bridging bone or fusion in the presence of stainless steel implants.

Typically, in the absence of fusion, such implants would fail within that two year time period. The key point, which I wish to leave with the Panel today is that there are

two important differences between the historical stainless steel instrumented posterorlateral fusion and the recent class of interbody cage devices.

First, interbody fusion is significantly different in terms of bone healing than the posterorlateral techniques. In the instance of interbody fusion constructs, Wolf's(?) law favors and expedites bone healing. What this means is that bone heals in compression of the inner body location and tends to resorb when exposed to distraction forces, the posterorlateral location.

Furthermore, the interbody location provides a consistently highly vascularized cancellous surface for bone healing when compared to the posterorlateral bone surface area. Secondly, over the past five to seven years, imaging capabilities for fusion have greatly expanded. This improvement has included imaging techniques, as well as the development of implants, which are much more favorable to imaging.

The overall result has been to greatly increase the accuracy of imaging, of bone healing and spinal fusion surgery. The thin slice CT scan with sagittal and coronal reconstructions coupled with the use of titanium instrumentation has become the most favorable monitor of

fusion. Specifically, in the case of interbody case devices, the autologous bone contained within the implants cannot be directly assessed with plain radiographs due to the fact that the implants are radiographically projected over the autologous fusion zone.

In this instance, the thin slice CT scan becomes the only direct method of visualizing the fusion mass. These scans, which employ 1 millimeter thick sections, allow multiple contiguous lyses through the fusion zone of each implant. In fact, it was by way of this type of CT study, including my specific cases shown to the Panel in May of 1996, for the spine tech PMA presentation, that ultimately led to the release of the BAK(?) device.

In my opinion, CT examination is the way to definitively evaluate fusion because of the sensitivity of the study. The historical two year arbitrary follow-up is no longer necessary, due to this capability of direct sensitive imaging of the fusion mass. The demonstration of successful bridging bone through and about the different interbody cage implants that I have had personal experience with has been reliably evident that the three to six month time point with no clinical or radiographic failures occurring thereafter.

It is for these reasons that I believe the current generation of interbody cage fixation devices permits us as clinicians to evolve beyond the old arbitrary two year imaging to determine fusion, looking for indirect evidence of pseudarthrosis(?) by way of implant failure.

Thank you.

DR. BOYAN: Thank you very much.

Any speakers from Sofamor Danek, 42 seconds. Hearing none, we will move on to U.S. Surgical Corporation. The first speaker will be Lou Mazzarese.

MR. MAZZARESE: Thank you and good morning, ladies and gentlemen.

Can I get some assistance on just a few overheads, if possible?

My name is Lou Mazzarese. I am vice president of quality and regulatory affairs at United States Surgical Corporation. With me this morning, I am joined by Dr. Scott Blumenthal from Texas Back Institute and Dr. Tushar Patel from the Yale School of Medicine.

I would like to open with a few brief comments from the regulatory standpoint of the debate and discussion regarding two year follow-up data. Clearly, in terms of BAK fusion cages, the practice so far to date has been that two

data are important and required to assess the proper amount of safety and efficacy data for these kinds of specific devices. That has been the practice. That has been recognized as the standard for these particular spinal implants.

Now, unlike some of the speakers who came before me, I was not at the June 1995 Panel meeting, but I marvel at how the transcript can be used by such divergent opinions to support each case. Nevertheless, our reading of that transcript clearly sees a restatement of the need for those two year data endpoints for this specific class of devices because it represented a proper balance between short term safety and longer term efficacy, i.e., two years.

Even though some comments were made that three to five years would be nice, the Panel and the FDA realized that could very well delay the introduction of these important new technological devices to patients and, therefore, two years was struck as a proper balance.

We certainly don't feel that today there is any change warranted from that position.

Unlike the June 1955 Panel meeting, I was at the May 1996 Panel meeting and at that Panel meeting, again, the two year recognized practice was seen with the approval of

two PMAs, one from SpineTech and the other from U.S. Surgical Corporation. There was and continues to be, in our opinion, a rationale for maintaining the two year requirement. And it is very simple.

These types of devices are not generic. They do achieve their success criteria at different times and I think that represents one of the most important elements in this discussion as to why two year data at this point should be maintained.

These devices will be implanted in thousands of patients and it is important to really look at how much data we have before we change a practice that has been observed for the last several years vis-a-vis the two year data requirement. If we are to reduce that two year data requirement, then I think a number of things need to be noted.

First and foremost, as has been said by all of the previous speakers, it should be done based on a careful consideration, comprehensive review by clinical experts of the need.

Number two, it should be done in an implementation scheme that is prospective in nature, new IDEs, new PMAs.

And third and fourth, the economic issue is

absolutely very real. We certainly feel that a level playing field should be maintained primarily because to not do so, to allow the regulatory process not to maintain that field serves as a very strong disincentive for research in the future.

Finally, there is nothing in the current Food and Drug, Cosmetic Act or the new FDA Modernization Act which would allow FDA to either actively or passively use clinical data from one PMA in its effect of approving another. That only can be done when certain criteria are met and in this particular case existing law or in the new law, those criteria have not been met.

So, we urge to be aware of these facts and please maintain that two year data point for now.

At this point, I would like to ask Dr. Blumenthal to make his comments.

Thank you.

DR. BOYAN: Thank you.

DR. BLUMENTHAL: My name is Scott Blumenthal. I am a faculty member for the Surgical Dynamics Corporation and a minor stockholder at U.S. Surgical.

You have also heard earlier from one of my associates at the Texas Back Institute, who must be a lot

smarter than me because he figured out he could send in a letter.

I am going to direct my comments towards that of a clinician. Basically, what we are talking about here is the guy out in the street, the orthopaedic surgeon, the spine surgeon. If you ask him what is an adequate assessment for a clinical study, two years comes to mind.

It is basically purported by the Journal of Bone and Joint Surgery, representing the AALS and the American Orthopaedic Association; specifically in spine surgery, the Journal of Spine requires this for clinical study. They represent multiple spine societies, including the North American Spine Society, which we have heard from already this morning.

The FDA in the past has represented that this two year follow-up for similar type devices has been the standard and I would see no reason to deviate from this. The fact that a double blind exists does not obviate the need for two year follow-up as long term results, complications, including bio-interactive interactions can also occur in that one to two year period.

If you were going to purport the opposite point of view, that is, is there a circumstance under which less than

two year follow-up may be adequate, well, that is a possibility for devices that have a long term standing record, those where certain characteristics and design characteristics have been agreed upon, but, frankly, the cage technology is not there yet. It is really too new.

There are significant design characteristics between the cages that are out now that do make a difference in terms of long term results and different types of complications that we see; thread design, occluded pitch, wall thickness, porosity and the use or lack of use of end caps(?) all take -- all are factors that we look at in long term follow-ups.

Also unique to the BAK interbody cages, and this is a very important point and why I keep talking about potential complications, is that contrary to what we have heard, there is no real agreed-upon method for assessing fusion. It is very difficult to assess when these cages are healed. The fact of the matter is is one of the factors that we look at are failure of cages to achieve fusion, that be the lucent line, the motion on flexion/extension in any type of migration or subsidence.

The fact of the matter is, and, frankly speaking, six months is not long enough to be able to determine

whether these are going to occur. Twelve months probably isn't and 24 is probably the minimum to see these type of potential complications.

At this point, there is no precedence for a lesser period of follow-up for clinical studies or bringing new products to market in the field of orthopaedics or specifically spinal implants. If this is, indeed, the goal, then careful due diligence needs to be done with well-scrutinized data that would support a change in that position.

At this point, I am going to turn it over to Dr. Patel, who is chief of the Spine Surgery Section and professor at Yale.

DR. PATEL: Good morning. My name is Tushar Patel. I am assistant professor and chief of the Section of Spine Surgery in the Department of Orthopaedics at the Yale School of Medicine.

I would like to thank you for the opportunity to address this panel. As an academician --

DR. BOYAN: Do you have any financial interests?

DR. PATEL: Sorry. I am a consultant to the company and I have no financial interest.

As an academician and a practicing clinician, I

feel that a two year follow-up should be required for premarket approval of new orthopaedic implants that are substantially different. This is not just my opinion but a requirement of the leading orthopaedic scientific societies and peer reviewed journals where our clinical research is presented.

One cannot underestimate the power of the placebo effect in the short term follow-up period. Every patient would like to believe that they will feel better as a result of the surgery performed upon them by their surgeon in order to relieve their pain. In the case of the BAK infusion cage, 76 percent of patients were fused at six months, at 12 months, 83 percent and 68 percent for the BAK patients. At two years, 92 percent of the BAK caged patients and 74 percent of the BAK patients were fused.

These numbers at 6, 12 and 24 months raise two major points; namely, that the fusion rates differ with longer term follow-up; 76 percent at six months compared to 92 percent at two years.

There are two similar devices, both cylindrical titanium, threaded fusion cages, that have significantly different fusion rates have two years; 92 percent versus 74 percent. Although the devices may appear similar,

biomechanical differences in terms of stiffness and stability, as well as technical differences in terms of instrumentation for implantation, bone grafting, end caps, thread design and porosity all contribute to these differences in fusion rates in the near term as well as differences in clinical outcome.

These factors may also account for differences in later adverse effects and complications. These later implant complications, including implant failures, such as migration and loosening and host failures, such as settling and bony failure need to be sought, observed and followed diligently.

This is the importance and role of postmarketing surveillance. These new orthopaedic devices will be implanted by hundreds of surgeons in thousands of patients for the rest of their lives in close proximity to vital structures. Cages are not designed to be removed.

The two year follow-up recommendations of the FDA and the requirement of our academic societies in peer reviewed journals should not be changed arbitrarily without much thoughtful deliberation and careful review of solid scientific data that may irreversibly impact thousands of patients.

Thank you.

DR. BOYAN: Thank you very much.

Any other comments from U.S. Surgical?

[There was no response.]

Let's go then to SpineTech and the first speaker from SpineTech will be Richard Jansen.

DR. DOWDLE: With the Panel's approval, we are going to change the order. I am Dr. John Dowdle.

Members of the Orthopaedic Advisory Panel and representatives of the Food and Drug Administration, ladies and gentlemen, my name is Dr. John Dowdle. I am a board certified orthopaedic surgeon from St. Paul, Minnesota. I have been doing spine surgery for the last 20 years. I do have a financial interest in SpineTech.

I have seen first hand the evolution of spine surgery over the last 20 years. I have been in practice when spine plates, pentacle screws and now intervertebral fusion cages were introduced and became widely used. I personally have used all these devices in clinical practice and followed patients who have had these operations and seen the results.

I was part of the SpineTech-FDA study and was one of the clinical investigators. I have personally done 350

fusion cage cases since the beginning of the study in March of 1992. The subject of discussion today is the length of follow-up necessary to determine the clinical outcome, safety and efficacy of intervertebral fusion cage devices.

In clinical practice, I want to know -- I want to follow a patient from the beginning of the episode of care to final resolution when they are discharged, either cured or medically stable and in need of no further treatment. I want to know their ability to work, their functional status, their complications and the necessary treatment to resolve any of those complications.

There is more at issue than just whether the bone fuses. If the bone fuses and there is no pain relief, no improvement in function or no return to work, should we really be doing these procedures? Success is not just bone fusion.

At issue today is whether one year follow-up data is adequate or a minimum of two years of data necessary to determine the safety and efficacy of these devices. The answer is it depends on what you want to see. If you want to see the results of uncomplicated cases and how they progress to fusion without difficulty and return back to work without a problem, one data will give those cases and

their results. It will not tell the outcome of the complicated cases where fusion was delayed or not complete and another surgery was necessary to complete the fusion.

The results of most of these cases will not be known -- will be known by two years. Two year data will give you a much better view of the overall results of this procedure. One year data will not tell you if the patient was able to complete the exercise program and return back to work to his or her pre-injury job and the overall functional result.

Two year data is much more likely to show the functional results of this surgery. Most of the reference journals will not accept for publication studies that do not have at least two year follow-up just for this reason.

Early data may not give you the clear picture. Two year data is a much better view of the overall results. One year data will not tell you the long term results of the medical or surgical complications that can occur from this procedure; the vascular injury, venous thrombosis, venous occlusion, the bone fracture, the migration of the implant or fracture of the implant.

Each of these complications can happen and have the potential to severely impact the quality of life of the

individual patient. As clinician, I want to know the potential risks for my patients and the potential long term solutions. These are the exact cases that need to be absolutely be understood and not ignored because the follow-up was too short.

As a clinician, I need to know. The patients have a right to know.

In summary, by requiring two year data, the FDA is acting in the best interest of the patient to ensure save and effective medical devices. It is clear that one year data for follow-up of this new medical technology is not adequate to ensure that these devices are truly safe and effective for all patients.

Thank you.

DR. BOYAN: Thank you.

MR. JANSEN: My name is Rich Jansen. I am vice president of regulatory and clinical affairs at SpineTech in Minneapolis, Minnesota.

The November 25th Federal Register stated that the FDA intended to raise with this panel the important public health issue of follow-up period necessary to evaluate safety and effectiveness of intervertebral spinal implants. The key to resolving this issue is the evaluation of the

experience of this type of device, the science that might support a change in follow-up requirement, the relative newness of intervertebral implant designs for the lumbar spine and the difficulty with which a sustaining and durable fusion can be determined clinically without a two year follow-up period.

Based on the continuing interaction with FDA and our view of relevant medical science, we strongly believe that a substantial portion of clinical data from a study evaluating intervertebral implants intended for lumbar spine must represent at least two years of follow-up, notwithstanding study design to ensure safety and effectiveness.

Although there is no regulation or two year rule that says that spinal implants need to be followed for two years, there certainly is substantial precedent with this. Since the Food, Drug and Cosmetic Act of 1976 went into effect, only two spine products have received PMA approval through the Orthopaedics Branch. Both of these were fusion cages; both of them last year.

SpineTech submitted its PMA on the BAK interbody fusion system in January 1995. That included one year data on 254 patients and two year data on 32 patients. FDA

refused to review this data and insisted that additional two year follow-up data for intervertebral implants were necessary to assure safety and effectiveness.

FDA's stated it was unable to conduct substantial safety and effectiveness without two year follow-up. In April 1995, a supplement to that original PMA was submitted to FDA with two year clinical data on 96 patients and fusion data on 69 patients. Again, we were told that 69 patients was insufficient two year follow-up.

In May 1995, we submitted another PMA supplement that included two year data on 123 patients and two year fusion data on 110, with a fusion success of over 90 percent at that point, considerably better than the historical literature controls that was used in that study, we were told the PMA could not be accepted for filing because there was insufficient two year data.

Again, in August 1995, SpineTech submitted another PMA supplement with clinical data on 160 patients at two year evaluation point and fusion data on 142. Finally, the FDA filed our PMA and began a substantive review. Previously, according to FDA, the data was inadequate to support either a safety or effectiveness judgment.

FDA was so adamant about the necessity for two

year follow-up that the Agency after filing our PMA requested additional two year data. With additional two year patients added for the study, this orthopaedics advisory panel reviewed two year follow-up on 284 patients with fusion results on 267 patients.

This was a 94 percent follow-up of all patients that had reached that two year mark. FDA's persistence in recently requiring two year data from SpineTech reflects medical and scientific considerations that data of that duration are necessary to reasonably assure safety and effectiveness.

To continue on this point, SpineTech submitted an IDE, including a randomized concurrently controlled trial for a new intervertebral, spinal intervertebral investigational device in July 1997. That is less than six months ago. We received a letter dated August 1, 1997, from FDA's Orthopaedics Branch denying this IDE. Although there were multiple reasons for the denial, one of the reasons for the denial was that we were told that all patients in the treatment and control group in the state of the art study designed required two year follow-up.

It is clear that the evaluation of safety and effectiveness of an intervertebral spinal implant still

requires two year follow-up, according to FDA's August 1997 determination.

Second, in 1994, FDA's Orthopaedics Branch took the position that all IDE studies should have three year follow-up. As a response to this, on June 12, 1995, an orthopaedic advisory panel meeting spent approximately four hours discussing the appropriate length of follow-up when considering a change from three year follow-up to two year follow-up.

The common wisdom was that all orthopaedic implants generally should have two year follow-up. Intervertebral spinal implants present a particular compelling case for two year follow-up at a minimum. Comments at the June 12 panel meeting support this view. For example, Dr. Ted Wendt(?) representing HIMA, the Health Industry Manufacturers Association, stated, and I quote, "The premarket clinical study is of real value in establishing minimal indications for use and to detect design problems, which result in precipitous failure. The two year follow-up time, which has been an IDE PMA requirement for several years, appears to meet this goal quite effectively.

Also, FDA's Mr. Mark Melkerson, branch chief of

Orthopaedics, stated that the FDA expects 85 percent follow-up and that the appropriate time period in general for spine has been two years.

Toward the end of that meeting, a panel member summarized the discussion with the following quotation: As I look at the studies in the literature and the one we saw this morning and Dr. Boran's(?) study and many others that I am aware of, it looks like the initial healing process is over the first two years and by the end of the second year or into the second year, there is clearly a leveling off that goes on for four to five years almost no matter what implant is used. So, it seems like the first decision point is logically two years."

Notwithstanding the common wisdom applicable to all orthopaedic implants, the newness of intervertebral spinal implants and the difficulty in assessing its successful durable fusion in less than two years provides compelling logic for a two year minimum follow-up for intervertebral devices.

Third, study design does not determine the length of follow-up for patients for receiving intervertebral implants in the lumbar spine. In other words, randomized trials do not change the required length of follow-up.

Because of the need for a two year follow-up to determine clinical success, study design cannot shorten this minimum requirement.

Finally, we must rely on scientific evidence. In the past, when SpineTech has proposed shorter length of follow-up for IDE studies, the FDA has asked us for scientific evidence to support that position. We have done a literature search and have been unable to find any scientific evidence supporting one year follow-up for lumbar spinal implants.

Additionally, neither FDA nor this panel may rely on data from SpineTech or U.S. Surgical's improved PMAs. The Safe Medical Devices Act of 1990 states that after approving PMAs for four devices of a kind and the passage of one additional year, the FDA may use the data from approved PMAs. As of February 20th, 1998, the Food and Drug Administration Modernization Act of 1997 will only permit FDA to use data for specified purposes six years after a device approval.

Simply put, neither FDA nor the Panel can rely on the two previously approved intervertebral spinal implant PMAs to draw any conclusions about similar devices. Of course, this applies to the devices being reviewed today.

I would like to conclude my comments about clinical study follow-up for intervertebral implants by stating that the issue being discussed presents significant safety and efficacy concerns. Unfortunately, the opportunity to fully air follow-up issues has been limited by inadequate notice for this meeting.

Notice for this meeting appeared in the Federal Register on November 25th, leaving only 11 business days from that day until this meeting. As a result it is impossible to bring before this committee certain leading medical and scientific experts who had prior scheduled commitments.

I believe it would be inappropriate to make a decision to change the length of follow-up for intervertebral spinal implants for several reasons. As stated, there has been a well-developed policy from the FDA regarding the need for two year follow-up for intervertebral spinal implants.

For this type of spinal device, there is no scientific evidence supporting a shorter period that will assure safety and effectiveness. This kind of issue cannot be reasonably addressed on short notice and one hour presentation. Moreover, a policy issue like this should not

be evaluated in the context of other PMAs.

I would encourage the Panel to conduct additional meetings in the coming months to obtain input from all interested parties, including all persons with appropriate expertise. At these meetings, patient follow-up for intervertebral spinal implants and other orthopaedic implants should be discussed.

To address current study guidance on short notice and in an ad hoc manner is not fair to patients or product sponsors and the public health requires full and open exposition on this issue.

It was stated that the Modernization Act of 1997 does request the use of least burdensome data. However, just eight days ago, Dr. Bruce Burlington stated that this should be established on presubmission meetings. To change its science-based policy with only limited discussion and to apply this retrospectively is inappropriate.

At this point, fusion cage technology is very young. The first post-approval studies requested by this panel of U.S. Surgical and SpineTech have not yet been completed and the first independent research by a non-IDE investigator has yet to appear in a major journal.

I would encourage the Panel to continue with the

policy of two year follow-up until a scientific review of the different product classes is complete.

Thank you for your consideration.

DR. BOYAN: Thank you.

I would now like to invite any other individuals in the audience that might like to address the Panel. Are there any other individuals that would like to address the Panel?

[There was no response.]

Okay. You have to speak now or forever hold your piece. All right. Since there are not other requests to speak in the open public hearing, we are now going to proceed to an open committee discussion.

I would like to introduce Dr. Celia Witten, division director of the Division of General and Restorative Devices to provide an opening statement, which will lead us into our first discussion of the day.

Agenda Item: Discussion of Clinical Endpoints

DR. WITTEN: Good morning. I will provide a short introduction to the question that is being presented this morning for Panel discussion on which we have already heard a number of comments this morning. And as I think is clear from this morning's discussion, the FDA and the orthopaedic

community have grappled on an ongoing basis with the most appropriate time and best method to assess success or failure of the orthopaedic implant.

Over the past several years, we have had several Panel meetings in which the Panel has discussed questions related to length of follow-up needed to obtain important clinical information. We are asking you today for a discussion of the appropriate duration of follow-up for marketing authorization for spinal implants.

I would like to mention that there have been a number of comments this morning relating to specific products and follow-up relating to specific products. We are asking for a general discussion of the considerations that would need to be taken into account in assessment of spinal implants overall.

We have previously accepted an assessment of spinal implants at 24 months. We are being asked by industry to consider the appropriateness of a one year assessment of these devices for the purpose of marketing authorization. Many of you are aware that we have two petitions before us regarding this time frame.

Today, we are asking you for a discussion of this question and the scientific issues that a one year

assessment would raise. Can information obtained at an earlier one year time point tell patients and clinicians about longer term expectations for the device?

What issues specific to the clinical indication and the product might need to be considered? For example, the expected benefit of the fusion device might vary depending on the area of the spine in which it was used and the expected time for the patient to receive the full benefit might also be different.

Different disease processes might call for evaluation at different times and, if so, some discussion of those different indications would be helpful. Different devices for the same indication may have a different mechanism and there may be different time frames to assess safety and efficacy for different device types.

We are also asking for a discussion of assessment -- of how safety and efficacy assessment should be performed at an earlier time. Safety endpoints for spinal studies have generally included such things as nature and frequency of complications, surgical interventions and unanticipated adverse events.

Is there important safety information that might be likely to be missed in a study with one year follow-up?

Are there markers that may predict later complications that would need to be considered as part of the safety assessment for a study with one year assessments?

We have also considered effectiveness endpoints to be fusion, pain, function and neurologic status. These are the types of questions that we have regarding these endpoints. What are the appropriate predictors of long term clinical success that can be assessed at one year? For example, what degree of pain relieve at one year predicts long term pain relief and is this different from the minimum pain relief at two years that would be felt to predict long term success?

For assessment of fusion, what consideration should affect selection of measurement tools? For example, can flexion and extension radiographs delineate fusion at 12 months? What percentage of patients would need to be fused at that time to provide adequate information to the clinician about the likely success of fusion of the device?

I am going to read the question since I see it is a little bit hard to read from here, for the audience.

We have previously accepted safety and effectiveness assessment of spinal implants at 24 months. The following safety and effectiveness endpoints have been

used at 24 months to study implants. Safety endpoints include nature and frequency of complications, surgical interventions, any unanticipated adverse events. Effectiveness endpoints include fusion, pain, function, neurologic status.

We are being asked by industry to consider the appropriateness of premarket approval of spinal implants based on 12 months rather than 24 months follow-up data. We are faced with a number of questions.

Question No. 1 -- what are the issues that would need to be considered in evaluating data at 12 months for a marketing authorization?

Question 2 -- what are the appropriate predictors of a long term successful intervention at 12 months? Are the endpoints and assessment tools we have used at 24 months appropriate for use at 12 months?

These are complex questions and I would like to say in closing, we recognize they don't have a simple answer. We appreciate your help in assisting us with a thoughtful approach to this question. We recognize that the time is short but since we are faced with the question today, we are opening our discussion of this very complex topic this morning.

We are not asking for a vote on this question today, but a discussion of the scientific issues that a one year assessment of a spinal implant would present. This may be one of a series of meetings in which we ask for your input on study designs for orthopaedic implants.

Thank you very much.

DR. BOYAN: Thank you, Dr. Witten.

I would like to remind everyone that at this point the Panel may ask for anybody in the audience, any of the previous speakers to answer a question, anybody from the FDA to answer a question, but the audience themselves cannot come forward and speak without being specifically requested.

So, we will now begin our discussion of the first -- of this general question. What I would like to do just to get this discussion going is to -- I have another business item. All Panel members remember you have to speak directly into the microphone and please state your name so that the transcriber can learn to recognize your voice with your name.

All right. Let's just begin and we will go around the table. We will start with Dr. Wilkinson and each of the Panel members please ask one of the comments that -- or make a comment that is just burning, the most important one that

you would like to get out short term.

DR. WILKINSON: You are asking for a statement and not just questions?

DR. BOYAN: No. Actually, I would like just questions now at this point, but if your question has a preamble, feel free to preamble.

DR. WILKINSON: I will hold the questions, but do want to make a statement.

DR. BOYAN: Okay.

DR. WILKINSON: Dr. Wilkinson.

DR. BOYAN: Dr. Laurencin.

DR. LAURENCIN: I hold questions. I can make a statement later.

DR. BOYAN: Okay. Back to Dr. Wilkinson. Go ahead and make your comment. Okay. I see what you guys are thinking. All right.

DR. WILKINSON: All right.

Harold Wilkinson, professor of neurosurgery and in that role, I teach neurosurgical residents, among other things, to operate on brain tumors, as well as on spinal problems. My residents are frequently upset about the way I approach spine surgery.

As you know, to teach a resident, the resident has

to be given progressive responsibility to do more and more of the case as they learn. Spine surgery is one of the last things I turn over to the resident.

The reason is the brain tumor patient will be dead in two years. The spine case will live. So, what you do to the spine patient has to be good and good for a long time. Now, there is adequate scientific data and clinical data that once hardware has been implanted in the spine, disc spaces narrow progressively and spinal hardware loosens progressively and that process often seems to accelerate during the second year.

As the device loosens and the interspace collapses, the patient becomes increasingly dependent on the bone fusion and no longer is the mechanical stabilization of the device sufficient. If the device in itself blocks fusion bone, if the device itself provides stress shielding to the point that the bone fusion is inadequate, that inadequacy of fusion will not be apparent until the device itself has loosened.

The simple presence of fusion bone or bridging bone is not enough. The fusion must be strong enough and over a large enough area to support the patient safely and comfortably and to allow that patient to improve

functionally.

How to determine that probably requires not only good quality CTs with reconstructions, but in my opinion, at least, certainly requires evaluable bending radiographic films. If there is no bending at adjacent levels, those films are not evaluable. So, you have to show that there is stability of the fusion construct.

What I teach my residents on this is summarized in a slide that I frequently show when I am asked to lecture about spine fusion. The slide simply states steel does not heal. The steel can help the fusion bone. It can also impair the fusion bone.

In my clinical practice, unfortunately, I wrote a book on the failed back. So, I see a disproportionally large number of patients, who continue to have symptoms after spine surgery. And I am seeing an increasing number of those patients more than one year beyond spinal instrumentation, more than two years beyond spinal instrumentation, who are sent to me because they are failed back patients.

My problem is to identify, of course, whether the fusion has anything to do with that, but quite often it does and that may not become apparent until well beyond the first

year.

The two year follow-up that we are asked to address I would view as a scientific standard for spinal implant support devices, but not necessarily a legal standard. As a scientific standard, I think there is more than overwhelming evidence that one year follow-up is simply not enough.

DR. BOYAN: Thank you, Dr. Wilkinson.

Dr. Laurencin.

DR. LAURENCIN: I just wanted to make a few brief comments from the perspective of someone who is a clinical orthopaedic surgeon and also a person heavily involved in biomaterials research.

First, if we look at the current standard of two years, we know well why that standard exists. We have examples if we look at areas such as ACL protheses and we have seen great data at one year. We have seen disastrous data at two years. These devices have sometimes been approved for use in the market, only to be taken off after long term data has occurred.

If we look at the areas such as the TMJ devices, other devices, again, the situations in which one year data looks good and the long term data looks very bad exists.

Now, what about in terms of orthopaedics, in terms of the spine? Well, you know that the requirement in terms of the literature, in terms of JBJS of the spine is a two year point. From my reading of the literature, I am not sure there is really anything in the literature that supports moving back to a one year time point in terms of follow-up data.

If we look at the data that has been presented in a number of the studies both here and also in the literature, we can see progression of fusion after the one year time point and see differences in fusion rate between one year and two year, again, supporting a longer term time point.

Now, in terms of spinal cages, I think that one point that was made by someone -- one of the speakers -- I think is very true. The devices that we are reviewing and looking at even today have designs that are very, very different and to suggest that a cage, quote, unquote, device will have a fusion rate -- a fusion that will be set at one year belies the fact that these devices vary in nature and should vary in results.

So, that in summary is all my thoughts about this. I do believe that this is something that the FDA should get

involved with. There was a statement that said that perhaps the FDA should not be involved with making this sort of determination or being a part of this discussion. I believe this is a discussion that the FDA should be a part of.

DR. BOYAN: Thank you.

Dr. Janosky.

DR. JANOSKY: It is actually a comment with a series of questions or at least one large question with a series of parts for it.

It seems to me that the issue is actually the issue of prediction and whether there has been a study looking at 12 month as predictor of a 24 month outcome clearly, in terms of indication for use, materials used and endpoints when you investigate it.

So, in terms of a question, has there been such study done? I did a quick review of the literature and haven't been able to find any either for cages currently on the market or -- it is an open question.

DR. BOYAN: I would like to address that question to anybody in the audience that can answer it because that may be privileged information that FDA -- yes, sir, in the black suit. Would you identify your name, who you are and any financial relationship you might have with any of the

discussion -- companies under discussion.

MR. MICHAELSON: Yes. My name is Gary Michaelson and I am the inventor of this technology and I have extensive financial involvement in this. I licensed the technology, both to SpineTech and to Danek.

I think that Dr. Patel alluded to statistics that showed that the fusion rates with these devices at one year differ only from the fusion rates at two years by the fact that none of the cases that are fused at one year unfused, but there is an additional 5 percent fusions that are detectable at two years.

I think that Dr. Dawson pointed out that that is not really that cases go on to fuse between one and two years. It is that in the SpineTech data, it lacked the sophistication of 1 millimeter thin sliced CTs that show cross trabeculation. So, what they were doing was looking for the much cruder signs of failure to get fusion, which is gross motion with flexion/extension or lucencies around the device.

But if, in fact, rather than looking for signs of failure, you can actually show by more sophisticated means of visualizing the fusion that it is fused, you can pick up those fusions earlier and the statistics that Danek now has

proves that point.

Does that answer the question?

DR. BOYAN: Yes, thank you. I just -- you have raised the one issue -- and I need to be careful with this. This is meant to be a general discussion and, so, it is important that we eliminate any mention of specific companies from the discussion because we are trying to get at a scientific question here.

Thank you very much.

DR. WITTEN: I would like to add to that just a reminder that we are not just talking about one type of device. It is not just one sponsor. We are not talking about one type of device, but spinal implants in general.

DR. JANOSKY: So, it seems to me that it is fair to come to the scientific conclusion that there has not been a scientific investigation as to whether 12 months, irrespective of cage type, outcome, et cetera, is predictive of 24 months.

DR. BOYAN: I think we won't go back to the audience, but as we go around, I would like to also raise another question and I think it is a follow-on to your question, which is is there a predictor at one year or two years that is a predictor of what happens at five or six

with this technology.

DR. JANOSKY: Exactly.

DR. BOYAN: Okay. Back to you.

DR. JANOSKY: That seems to be the paramount issue, at least to me. That seems to be the issue we are dealing with.

DR. YASZEMSKI: My name is Michael Yaszemski.

I think that Dr. Boyan made a nice lead-in to the -- I guess I will call it an issue and ask for comments from anybody in the audience who might like to comment on it.

DR. BOYAN: Wait, wait, Michael. We are not going to let anybody comment, maybe one or two.

DR. YASZEMSKI: Thanks, Barbara.

When we are talking about outcomes, I think perhaps we should at least consider the difference between the patient versus the -- what I would consider a measure of the effectiveness of the device and that is to see whether a fusion happens and to separate the concept of a fusion happening versus the concept of whether the patient considers their clinical outcome successful, because I think those are two things.

I think for considering the patient's clinical outcome to be successful, we are asking questions, such as

is an anterior lumbar interbody fusion for the specific disease process that the patient has effective in making that patient feel that they have been helped in return for the risks they have undergone with surgery.

There are many different ways to get that fusion to happen and from my perspective, for this discussion, I would focus on a measurement of effectiveness of is this particular device or devices that we are considering effective in making a fusion happen and separate that from is the patient's outcome effective, and perhaps one or two years might be okay given the considerations that our statistical consultant just made.

DR. BOYAN: Thank you very much, Dr. Yaszemski.

I think we will go -- I have one question and I have a specific individual that I would like to address it to, since that individual raised the question. Let's see if I can find it. You will recognize you when I get there.

My question really has to do with the fact that a comment was made in the overall discussion that stents(?) and things that promote fusion between any two bones are merely aids and once the fusion has occurred, the surgeon can go in and remove those aids and the fusion would be stable over time.

But in most of the spine cages that we have had an opportunity to think about or look at in the literature or discuss, this is a technology in which the cage remains. It isn't removed and, therefore -- and I would like to preface this by sharing with you all, I am not an engineer. I am a biologist. So, I really want to get at the biology of what we might be looking at over time, where a device remains and, therefore, contributes to the overall characteristic of the fusion mass.

It isn't independent from the fusion mass. It is actually a part of the fusion mass and will remain in the human for the duration of their lives, hopefully, but certainly until a failure may occur at some future time. I enjoyed hearing myself quoted about the nothing much happens over the intervening four to five years, but it isn't really known about this technology whether that is the case.

So, I would like to raise the issue and have all of us consider the issue of the biology of fusion mass that in this technology, the fusion mass is not only the patient's own bone, but the patient's own bone together with whatever gets put into the site.

Is there anybody that would like to comment on that?

DR. HEIM: My name is Steve Heim. I am an orthopaedic spine surgeon from Chicago. There are several studies that have looked at that particular issue. In the instance of posterorlateral fusion, Dr. Craig out of Vermont did a very elegant study where he looked at the stiffness of fusion that was instrumented by pedicle screw devices versus non-instrumented fusions. The issue he was addressing is the risk of adjacent segment failure due to whether instrumented fusion was disproportionately stiff over a non-instrumented fusion.

What he found was there was only a 4 percent difference once the fusion mass had fully matured between the non-instrumented and the instrumented fusion. His ultimate conclusion why they were seeing adjacent segment failure was because of how the pedicle screw device was injuring the facet joints, rather than the fusion being disproportionately stiff.

In the instance of interbody cage devices, the cage itself provides initial immediate stability to the motion segment. These are interbody fixation devices versus interbody spacers, which themselves do not provide the motion segment stability, such as femoral ring. Once the bone heals, then we know biologically with spinal implants

that there is a fibro-reactive tissue that develops at the margin of the bone in the implant itself.

This occurs over the time during which the fusion is occurring. The bone is solidifying and basically the bone begins to take the overall stress, that being represented by the 4 percent difference seen by Dr. Craig. Essentially, the way I explain it to patients is when the bone is solid, the cage itself is similar to a toothpick frozen inside an ice cube. It is physically there but it is not functional.

DR. BOYAN: And I would argue with you from the biology, that that may be true short term, but long term with all of the experience with orthopaedic devices that has gone on over time, that that is not a benign, non-reactive interaction between the implant and the bone and that eventually, as has been pointed out by Dr. Wilkinson, there will be loosening because there will be resorption of bone adjacent to the implant that needs to be considered as we think about 12 months versus 24 months as this date of when we actually determine how successful a fusion is long term.

Thank you.

Dr. Hale.

DR. HALE: I guess my general comment would be

that I agree with what has been said about the device as being different and I think it is in the best interest of the FDA to have the flexibility to determine what the appropriate time periods are for evaluating those devices.

That said, as far as the issue of fusion is concerned, particularly in the area of the spine, from a more mechanical point of view, I am not sure that I am entirely comfortable with the idea of making a decision on these devices based on a one year time period. As was just pointed out, because of the nature of the biological interface, the load sharing that goes on throughout the initial time period, whether it is one year or two years changes as the bone fuses and potentially absorbs more of the load, carries more of the load than the cage does or whatever the device might be.

As loosening progresses, that may shift back in terms of the device carrying more of the load. So, I am not sure that that is a -- that you can clearly say once fusion occurs that the device is no longer part of that structure or no longer bearing load or that that is the way it is going to continue to be.

We have a very dynamic biological interface that potentially changes the role of that implant and potentially

changes or affects the patient's long term outcome.

From a purely mechanical point of view, I guess I also have some concern with using medical imaging modalities to establish the mechanical integrity of the repair tissue. Simply saying that there is trabecular bridging does not say that the mechanical properties of that structure are the same as solid bone or alternately because the device we are putting in there does not have the properties of the biological tissue it is replacing, whether that is the optimal structure to be creating.

My considerations, I guess, are that -- or comments are that one year may not be adequate to make that determination.

DR. BOYAN: Dr. Grobler.

DR. GROBLER: Thank you. Leon Grobler. I am an orthopaedic spine surgeon.

I think, listening to all the comments, we may need to go one step back and decide what is this implant's goal. The implant's goal in most cases or at least what we are discussing now is to obtain fusion. It is not a treatment for pain. And there are some elegant studies showing that from a teaching point of view you can have the totally correct implant, great biological fusion and the

patient still has pain and it is like doing a perfect landing at the wrong airport.

It just seems to me that we should go back and just realize we are trying to assess has biological fusion taken place. And I totally agree with a previous speaker, that it is not the metal that heals, it is the bone that heals. And I think -- and the follow-up to that is we need to have a method, and technology is improving, to be able to find out if fusion in actual fact has taken place.

If we do elegant CT scan sections at different planes, parasagittal and sagittal reconstructions, we can in most cases determine if fusion has taken place. I think that should be the decision point not just history. Quite frankly, I don't think that 24 months has been shown, as we discussed, as being a good scientific figure.

The other point I also think needs to be -- this is a general discussion and I think the FDA is the ideal forum to assist studies that go for 12 months and eventually go for 24 months and see, in fact, if, from a scientific point of view, there has been a change.

Then I also need to just add one other comment. I think you should go, as they say, one client at a time. You should go one implant at a time. You cannot compare apples

and pears. There is a great difference between fusion in the antibody level where there is distraction. There is load bearing compared with an inter-transverse fusion. We should take one implant at a time.

And I think if we don't do that, we are actually not taking technology by the throat and bringing it near to research.

That is basically what I -- the comments I have and I agree, this could only be a starting point. Some speakers mentioned that this was too short a time to make a decision. I am hoping we are not going to make a decision today. This is purely a forum for discussion.

Thank you.

DR. BOYAN: Thank you.

Dr. Kerrigan.

DR. KERRIGAN: I guess I have one sort of burning issue here from the standpoint of a non-surgeon. I am a physical medicine and rehabilitation specialist, who also specializes in biomechanics of movement, particularly walking. My concern is what the effect of theses fusions will have on other parts of the body and, most importantly, on the levels above and below where the fusion is taking place.

I mean, just common sense tells us that if you fuse one part of the spine, other motions have to occur to achieve any kind of movement that you do in life. So, it makes sense that these other motions are going to occur and if you are predisposed to degenerative disc disease at one disc, I would assume you would be predisposed at other levels as well. We don't know what the effect of a fusion -- and whether it is effective from a radiologic standpoint or not, I think, is sort of beyond that issue.

What is going to be the long term consequences, again, as everybody has pointed out, you know, these devices and the fusion are going to stay in for a long time and I am concerned about just the effect on these other levels' mobility as well, which I really don't think one year is going to be able to determine those.

DR. BOYAN: Dr. Cheng.

DR. CHENG: Dr. Boyan, I have two questions I would like to ask specific people and then I will make a comment.

DR. BOYAN: Okay.

DR. CHENG: The first question I would like to address to Dr. Kahanovitz. and that is you spoke on behalf -- is Dr. Kahanovitz still here?

DR. BOYAN: Spell the name because I don't think they recognize --

DR. CHENG: Dr. Neil Kahanovitz.

PARTICIPANT: He is not here.

DR. CHENG: Well, the question I would have asked and perhaps someone else knows the answer from the Spine Society, North American Spine Society, would be that does the society, which represents the spinal surgeons of North America, do they wish to deviate from their position as stated in their journal of a minimum 24 month follow-up requirement?

The second question I would like to address to Dr. Steve Heim and this was raised by Dr. Janosky already and that deals with we are basically assessing fusion of bone and we have no good means of doing that, although Dr. Heim has proposed the use of thin slice CT scanning to assess this problem.

I would like to know have you done any long term studies correlating the thin slice CT fusion results with the more accepted long term standard fusion, i.e., stability, although they may be more cruder measurements of stability and have you published them.

DR. HEIM: Again, my name is Steve Heim. To

address the second question regarding thin slice CT imaging as final fusion, in 1990 or 1991, I began the clinical trials with titanium alloy, pedicle screw devices. And since that time have used thin slice CT as an imaging and diagnostic criteria for success of fusion.

I have had in my own practice no instances of pseudoarthrosis not apparent on thin slice CT either with the interbody cage devices of several different manufacturers or with the titanium alloy posterolateral fusion instrumentation systems. Probably more to the point is as the initial IDEs for the antibody devices were being developed, they did rely primarily on plain radiographs, including flexion/extension. It was based on my past experience working with the posterolateral titanium implants that I began on my patients, also studying in 1994 my interbody cage patients with thin slice CT, sagittal and coronal reconstruction.

The images shown amongst the various investigators for that device were very impressive. The other investigators began utilizing the IDE, although it had not been originally written into the IDE protocol. To further objectify that, Dr. McAffee(?) did an animal study where he implanted interbody titanium cages, studied those animals

with a thin slice CT, though at that point it was only 1.5 millimeter, not 1 millimeter sections, sacrificed them and found the accuracy of the CT, as well as the specificity of the CT to document fusion to be remarkably high.

In my mind, I believe that was an 88 to 90 percent sensitivity and specificity correlation.

DR. CHENG: Steve, the question I have asked you, though, is do you have data which accurately can show that thin slice CT will predict fusion on a long term basis by most standard accepted criteria?

DR. HEIM: I always follow my patients for at least two years after any spinal instrumentation procedure, whether it is posterolateral or the interbody cage devices, whether it is an IDE or not. In any instance where I -- and I typically obtain my thin slice CTs at the three to six month time period.

In any instance where I have seen what I believe to be a solid fusion in that three to six month time period, I have at least the two year follow-up, I had no instance of a later diagnosed pseudoarthrosis. I have had instances where I have been concerned at the three to six month point, whether there had been fusion or not. I followed those patients and it appears that the majority of those did go on

to heal, though some did not, but the CT scan was very specific in raising my concerns to watch those patients more closely.

So, in that regard, I would say "yes."

DR. CHENG: So, you are basing this on your experience and not on peer reviewed published data. Am I correct in that?

DR. HEIM: That would be correct.

DR. CHENG: Okay.

My comment then on this subject would be that I think the evidence to date is overwhelmingly in support of the 24 month time period and I would base that on the following reasons. I believe that this is not necessarily a device specific criteria and that it is a disease process specific issue we are talking about and that is fusion of bone and that process is a biological process and cannot be shortened.

The complications which may develop as a result of a device, they may show up early or they may show up late. The failure of fusion, which may be one criteria -- or the failure of a device, which may be one criteria of fusion may not be evident until 24 months. I see no particular advantage to shortening the time period for follow-up from

24 to 12 months; that is, no particular medical advantage.

In addition, I think the history of orthopaedics is replete with examples where devices have been implanted with excellent early term results; however, in the long term have failed to hold up and complications have also developed at a later time point.

Finally, I think that although fusion is difficult to assess with the cage devices, which is what we are talking about today, the mere presence of some spicules(?) of trabecular bone bridging across two bony sites does not indicate evidence of a durable, long term, functional fusion and that over longer term, at 24 months, that may not hold up.

DR. BOYAN: Thank you, Dr. Cheng.

Coming around, I would like to hear from Dr. Silkaitis -- oh, did --

PARTICIPANT: [Comment off microphone.]

DR. BOYAN: No. We can't have any participation from the audience that is not specifically requested.

Dr. Cheng, do you feel like your question was answered sufficiently already?

DR. CHENG: No, it wasn't actually. Go ahead, if that is all right with Dr. Boyan.

DR. BOYAN: Yes.

DR. BLUMENTHAL: Scott Blumenthal.

To answer your question on the position statement read by Dr. Kahanovitz and as a card carrying member of the North American Spine Society, I do not believe that the society in general is willing to give up the standard criteria nor the Journal of Spine is willing to give up the standard criteria of two year follow-up for clinical studies. I don't know what meetings went on behind closed doors, but the general membership of NASS is not aware of that.

DR. CHENG: Thank you.

DR. BOYAN: Thank you.

Dr. Silkaitis.

DR. SILKAITIS: Yes. The industry recognizes that there are multiple options to evaluating a medical device and that the criteria that is used to evaluate that device, for example, what has been done in the past is basically uncontrolled studies. So, therefore, when looking at maybe more rigorous type of studies, where they are prospective and controlled, additional information can probably be gleaned for those types of studies.

So, I would like the Panel to consider those type

of studies, as well as the technologies that was addressed a little bit earlier and help clarifying for the industries what type of follow-up would be required.

DR. BOYAN: Thank you.

Finally, from our consumer rep, Dr. Holeman.

DR. HOLEMAN: I think I want to speak to that in reference to the outcome, not so much as to the device, and the effectiveness of the device in serving the purpose for which it was implanted, but to also provide enough time to assess to what extent has that device helped the patient with the activity of daily living, recognizing that when surgery is done, that different physicians considering different variables associated with that individual patient, that activity levels are increased at different times and to be able to understand to what extent a patient can function would require a longer period of time and could not adequately be assessed at 12 months.

DR. BOYAN: Thank you.

At this point, Dr. Witten, I was going to turn this to your questions. Are you -- would you like to make a comment?

DR. WITTEN: Well, actually if you are going to turn to the questions, you may cover it, but just following

along with what Dr. Holeman just said, there has been a lot of discussion about fusion and when you can detect it or determine it and not as much discussion about when pain and function, neurologic status can adequately be determined.

I don't know if there are any additional comments about that, but if there are, we would be interested to hear them.

DR. BOYAN: Yes, Dr. Wilkinson.

DR. WILKINSON: Well, as long as I have the finger pointed, in my statement earlier, I tried to emphasize that in my perspective simply achieving bony bridging or bony growth is not adequate, that the fusion needs to be strong enough to support the patient safely, comfortably and to permit functional improvement. I think the point raised by Dr. Kerrigan is also very important.

The impact of the device on the adjacent motion segments has to be considered part of the safety and effectiveness of any device. So, in my perspective, yes, we are trying to achieve bone fusion, but the reason the patient came to the doctor, the reason the doctor chose this operation was because the patient had symptoms.

If the device achieves something that is ineffective in relieving those symptoms, the device is

ineffective. The device can achieve a secondary goal without achieving its primary goal. That is ineffective. The patient didn't come to you and ask make my bones strong. The patient came and said make my back better. Make my life better.

So, to me, that is the goal of all medical input.

DR. BOYAN: In that context, I think then it is a good time to turn to the questions that have been put before the Panel. There are two primary questions that we need to discuss, but in the -- what you can see on our view screen is that we are to consider as we answer these questions, the safety endpoints, things that refer to the kinds of complications that might occur, how frequent they might occur, whether there needs to be surgical interventions and any unanticipated adverse responses and events.

We also need to consider effectiveness issues, whether there has been successful fusion, whether the patient is experiencing pain, whether they have returned to function, what their neurological status is. And I think we have heard Dr. Kerrigan bring in also the issue of what is the effect on the other parts of the skeleton, of the spine.

Then finally let's go to the first question, which is to address what issues that would -- if a 12 month review

was considered to be adequate, what issues would need to be assessed at 12 months in order to determine that a device is both safe and effective.

Just so no one thinks that they are going to have the same order as they had before, this time we will start with Dr. Kerrigan and come this way. So, the last speaker to the question will be Dr. Cheng.

DR. KERRIGAN: I think that is a problem for me to figure out the issues -- if we did say 12 months. The problem is that there is nothing in the issue that I brought up is -- I don't think that there is anything that you could evaluate. There is nothing that is going to -- you just wouldn't know at 12 months. That is why I don't think it is possible.

DR. BOYAN: Dr. Grobler.

DR. GROBLER: I think it is possible, as I discussed, that you can, with technology decide on fusion, but I think that is what the whole issue is. You are not dealing with a patient that came to you to ask about a fusion. You are trying to treat a patient's pain.

I was involved in a few recent studies where patients were asked by an independent medical person, not part of a surgical team, without even looking at the x-rays

and assessing the patient's outcome and I think that is what Dr. Holeman referred to. It is the outcome of the patient. What did the patient come to us for? I think that is important.

It is not just the fusion. That is why those four parts were put up there. It is the fusion, the pain, function and I don't think -- I think we have to acknowledge that we can approve the technology Dr. Heim talked about. I am sure all spine surgeons in this room are using that technology. It is not unique. It is available. We all use it.

I think it is the issue about the pain and function and I don't think we can easily assess that after 12 months. If we go one step further -- and that is going to become on the horizon in a couple of years -- we are going to start looking at disc implants. Now, I am totally sure that I won't be convinced that a disc implant is successful after 12 months, not knowing what happened to the adjoining levels.

Then one other is I very often see patients coming into my office --

[Fire alarm -- people exit room.]

I am actually recovering from bomb shock. I

thought I was saying something wrong when that bell went off. I will try to recap my thoughts.

I think looking at the four effectiveness issues that were put for us to decide is there anything that can be seen at 12 months to detect what is happening at 24 months, I think with technology and probably that is even improving more as time goes on -- fusion most probably is the one thing we can do, but I would be very hesitant to say that a patient's pain function, as well as neurological status can change within 12 months. Quite frankly, it may be to the advantage of industry that a patient is given time to improve their neurological function over time. That may have occurred at or shortly after surgery.

I also think that we should realize that fusion is not an issue of the endpoint. I am sure Dr. Dawson will not redo a fusion at one year if the patient is pain free. Therefore, we all patients for one or another reason do not have a total fusion or a solid fusion and still seem to be pain free.

So, I think, in summary, I would be very hesitant to go beyond 12 months for the other factors, although I think if our goal with the implant is purely for assessing fusion, I think technology is there and will be there sooner

than we think to be able to assess that.

One other aspect that was mentioned about the creation of load bearing of the bone graft, there are two ways how -- coming back to cages, although this is a general discussion -- coming back to cages, there are two ways that a cage can be stable. One is by interlocking bone. That doesn't imply that the bone has grown through from top to bottom and, secondly, by a through growth of the bone and in that case, the implant really does not take any weight. Therefore, the bone graft should be taking weight at that point.

So, I think that, to conclude, the fusion is something we can assess, but I have a problem with the standards we are using at this time for the other three factors to be adequately assessed.

Thanks.

DR. BOYAN: Thank you very much.

Dr. Hale.

DR. HALE: As may have been apparent from my previous comments, I am not a clinician. I am an engineer. So, I don't have a lot of clinical background to make assessments as far as pain and neurological status. But I do think it is reasonable from a mechanical point of view,

going back to the issue of changes in adjacent segments to think that because we are implanting an engineering type material to replace biological materials, that there may be subsequent changes not detectable at the 12 month time period in some of these clinical indices.

I think it is entirely reasonable to look at those at longer time periods and expect that to be a dynamic process that is going to change in response to mechanical changes in the composite structure of the implant and the biological tissues that are present.

DR. BOYAN: Thank you, Dr. Hale.

For my comment, I would like to point out to staff, FDA staff, that some of the issues about predictors of what might be occurring at 12 months, whether it is predictive for 12 months or 24 or 48 or 72 or whatever the number happens to be, could be gleaned from preclinical studies that are appropriately designed and that in looking at making a decision for an IDE, that maybe the design of the preclinical study could be better evaluated and that kind of data could be better assessed.

It could eliminate a lot of the questions that are coming up at this time in the discussion.

Dr. Yaszemski.

DR. YASZEMSKI: I think I am going to restate a comment similar to the one I made earlier and actually that has just been said again in large part by Dr. Grobler.

And that is that we do want to look at the patient's clinical outcome and solve the problem that they came to us for. I raise as an issue whether we can attach any causality between the type of implant that we choose for an anterior fusion and outcomes that are measured by the patient as their neurologic function, their pain score, et cetera.

I think I will reemphasize what Dr. Grobler said, that I feel the only thing we can objectively look at by device really is fusion. With respect to the safety issues, I feel that most of them can certainly be addressed by 12 months. With respect to the effectiveness issues, we will have to consider as an issue whether to link any causality between type of implant and patient outcomes, which is our goal as their physician and -- or rather, excuse me, between that implant and the fusion.

With respect to the biologic and mechanical nature of the implant, to a certain extent we have a little guidance in something that we do clinically because one of the things -- the techniques that we can use for an anterior

fusion is femoral allograft ring and the difference in biology between the incorporation of cancellous graft and cortical graft is such that cortical graft may not be incorporated for a long period of time to measure years and may from a mechanical perspective simulate an engineering material, if you will, that we put in, such as a gauge.

DR. BOYAN: Dr. Janosky.

DR. JANOSKY: If I look at the way that safety endpoints and effectiveness endpoints are defined, it seems to me that the definitions that are proposed on the slide are appropriate. I do, in listening to all the other Panel members discuss this issues, think that there needs to be a hierarchy presented, both in terms of effectiveness and safety; effectiveness looking at clinical outcomes as to what the patient is experiencing, what the patient is reporting and then also what is happening in terms of the spine and making a decision as to which one of the two needs to override is something that I think needs to be approached cautiously.

In terms of the issue about 12 months being predictive of some future point, whether that is 24 months or some other point, I feel that there is not enough information to make that decision.

DR. BOYAN: Dr. Laurencin.

DR. LAURENCIN: I think that I would have to reiterate what has been said and to a certain extent I think that there is really not enough information that we have to say that even at a 12 month time point will be sufficient under any set of criteria to be able to predict what the long term outcome could be.

If we look at what the possibilities in terms of 12 months, though, I think there may be two different situations in which a 12 month outcome could be accepted. One is if you have -- and actually, I think, Harry Skinner mentioned this in his memo -- if you have a study design that has such a high significance of P value in terms of a 12 month analysis, in terms of what the criterias are for success, that if you meet those very high criteria for success, then that may be -- it may be accepted.

The second is if you have a clinical situation -- I think this is what happens in the FDA right now -- if you have a clinical situation in which you are treating a disease that is perhaps tragic that has -- that there exists no adequate treatment right now, then, again, you may be able to make a case for making a shorter time point in terms of looking at what the clinical success is because you are

treating a disease that doesn't have an adequate treatment.

I am not sure that it really applies in terms of what we are talking about here.

DR. BOYAN: Dr. Wilkinson.

DR. WILKINSON: Yes, the question is what are the issues that would need to be considered and to my way of thinking any particular device would need to demonstrate that it cannot reasonably be expected to have any device-related delayed issues in safety or these four criteria of effectiveness endpoints. So, this has to be a device-specific determination. How can a device show that it is not going to change in the second year without doing a second year study?

Well, animal data might help. Equivalent data might help. Penicillin does cure pneumonia in less than two years or if the device was shown to have a near 100 percent success or failure. So, if at 12 months the total abject failure, then you don't need the second year and if it is as effective as penicillin was for pneumonia, you certainly don't need the second year.

So, I think it has to be a very device-specific decision.

DR. BOYAN: Dr. Holeman.

DR. HOLEMAN: I am going to go back to the assessment of pain and function. In listening to what has been said, it seemed that it is kind of difficult to get at the outcome of the patient as far as effectiveness. And I think there need to be built into the scientific design collaboration with other health care professionals, who may be better able to detect the extent to which a patient can function after activity has been resumed.

There are measures out there to get at the quality of life issue. There are self-care instruments. There are quality of life instruments that have been proven to be effective in assessing that aspect of outcome. And I think some of those instruments should be utilized in scientific research, such as this and also that collaboration with other health care professions who -- or with the patient, much more frequently and for a longer period of time than a physician who could adequately assess and provide objective data as well as subjective data.

DR. BOYAN: Dr. Silkaitis.

DR. SILKAITIS: Yes. With regard to the issues that would be needed to be considered, basically what the industry would like to see is very clear guidance, whether that be a guidance document, as to what the criteria for

evaluation. Obviously, it is a body of evidence that would support one time period of clinical follow-up versus another. If a company would choose to do a non-randomized trial, maybe that would require one length of follow-up, maybe because of the technology that is used to evaluate that particular device.

If a more sophisticated technology is used in a rigorous controlled clinical trial, well, maybe a shorter clinical time period should be considered. But those guidances should be helpful if they are clearly delineated in a guidance document to the industry. We do see articles in the literature that report on shorter periods of clinical follow-up and we as industry feel we shouldn't be held to, quote, unquote, a higher standard than what is typically done amongst clinical researchers in the area.

There are other tools that are available to the FDA to follow devices and their successes and failures and that is -- excuse me -- they are outcomes from a failure perspective and that is the MedWatch Program and companies as a requirement are to report those to FDA, any complaints and failures of the device that I reported to them.

So, basically, I think it is a body of evidence that should determine length of follow-up for a particular

device. That includes preclinical testing and mechanical testing.

DR. BOYAN: Thank you.

Dr. Cheng, you get the final word here.

DR. CHENG: I get the final word both times, I guess.

I want to reiterate that I am a strong advocate of maintaining the 24 month requirement and I say that for three reasons. Number one, given all of the evidence that was presented here this morning, most of it is opinions. I see no hard scientific evidence that 12 months is a good data point on which to base your outcome and assessment. Perhaps with fusion, as we talked about with some CT scanning, however, there is hard science showing that it is a good predictive factor as of 1997.

The second reason is that I see no advantage to approving devices earlier at 12 months. There is no medical reason. There is no medical advantage that I see to doing that. I can see there are other advantages but no medical advantage to doing that.

The third is that if there is some device, which a physician or industry should come up with, which is a miracle device, which does something truly extraordinary,

different from other devices that are currently on the market and that can be approved with a faster means of approval through an expedited review.

DR. BOYAN: Okay. Now, I think that we moved ourselves into question No. 2 by our answers, which the question No. 2 being what are the appropriate predictors of a long term successful intervention at 12 months and what would be needed -- are the tools we are using at 24 months appropriate for 12 months?

I think we covered the predictor question pretty adequately. Let's just call for anybody that wants to add anything to the predictor question.

[There was no response.]

Okay. The other side of question No. 2 is are there different criteria at 12 months versus 24 months that need to be considered? Does anybody want to address that?

[There was no response.]

So, I think that by your silence you are stating the criteria would not be different. Am I correct?

DR. WILKINSON: With the exception that I -- this is Dr. Wilkinson -- with the exception that I put forth. At 12 months, the manufacturer would have to show that there is no reasonable expectation of delayed changes in safety or

effectiveness.

DR. BOYAN: Okay. So --

DR. LAURENCIN: I just have one other point. One thing that sort of concerns me is that -- I wasn't here for last year's panel in 1996, but that in terms of the discussion here, two manufacturers came up and made statements regarding that meeting and one statement was -- sort of implied that different dates can be used and the other statement implied that there was a really -- we were being really -- that that panel felt that we were being held to a two year standard in terms -- they were being held to a two year standard in terms of the dates.

When I was listening to what was being said in terms of Ed Cheng's comments, they rang very true to me in terms of some of these criteria, that, in fact, the 24 month criteria really has good scientific evidence and there really is no evidence that a 12 month criteria, that there should be. There is an expedited review process.

Is that a consensus of the Panel in terms of what people think and, if so, maybe we should make some statement to that effect because what I wouldn't want to have happen is --

DR. BOYAN: Are you asking, for those of us that

were at that former Panel meeting, did we have a sense that we were being told that there was a certain number --

DR. LAURENCIN: Well, was there a sense of a consensus at that panel in terms of some of the -- in terms of what sort of criteria -- was that 12 month criteria a reasonable criteria, that it should be at least 24 months or what was the consensus at that Panel meeting?

DR. BOYAN: I don't think that was the subject of that Panel meeting. I don't think it was an issue at that Panel meeting. I think the discussion now is a new one, which is over the past year the question has come up to us from FDA to consider whether or not a shorter time frame than two years would be appropriate, which is a different question.

Is that -- Celia, am I covering it -- Dr. Witten?

DR. WITTEN: That is right and what we really were asking for was the scientific discussion, I think, that you have provided about what issues or what would need to be considered in that time point.

DR. BOYAN: Right. And I think we have answered that. Dr. Holeman, you were also present. Are your -- I think that is pretty much what happened.

Yes, Dr. Grobler.

DR. GROBLER: I just have one comment.

If we do go or decided to look at 12 months, I think there should be no shortcuts taken then. I think then it should be prospective studies. Every single patients should have gone through the whole process and I don't think there should be any issues then because then that is really the epitome of doing a scientific study, 12 months, taking everything and the deciding if that is, in fact, a case that could qualify for 12 months.

So, what I am saying is we should be very strict at 12 months if at any time that becomes a possible standard.

DR. BOYAN: Okay. Hearing no further discussion, then I would like to bring the general discussion to a close and just one last statement, I would like to ask FDA if we have addressed their questions adequately.

DR. WITTEN: Yes, thank you.

DR. BOYAN: Okay. Then we are on to the next part of our meeting. Only now there is no break between now and the next part because we have had that break. So, I have to go immediately to Plan B.

Okay. Now, for this next part of the meeting, I need to remind everybody that this is open to the public in

terms of hearing, but not in terms of participating. So, the public is, again, reminded that they are not invited to come forward and speak.

Agenda Item: Open Session -- Sofamor Danek Novus LC Interbody Fusion System

We are going to begin the discussion of the first PMA being presented before the Panel, Sofamor Danek's Novus LC Interbody Fusion System. We are going to start with the sponsor's presentation and I ask again that each speaker state his or her name and affiliation to the firm before beginning their presentation.

We scheduled one hour for sponsor presentation.

MR. LIPSCOMB: Members of the Orthopaedics and Rehabilitation Devices Advisory Panel, my name is Bailey Lipscomb and I am the group director of clinical and regulatory affairs at Sofamor Danek.

We are pleased to have the opportunity to present to you the preclinical and clinical data pertaining to the Novus LC Interbody Fusion System and to answer any questions which you may have.

The Novus LC Interbody Fusion Device is a threaded cylindrical titanium alloy cage, which is packed with autogenous bone and implanted between vertebral bodies in

the lumbar spine. Typically, two are implanted at a spinal level. The system also includes instruments which are necessary to implant the cages.

The Novus LC Interbody Fusion System is indicated for single level anterior use with autogenous bone graft in patients with degenerative disc disease in the lumbar spine. This is not the first time this panel has reviewed interbody fusion devices.

Last year, the data from two PMA applications of interbody fusion cages, the BAK and the RAY TFC device were presented to this panel and the Panel recommended to the FDA to approve both of them. Within several months, the FDA had approved both implants.

In comparison, the Novus LC cage is very similar to these two approved devices. In fact, the BAK device and the Novus LC device are covered by the same patent and we believe the same is true for the RAY TFC device.

In addition, all these devices are indicated for degenerative disc disease in the lumbar spine. Historically, the IDE for the Novus LC Interbody Fusion System received conditional approval in March 1995. At that time, the device was referred to as a threaded interbody fusion device or TIBFD. The IDE study featured a

prospective, randomized, controlled design.

Data from studies such as this have rarely been presented to this panel in support of a PMA application. Based on our review of previous Orthopaedic Panel transcripts, we believe this is the first time that a PMA applicant has relied on prospective randomized control IDE data from a clinical trial of a low band metallic orthopaedic implant device.

The first Novus LC study surgery was performed in April 1995. The PMA application for the Novus LC Interbody Fusion System was submitted to the FDA in March 1997 and accepted for filing by the Agency in May. In October of 1997, we submitted an updated report of the IDE results and responses to PMA deficiencies.

The preclinical and clinical data clearly demonstrate the safety and effectiveness of the Novus LC Interbody Fusion System. This supporting information will be discussed in the following presentations.

Mr. Larry Boyd of Sofamor Danek will first review the preclinical testing. Next, Dr. Scott Kitchel, an investigator in the IDE clinical study, will discuss the IDE clinical results. Dr. Jeffrey Toth(?) of the Medical College of Wisconsin will discuss the histological findings

of an explanted but fused Novus LC device.

Dr. Grant Somes of the University of Tennessee and a member of the Data Safety Monitoring Board for the study, will critique the statistical methods used in the analyses of the clinical data. Next, Dr. Allen Edmonson of the University of Tennessee will present the findings of the Data Safety Monitoring Board's review of the IDE results.

I will then conclude the formal presentations. In addition, other individuals are present today and are available to answer Panel questions. These individuals include three clinical investigators, one of whom being the inventor of the device, a surgeon who conducted the animal studies and the radiologist who read the x-rays and CT scans.

We also have a statistical consultant with us as well. At this time, I will turn the podium over to Mr. Larry Boyd of Sofamor Danek.

MR. BOYD: Good morning. My name is Larry Boyd. I am an employee of Sofamor Danek.

The Novus LC Interbody Fusion Device is a threaded, hollow, perforated, metallic closed cylinder. You can see the screw and NCAP in the figure this. This NCAP allows the surgeon to preload the device completely with

autogenous bone prior to implantation. Instrumentation allows the surgeon to position the devices accurately and allows for the ability to revise or remove the device simply by reengaging the threaded end.

The continuous thread form acts to resist expulsion of the device and facilitates fusion by aligning and stabilizing adjacent vertebrae. Holes in contact with the vertebrae further facilitate fusion by allowing for bony attachment and anchorage of the device. The Novus LC device is made entirely from titanium allow with no polyethylene components as found on previously cleared devices. Devices are available in sizes ranging from 12 millimeters to 20 millimeters in outer diameter and in lengths from 20 millimeters to 29 millimeters.

This slide shows a comparison between the Novus LC device porosity and the BAK device previously cleared by the FDA. All threaded devices, although somewhat different in structure have essentially the same amount of open total surface area for attachment to and integration with the vertebral bodies.

Next, I am going to discuss the preclinical testing. Instead of testing, the fixture shown was used to apply to axial loads. Load to yield for the device was

80,000 newtons. This is more than 18,000 pounds. As you can see in the figure, this represents a value far greater than that seen physiologically.

Maximum reported values are also shown for cancellous bone, cortical bone, the RAY threaded fusion cage, Brantigan cage and maximum reported value for the BAK device.

The more important data for a device such as the Novus LC is the fatigue or dynamic performance of the device. Syphlic(?) axial compressive loading was applied from 800 to 9680 newtons at 15 hertz over five million cycles. No observable microscopic damage or deformation was seen.

In order to better simulate the loading seen clinically, bending moments, compressive loads and a sheer load across the interface were also applied. You can see the fixture in the coordinate system in the figure here.

The load cycle life diagram shown here revealed a run-out value for the Novus LC device of approximately 1,500 newtons. This translates into a bending moment of approximately 135 newton meters. Adam states that this final motion segment can resist a bending moment of 33 newton meters before it sustains damage. Therefore, the

Novus LC provides a significant safety factor in bending loads also.

The stability of various constructs was tested, using a construct stiffness test in lumbar capsfines(?) using the fixture shown here. The stiffness of the construct in various loading methods was normalized versus the intact value. The stiffness is shown with the implanted value over the normal spine value.

The prior work by this group is shown in orange with the most recent data in yellow. As you can see, the Novus LC Interbody Fusion Device showed a significant increase in construct stiffness over intact and performed well versus both the bony pliff(?) and the SpineTech BAK device. Additional testing determined the installation torque and push-out load for the devices. Human spine segments between L3 and L5 were harvested and a BAK or Novus LC Interbody device was randomly inserted.

The maximum push-out load provides important information on the performance of the devices. The devices were installed with torque measured and a push-out force was applied. No catastrophic failure was noted and the push-out was via failure of the bony end plates with no implant damage seen.

This figure shows the push-out values for a smooth bone plug, a carbon fiber cage and the BAK Novus LC devices. Animal testing, using a sheep lumbar model was conducted. Animals were sacrificed at six months, followed by radiological analysis, non-destructive biomechanical testing and histological analysis.

Radiographs at 0, 2, 4 and 6 months allowed for assessment of changes in disc space height. As you can see, the surgeons were able to initially distract the disc space above the preoperative height for both the fusion case and the cancellous bone dowel, as shown by the increase in the line in the center of this figure.

The sham never exceeded its reoperative value. Over time all showed some subsidence; however, the fusion device appeared to level off at values still above preoperative levels, while the autogenous bone dowel and sham cases collapse well below preoperative levels.

Histological studies show bony incorporation and bridging bone around the titanium cage filled with autogenous bone. This view shows a transverse section of the device demonstrating secure attachment to the adjacent vertebral bodies. Both standing and microradiography support bony incorporation into the devices. Minimal areas

of fibrous tissue are seen.

Based on the previously reviewed preclinical testing, we have shown that the Novus LC device has adequate strength and performance characteristics to support its use in humans.

I will now turn the presentation over to Dr. Scott Kitchel, who will present the clinical data from the Novus LC prospective randomized clinical study.

DR. KITCHEL: Thank you, Larry. Good morning.

I am Scott Kitchel. I am an orthopaedic surgeon from Eugene, Oregon. The primary focus of my clinical practice is spine care and I am a clinical assistant professor of orthopaedic surgery at the Oregon Health Sciences University. I have no direct financial interest in either the product being reviewed today or the company which is its sponsor.

My presentation this morning is based upon my role as a clinical investigator in the IDE study of the Novus LC Interbody Fusion Device. I have also participated in two other clinical trials through the IDE system of interbody fusion devices very similar to this and have a large wealth of experience with interbody fusion surgery in the lumbar spine.

Before I move into the details of my presentation, I would like to briefly address the issue of 12 versus 24 month data, one of the topics under discussion today. I believe that there are a number of factors that make 12 month data sufficient to show the safety and effectiveness specifically of the Novus LC device with this study design.

First, our rigorous study design ensures in the most effective way possible that our results are valid and that all confounding variables have been controlled.

Second, as demonstrated in our study, spinal interbody fusion generally occurs between 6 and 12 months. The minority of patients who have not fused by the 12 month point are considered failures and are candidates for a revision surgery.

As we have heard this morning, surgeons in general will not wait an additional 12 months to assess whether or not fusion occurs in a clinical situation. Finally, we do have 24 month data, which as you will see, supports our argument that 12 month outcomes are predictive of 24 month outcomes.

There are three key issues that I would like to stress and ask you to keep in mind as you review the results of the study this morning. These key issues are the study

protocol, the success rate of the device in helping patients achieve fusion and the overall safety and effectiveness of the device, as compared primarily with the control group, but also with the two previously approved interbody fusion devices.

My comparison with these previously cleared devices is based solely upon publicly available information.

First, looking at the study protocol, we conducted a multi-center, prospective, randomized, controlled clinical trial. As this panel knows, this study design is highly desirable in evaluating medical devices. Although it is more difficult to execute, the resulting clinical data have greater reliability and are, therefore, a better assessment of the device's safety and effectiveness.

Fusion, the second key issue, is the primary endpoint of the study. At 12 and 24 months following surgery, the fusion rates of the Novus LC patients were 95 percent or greater. These rates are statistically significantly higher than those of the control group and are comparable to results reported in publications for the two commercially available fusion devices, which have previously been approved by this panel.

The third key issue is safety and effectiveness.

Measurements of safety were based on the occurrence of adverse events. Measurements of effectiveness were based on fusion, pain improvement, disc height, health status and neurologic status.

The Novus LC results were comparable to or in some cases better than the controlled patients in all primary outcome measurements at both the 12 and 24 month data points.

I would now like to elaborate on the clinical study and its results. The goal of the study was to evaluate the safety and effectiveness of the Novus LC Interbody Fusion Device in the treatment of patients with symptomatic lumbar degenerative disc disease.

As I previously mentioned, we conducted a multi-center, prospective, randomized, controlled clinical study of the Novus LC device. Clinical data were collected from study patients implanted with Novus LC devices, which had been filled with iliac crest autogenous bone. These data were compared with clinical data collected from a concurrent control group of patients. The control patients were implanted with femoral ring allograft filled with autogenous bone taken from the iliac crest.

The statistical hypothesis for the primary

endpoints of the study were based on the Novus LC results being equivalent to or better than the control results. Patients admitted to the study had single level, symptomatic, degenerative disc disease in the lumbar spine, as noted by intractable back and/or leg pain, supported by radiographic findings.

In addition, these patients exhibited spinal instability as documented on function extension radiographs. There were a number of additional inclusion criteria, such as age range 18 to 65 years, single level, symptomatic involvement between L2 and S1, no greater than Grade 1 spondylolisthesis and at least six months of non-operative treatment prior to any surgical intervention.

Patients were excluded from the study if they had a previous anterior interbody fusion procedure at the involved spinal level, required drugs, which interfered with fusion, had metabolic bone disease, an infection or fever, had a known metal allergy, required bone growth stimulation, were tobacco users at the time of surgery, were pregnant or had any compromising psychosocial factors.

Al the study procedures were single level and performed with an open anterior approach. Patients involved in the study were evaluated preoperatively during surgery at

3, 6, 12 and 24 months following surgery. At the time of the closure for the most recent data analysis, a total of 207 patients had entered the study and had received surgical treatment.

These patients were contributed by 22 surgeons participating at 16 individual investigational sites. Of the 207 patients, 145 were Novus LC patients and 62 were control patients. It is apparent from the Novus LC cases and control cases that there is not a one-to-one ratio as prescribed in the original study protocol.

This was due largely to a protocol amendment, which was approved by the FDA to allow investigators to enroll only Novus LC patients without randomization. Sofamor Danek purposely delayed initiating this approved amendment for nearly six months in order to maximize the number of randomized control patients.

The purpose of the protocol amendment was to foster lagging patient enrollment, due in part to the availability of competitive cages in non-randomized studies and to the reluctance of the investigators and their patients to participate in a randomized treatment allocation study.

Approximately 68 percent of the patients enrolled

in the study were part of the randomization process. The primary clinical support of the PMA application of the Novus LC device is the 12 and 24 month follow-up data. It is important to note that all study data at 12 and 24 months following surgery are from patients who are part of the randomization process.

Therefore, the following data presentations will pertain only to the randomized patients, except for comments regarding safety, which will pertain to all study participants. A total of 97 randomized patients have been evaluated at 12 months following surgery and 38 randomized patients at 24 months.

Postoperative follow-up compliance has been excellent, exceeding 92 percent at all study time periods. Demographically, the Novus LC patients were similar to control patients. The mean age and weight of patients in the two treatment groups were virtually identical. Thirty-eight percent of the Novus LC patients were male, as compared to over 53 percent of the control patients.

These rates were not statistically significantly different. Over one-third of the patients in both treatment groups were workers compensation cases and approximately 10 percent of the patients in each group were involved in

litigation. Approximately one-third of the patients in both treatment groups later admitted to tobacco use.

The radiographic inclusion criteria results were virtually identical between the two treatment groups. Disc disruption was the most prevalent diagnosis and it was noted in over 80 percent of the patients in both treatment groups. Disc collapse was reported in 72 to 74 percent of the patients and disc herniations were noted in approximately one-third of the patients in both groups.

In terms of preoperative spinal instability, there were no statistically significant differences in the mean translation in angulation measurements between the two study groups. The rates of patients having at least one prior back surgery were similar.

None of these distributions were statistically different. Looking at surgery results, there was no statistically significant differences between the treatment groups for mean operative times or blood loss. A majority of the surgical procedures in both treatment groups involved the L5, S1 level and more retroperitoneal approaches.

The average length of hospital stay for the Novus LC Interbody and control patients were virtually identical. Data pooling analysis confirmed that the data could be

pooled across all investigational sites.

I will now review the clinical results of this study, first beginning with the safety information. The safety of treatments used in the IDE study was assessed as a function of the nature and frequency of adverse events. Overall, 60 of a 142, or 42.3 percent of the Novus LC patients experienced at least one adverse event. Comparatively, 27 of 62 or 43.5 percent of the control patients experienced one or more adverse events. The rates of patients who had adverse events were not statistically significant between the experimental and control groups.

A majority of these adverse events occurred infrequently. However, there were five types of adverse events, which occurred at rates of greater than 5 percent in either treatment group. This slide shows these events and their rates.

The Novus LC patients experienced statistically significantly lower rates of implant breakage, implant displacement and non-unions in the control patients. The control group had a lower incidence of vessel damage and bleeding than the Novus LC group but the rates were not statistically significantly different.

Neurologic complications occurred at similar

rates. The adverse event information has been presented without much discernment to the severity of event or to distinguish between adverse events and complications. Many of the events were not clinically significant and readily resolved on their own.

Additional surgical procedures were classified into one of four categories; removal, revision, supplemental fixation or reoperation. Thirteen Novus LC patients or 9.2 percent of the cases, had a total of 14 additional surgeries. Twelve control group patients experienced 15 additional surgeries for a rate of 19.4 percent. The rate of patients who had additional surgical procedures in the Novus LC group was lower than the control group and the difference approached statistical significance at a P value of 0.06.

The one device removal in the Novus LC patient provided valuable information. This patient had continuing postoperative dysesthesia, tingling and irritation in the left leg because the left cage was malpositioned at the time of surgery. At approximately 16 months following surgery, the device was removed without any problems using a trefine(?) that cores about the device.

The resulting cavity was filled with bone dowel.

Radiographs and CT scans indicated the device and spinal segment had fused. Following its removal, the device underwent histologic and metallurgic analysis. Histologic analysis showed bone growth surrounding the device and growing through the superior and inferior holes, which is clearly indicative of fusion.

Dr. Toth will present a more in-depth review of these analyses later in this program.

In summary, the Novus LC device was found to be at least as safe as the control treatment. The overall rates of adverse events were comparable between the two treatment groups. The control group also had a nearly statistically higher rate of second surgeries than the Novus LC group.

In terms of device effectiveness, as evident from the nature of the product, the primary purpose of the Novus LC device is to facilitate spinal fusion. Therefore, we consider fusion to be the most important of the primary endpoints. Fusion of the surgically treated vertebral bodies was determined using AP, lateral and flexion/ extension radiographs. The radiographs were interpreted by an independent, board certified radiologist.

Fusion status of all study patients was assessed at 6, 12 and 24 months following surgery. According to the

protocol, a determination of fusion is based on the following rather stringent criteria. Evidence of bridging trabecular bone, translation and angulation stability and radio lucent line criteria.

The fusion results, according to these criteria, are presented in this slide. At 6, 12 and 24 months following surgery, the fusion rates of patients treated with the Novus LC device met or exceeded 95 percent. These fusion rates take into account the x-ray findings, as well as patients having second surgeries for failed fusion. If these second surgeries were not included, the fusion at the 12 month period would be a hundred percent.

By comparison, the control group patients had much lower but progressively increasing fusion rates of 23, 40 and 50 percent at the three time intervals. The differences in the fusion rates between the Novus LC and control patients were highly statistically significant at all time periods.

The Novus LC IDE protocol prescribed that CT scans be taken preoperatively and at 12 and 24 months postoperatively. The CT scans were intended to provide supportive information to the findings from the AP and flexion/extension radiographs. At 12 months following

surgery, trabecular bone bridging across the entire span of the implant was detected in the CT scans of over 95 percent of the Novus LC cases. This compares to only 66 percent of femoral ring allograft cases.

At 24 months following surgery, spanning, bridging, trabecular bone was noted in a hundred percent of the Novus LC cases as compared to 80 percent of the control group. We believe the CT information is particularly relevant to the determination of fusion in study patients since it provides the best direct evidence of bridging trabecular bone within the cages.

Analyses were performed to correlate CT results by comparing the evidence of bridging bone as determined by CT scans to the evidence of bridging bone as determined by plain x-ray. Although evidence of bridging trabecular bone could not always be determined using the x-rays, the radiologist was able to make these assessments for 80 percent of the 12 month patients and 97 percent of the 24 month patients.

These assessments were typically characterized by the observance of increased bone density at the implant interface when compared to preoperative films, which is consistent with bridging trabecularly. The findings as

shown on this slide show strong agreement between the CT scans and the x-rays. The percent of patients with evidence of bridging trabecular bone as determined by the use of CT scans was remarkably similar to that determined by the use of x-rays.

This was true for both the Novus LC and control groups at 12 and 24 months. These results show that CT scans are useful tools for evaluating fusion in the interbody space. The use of CT scans in assessing viable bone in cages was corroborated with animal data from two species. These animals were implanted with cages similar to the Novus LC designs and CT scans were taken postoperatively.

Histologic analysis following sacrifice revealed the presence of viable bone, which was likewise evident with CT scans. Dr. Vin Sandhu from the Hospital for Special Surgery in New York is here today and is prepared to provide additional information regarding these animal studies and to answer any questions, which you may have.

Further data showing CT scans as a means of evaluating fusion occurred with the analysis of the previously mentioned explanted fuse Novus LC device. In this case, CT scans and x-rays were taken prior to

explantation. The CT interpretations were consistent with the x-rays and the histological analysis.

In summary then, patients treated with Novus LC implants had higher fusion rates based on radiographic analysis at 6, 12 and 24 months following surgery. The differences at these time periods were highly statistically significant. In addition, the CT information was found to be a valid means of assessing fusion and the results were supportive of plain x-ray findings.

To determine if the Novus LC device and femoral ring were effective in helping to restore and maintain vertebral body spacing, preoperative intervertebral disc height was compared to postoperative disc height. If the postoperative disc height was determined to be greater than or equal to the preoperative disc height, the result was considered a success.

Disc height was evaluated based on radiographic measurement using an AP view whenever possible. The disc height information as presented in this slide does show that the Novus LC device, as well as the femoral ring allograft restored and maintained disc height.

None of the comparisons between the treatment groups revealed a statistically significant difference.

Pain in the IDE study was assessed using the validated and commonly used Oswestry Low Back Disability Pain Questionnaire. This is a self-administered questionnaire, which was completed prior to surgery and at each postoperative visit by the patient independently. The Oswestry questionnaire has ten questions, which focus on pain intensity and disabling affects of pain on daily living, such as personal care and walking.

The total Oswestry score combines the responses to the ten questions and were reported as a percentage ranging from zero to a hundred, with a lower numeric score representing a better or less painful condition. The IDE study protocol indicated that Oswestry results would be evaluated on a categorical basis. The patients were categorized as having successful pain outcomes if the postoperative Oswestry score was lower than the preoperative score; therefore, indicating pain reduction.

At 12 months following surgery, the Oswestry success rate for the Novus LC patients was approximately 88 percent as compared to approximately 70 percent for the control group. This difference in rates was statistically significant. The 24 month Oswestry success rates for both treatment groups were similar and exceeded 90 percent.

During their review of the PMA application, FDA raised a question regarding the clinical relevance of one point difference in Oswestry score.

To respond to this question, we used an additional method to analyze this data. This method was based on successful pain outcome being defined as an improvement of at least five points between preoperative and postoperative scores. This criterion was based on data from an article by Beurskens in the Journal of Pain.

Beurskens reported that a four to six point difference in Oswestry scores would discriminate a clinically relevant difference in pain. For the Novus LC study device, a five point improvement in Oswestry scores represents an approximate 10 percent change from the preoperative condition.

This slide represents the results of these additional analyses. The success rates are naturally lowered by this method as compared to that originally described in the protocol since the criteria for a successful pain outcome is more stringent. However, the rates of successful pain outcomes were comparable for the two treatment groups at each of the time periods of evaluation.

The higher 12 month success rate for the Novus LC patients approached statistical significance. Regardless of the method used to calculate pain success, the Oswestry questionnaire was found to be a useful instrument in measuring pain. The Novus LC pain improvements were comparable to if not better than those of the control group. The IDE study also utilized the short form 36 survey to assess the health status of all study patients.

The SF-36 is a self-administered questionnaire, which was completed prior to surgery and at each postoperative visit. The SF-36 questionnaire contains 36 questions that are used to produce eight scales that measure different aspects of health status.

According to the IDE protocol, in order for there to be a successful outcome for the physical functioning, role physical and pain index scales, the postoperative score must be greater than the preoperative score; that is, improvement in the results.

For the remaining five scales, a successful outcome occurs when the postoperative score is greater than or equal to the preoperative score; that is, maintenance or improvement of results. According to the approved protocol, overall success for the SF-36 is based on at least five of

the eight subscale parameters being successes.

As you can tell in this slide, the distribution of successful SF-36 outcomes were very similar for both the Novus LC and control groups at each of the postoperative study periods. During their review of the PMA application, the FDA raised a question regarding SF-36 successes also. To respond to this question, we have used an additional method to summarize the SF-36 results. This method was created and validated by the developers of the SF-36 questionnaire.

It combines the eight scales into two primary measures, the physical component summary or PCS and the mental component summary or MCS. Based on the normative data presented in the developer's manual for these summary scales, the condition of people with back pain and sciatica in the United States population would be classified as better or worse one year later if the differences in the PCS score was 5.42 points and the MCS score was 6.33 points.

Utilizing these values and the premise that successful classifications for functional measurements are traditionally based on maintenance or improvement in scores, the PCS and MCS success rates were calculated and are presented on this slide.

The success rates were similar for both treatment groups and no statistically significant differences were noted in the comparisons. The PCS and MCS success rates tended to be higher than the overall SF-36 success rates calculated by the previously described Novus LC protocol methods. The neurologic status of patients participating in the study was assessed preoperatively and postoperatively at every follow-up visit.

The neurologic evaluation consisted of measurements of motor function, sensory function, reflexes and the degree of straight leg raise, which reproduced leg pain. An algorithm was developed to reproduce the detailed scoring for each parameter into an overall success/failure classification for that parameter. Success/failure is predicated on the neurological status of the patient at a given time postoperatively, as compared to their preoperative neurologic status.

Every patient in both the Novus LC and control groups were successes in terms of motor function at all postoperative time periods. In addition, the distribution of patients showing sensory, straight leg raise and reflex successes were very high and similar for the treatment groups at all postoperative observation periods. The

overall neurologic status success rates were comparable for the Novus LC and control group at all times following surgery. None of the distributions for any of the neurologic parameter successes for the two treatment groups were statistically significantly different at any postoperative time period.

According to the IDE protocol, in order to be considered an overall success, the patient had to have successful outcomes for all primary endpoints; fusion, disc height, pain, health status, neurologic status and no secondary surgery, which was classified as a failure, obviously, requiring that a patient successfully meet every one of these criteria in order to be considered a clinical success is extremely rigorous.

For an example, a patient could be fused, pain free, in good general health, have no neurologic deficit and still be considered a failure on the basis of this criteria. The overall success rates for study patients of 6, 12 and 24 months following surgery are presented in this slide. All rates are higher for the Novus LC patients and the rates at 6 and 12 months are statistically significantly better.

The apparent drop between 12 and 24 months in the Novus LC group is an artifact resulting from evaluating two

different patient populations. When the same population group was evaluated utilizing a GEE analysis, no difference in overall success rates was observed over time.

At 12 months following surgery, the Novus LC success rates for all five primary outcomes and for overall success are statistically equivalent to or better than the control group rates. The delta values for this are less than 5 percent. These calculated delta values for equivalence were found to be far less than those described in the approved IDE protocol as being considered minimally, clinically significant.

Therefore, the objectives of the study have been clearly and unequivocally met, using the 12 month follow-up data, which has been presented. Notwithstanding the data comparisons from the IDE study, which definitively show the Novus LC device to be as safe and effective as the control treatment, the Novus LC results also compare favorably to publicly available information regarding two very similar FDA approved devices, the BAK and RAY TFC cages.

Because of the close similarities in the device designs, the clinical outcomes between the Novus LC device and these two commercially available devices were expected to be similar. Indeed, the IDE data shows this to be the

case.

I would now like to take a moment and present a clinical case from my practice, which is demonstrative of the use of this device.

This is a case of a 39 year old woman, who had severe disc degeneration and accompanying back and leg pain at the L5, S1 level. She had a three year history of this pain, which had been unresponsive to multiple conservative treatment modalities before she underwent surgery. This is her preoperative radiograph, which, indeed, shows the collapse of the L5, S1 disc space, as well as retrolisthesis, which is consistent with the inclusion criteria for the IDE study.

This is an immediate postoperative radiograph, which shows nice creation of a disc space and resumption of disc height, as well as some improvement in the sagittal plane alignment and appropriate positioning of the device. This is a film now at 12 months, which again shows good recreation of the disc space, with no evidence of any lucency or migration of the device.

These are axial cuts of a CT scan that were done at 12 months, which, again, confirm appropriate positioning of the device with no loss of integrity of the spinal canal

and you are also able to see at this point the density of cancellous bone at 12 months within the cages, consistent with that density that you see in other spots of cancellous bone in the spine. That is present at all the cuts that were evaluated.

This is a sagittal reformatting of that same CT, which also shows cancellous bone density comparable to that of the vertebral bodies within the cage itself, as well as the fact that bone is now growing into the interstices of the threads at each level. This is a special subtraction technique of the CT scan at 12 months, which actually has removed the titanium. The titanium cage would be present here in each of these, but through the subtraction techniques, that has been taken out.

You can only see the halo effect of that, but I think the interesting fact is that you can see cancellous bone in continuity growing from the vertebral body above through the device into the vertebral body alone, which would certainly meet our criteria for fusion. This is now a two month -- I am sorry -- a two year follow-up, again, shown primarily to show that there has been no migration of the device, no loosening and no loss of correction of the disc height.

The two year CT shows even more clearly the bone within the device, consistent with cancellous bone; again, not significantly changed from the bone seen at the one year point. Again, a sagittal reconstruction of that same bone, cancellous bone, same density as the adjacent bone in the sacrum and the L5 vertebral body.

And, lastly, another reconstructed CT, this time without removing the titanium, without that technique, but interestingly showing complete incorporation of the devices with bridging bone growing laterally on each side of them, as well as clear indication of bone growing through the pores, cancellous bone growing through the pores of the device that you can see in a number of different locations.

In conclusion, we conducted a prospective, randomized, controlled study of the Novus LC Interbody Fusion System. The Novus LC device produced superior fusion results at 6, 12 and 24 months, as compared to the control treatment. In addition, the Novus LC device produced comparable or better clinical outcomes and comparable safety profiles, as compared to the control treatment at all times checked.

Therefore, I believe that the results of this study, of the Novus LC Interbody Fusion System show the

device to be safe and effective in the treatment of lumbar symptomatic degenerative disc disease.

This concludes my presentation. Thank you for your attention. I will now turn the podium over to Dr. Toth, who will discuss the results of the retrieved implant analysis.

DR. TOTH: Thank you, Dr. Kitchel.

Good afternoon. I am Jeffrey Toth from the Department of Orthopaedic Surgery at the Medical College of Wisconsin. I have no financial interest in the product or company being reviewed here today nor do I have a financial interest in any orthopaedic device or device manufacturer.

During my few minutes with you, I will discuss the results of a retrieval analysis we performed on a single cage from one of the Novus LC patients. Retrieval analysis can be used to evaluate the materials response to implantation. For metals, this is primarily done by metallurgical analysis. Retrieval analysis is also used to evaluate the host response, which is essentially the biocompatibility of the implant material.

The retrieved Novus LC threaded titanium interbody spinal fusion cage was initially implanted into a 50 year old, 209 pound male patient with degenerative disc disease

at L5, L6. The cage was also packed with autograft at the time of surgery. The device was retrieved 16 months postoperatively.

Some clinical information, the device was removed due to initial malposition and continuing postoperative dysesthesia, tingling and irritation in the left leg. The attending surgeon indicated that no migration was observed from the postoperative radiograph and that no frank pseudoarthrosis caudaquina(?) syndrome, dysfunction or root deficits were associated with this cage.

In this slide you can see the CT scan at the time of revision surgery. On the left, malpositioning of the left cage caused impingement of the nerve root. This slide shows the trefine placed over the implant driver, which was used to explant the device. The trefine allowed the cage to be overreamed to preserve the tissue interface. Removal of the device occurred after all but the last few threads were cored over and the cage still needed to be slide hammered to be removed.

Gross observations of the intact device, there was no evidence of fractures of the device and the cage was surrounded by bone and dense connective tissue. In this slide, you can see the gross appearance of the cage with

bone and dense connective tissue in the threads of the device. The device was bisected in the sagittal plane.

Gross observations of the bisected device, bone and connective tissues were found surrounding the threads of the device. A mosaic of bone and dense connective tissues was found inside the device. In this slide, you can see the gross appearance of the bisected device with bone and fibral cartilage present inside the device.

The specific aims in evaluating this cage were to evaluate the host response to the retrieved device, to assess tissue ingrowth and tissues at the implant interface and to evaluate the fate and presence of autograft packed into the implant. Metallurgical analysis was performed on one-half of the bisected device. The tissues removed from that half of the device were processed by decalcified analysis, which included routine H&E staining, as well as special staining.

The other half of the device with intact tissues was processed by undecalcified histology, which included imbedding in polymethyl methacurate(?), sectioning in the sagittal plane, multiple staining and microradiographs. In this stained, undecalcified section, bone is blue. Fibrocartilage is purple and fibrovascular tissue is pink.

In this slide, an extensive amount of de novo(?) bone formation has occurred in the cage, as well as in the pores of the device. One area of fibrocartilage is observed. The next slide shows the corresponding microradiograph for that section. Note the extensive bone formation.

Radiolucencies in the center of the device and around the front of the device correspond to fibrocartilage and fibrovascular tissue respectively. As we progress more laterally through the cage, an extensive amount of bone formation is observed, as well as bone in the pores of the device.

This corresponding microradiograph shows bone in the pores of the device. The radiolucencies correspond to fibrovascular tissue as observed on this stained section. Fibroblasts were observed surrounding the implant. No inflammatory response was observed.

In one of the most lateral sections, you can see a few autograft fragments surrounded by fibrovascular tissue. Viable bone is present in the threads of this device. With regard to histological results, continuous bone growth through the device from superior to inferior was observed. The threads of the device were primarily surrounded by

fibrous tissue and one area of fibrocartilage was observed.

In decalcified sections, Taludin(?) blue was used to differentiate dark blue autograft fragments from de novo bone. At low magnification, you can see the autograft fragments consolidated into the fusion mass. At higher magnification, you can see the differentiation of the autograft fragments on the basis of color, as well as different lamellar(?) orientation.

Note the empty osteocyte lacuni(?) of the autograft fragments. Autograft fragments were incorporated into de novo bone. No osteoclastic activity was observed. Intratrabecular spaces were filled with hematopoietic marrow, as seen in these stained undecalcified sections.

Normal red and white cell lines were observed. Note the prominent megacaryocyte in this section indicative of normal hematopoiesis.

In summary, no inflammatory response was observed. Intratrabecular spaces were filled with normal hematopoietic marrow indicative of well-vascularized viable organ bone. Lastly, no osteoclastic activity was observed.

In conclusion, continuous viable bone through growth from superior to inferior resulted in a fusion. The autograft fragments from the iliac crest were consolidated

into the fusion mass.

Thank you. I would like to now turn the podium over to Dr. Grant Somes, who will provide a statistical review of the Novus LC study.

DR. SOMES: Thank you, Dr. Toth.

Good afternoon. My name is Grant Somes. I am professor and director of the Division of Biostatistics and Epidemiology at the University of Tennessee College of Medicine. I am also a member of the Data Safety Monitoring Board of the Novus LC Interbody Fusion System IDE study.

I have no financial relationship with Sofamor Danek and I do not own any Sofamor Danek stock.

I have reviewed the results of the prospective, randomized, clinical study and believe the statistical analyses performed by the company have clearly demonstrated the safety and effectiveness of the product. The Novus LC IDE was a prospective, randomized, controlled clinical trial. The random assignment of patients in a clinical trial assures to the greatest extent possible that the study groups will be equivalent for all possible confounders that might bias the results when comparing outcomes.

In fact, the company has demonstrated that possible known confounders were balanced across the

investigational and control treatments. Note that while statistical analysis in non-randomized studies can allow for partial adjustment of known confounders, there is no way to adjust for the influence of unknown variables that are not measured and that may affect the outcome.

It is on this point that a prospective randomized controlled clinical trial, such as has been presented, has a unique advantage that cannot be equivocated. I have evaluated the statistical procedures used by the company, both those related to specific time points and those related to longitudinal analyses.

I believe these were the most appropriate statistical procedures for this type of study and that they were properly applied. The 12 month data showed the Novus LC group was statistically superior to the control group for fusion, the Oswestry pain score and for overall success and at least equivalent for the other parameters.

In order to increase my confidence that earlier data predictive of 24 month data, I used common statistical extrapolation techniques to project the overall success rates from 6 to 12 to 24 months for both treatment groups. These projected rates are based on 84 patients in the Novus LC group and 50 patients in the control group and are very

close to the actual 12 and 24 month rates.

Based on these projected rates, it would be virtually impossible that the Novus LC group would not be judged equivalent to or better than the control group.

In conclusion, the study design used by Sofamor Danek was the best design possible. The conduct of the study and statistical analyses were appropriate and the results of these analyses support the conclusion that the Novus LC device is at least as safe and effective as the control.

I will now turn the podium over to Dr. Edmonson.

DR. EDMONSON: Thank you, Grant.

I am Dr. Allen Edmonson and I am presenting the recommendations of the Data Safety Monitoring Board. I am an orthopaedic surgeon at the Camel Clinic in Memphis and professor and chairman of the Department of Orthopaedic Surgery at the University of Tennessee College of Medicine. I am a spine surgeon. I am a founder and past president of the Scoliosis Research Society.

I do own some Sofamor Danek stock, making up less than 2 percent of my assets. None of the work on this device was done in our department or by any of our faculty.

The other two members of the monitoring board are

Dr. R. A. Kelandrusio(?), an orthopaedic surgeon and former president of the Orthopaedic Research Society and past president of the American Academy of Orthopaedic Surgeons and Dr. Grant Somes, the statistician, whom you just heard.

The three of us agreed to participate on this board prior to the commencement of the study of this device. In the first part of this year, we reviewed the data on the Novus LC Interbody Fusion System with a database closure of February 1997.

The more recent data, which has just been presented, were compiled with a database closure of September 1997. At that time in February, we believed that the controls part of the study could be stopped. The data were so overwhelming in favor of the Novus LC fusion system, that last February we unanimously agreed that it was not necessary to continue to randomize patients into the femoral ring control group.

We have also reviewed the data from the September closure of the database and believe that the information is now even more compelling than it was last February and we believe even more strongly that this prospective, randomized, control study has statistically and unequivocally demonstrated that the Novus LC Interbody

Fusion Device, Interbody Fusion System, is safe and effective for anterior, lumbar interbody fusion procedures in patients with degenerative disc disease.

I will now turn back to Bailey Lipscomb.

MR. LIPSCOMB: As we have clearly demonstrated in these presentations and in the information submitted in the PMA application, valid, scientific evidence of the highest order, that is, from well-controlled investigations, has been used to support the PMA application. Based on this valid scientific evidence, we have shown that there is a reasonable assurance that the Novus LC Interbody Fusion System is safe and is effective for its intended used. That is the criterion for PMA approval.

This concludes Sofamor Danek's presentations. We are available to respond to any Panel questions. Thank you.

DR. BOYAN: We are going to save all the questions under the end of all the presentations and we will now move to recess until 1:30.

[Whereupon at 12:34 p.m., the meeting was recessed, to reconvene at 1:36 p.m., the same afternoon, Thursday, December 11, 1997.]

A F T E R N O O N S E S S I O N [1:36 p.m.]

MS. NASHMAN: We are about ready to get started for the afternoon. I will turn the Panel back over to Dr. Boyan.

Agenda Item: Sofamor Danek Novus LC Interbody Fusion System -- FDA Presentation

DR. BOYAN: We are now ready to begin with the FDA's presentation. The first speaker will be Paul Williams, who is the lead reviewer.

MR. WILLIAMS: Good afternoon, ladies and gentlemen of the Panel. My name is Paul Williams and I am the lead reviewer for the Novus LC Interbody Fusion System PMA sponsored by Sofamor Danek.

I would like to first apologize that we had some difficulty with some of our slides and, therefore, we had to make last minute overheads. So, during our presentations, we may be referring to slides on the right hand side.

I will be giving an introduction, which includes a device description, the indication being sought and the results from the preclinical testing.

Dr. Orlee Panitch will then present the results of the clinical study followed by a brief presentation on the statistics by our statistician, Mr. Gary Kamer.

At the end of Mr. Kamer's presentation, I will then ask the Panel for input on some major concerns regarding the sponsor's application.

As previously stated by the sponsor, the Novus LC Fusion Device is a hollow, threaded, cylindrical, titanium alloy cage, which is intended to be implanted between two adjacent vertebral bodies in the lumbar spine. The cage will be available in several lengths, ranging from 20 to 29 millimeters, for diameters of 12, 14, 16, 18 and 20.

The indication being sought is as follows: The Novus LC Interbody Fusion System is indicated for use with autogenous bone graft in patients with degenerative disc disease at one level from L2 to S1. These DDD patients may also have up to Grade 1 spondylolisthesis or retrolisthesis at the involved level.

Novus LC devices are to be implanted via an open anterior approach. Additionally, DDD is defined as discogenic back pain with degeneration of the disc confirmed by history and radiographic studies and these patients should be skeletally mature and have had six months of non-operative treatment.

Next, I would like to summarize the preclinical. The next three slides are summaries of the mechanical

testing results. I believe the sponsor has adequately presented the results and, therefore, I have no additional comments. So, I will just flash the slides.

As previously presented by the sponsor, animal testing was conducted on 20 sheep. Eight were implanted with the Novus LC, which was previously studied as the TIBFD with autograft, six with dowel graft and six with shams. All sheep were sacrificed at six months. Radiographs were taken at 0, 2, 4 and 6 months. Postoperative subsidence was looked at. All three treatments underwent some degree of subsidence with the TIBFD sites subsiding the least.

In summary, from the preclinical results provided by the sponsor, it can be concluded that the Novus LC cage should be able to withstand typical physiological loads that would be expected in the lumbar spine.

I would now like to turn the presentation over to Dr. Orlee Panitch, who will present the clinical data.

DR. PANITCH: Thank you very much.

I must apologize in advance. There will be some redundancy to this presentation.

As previously mentioned, the study design was a randomized, prospective, multicenter, concurrently controlled study with the Novus LC device stuffed with

autograft versus femoral allograft with autograft. The study design was changed mid way to a non-randomized treatment group only protocol due to difficulty in patient enrollment.

The original study was approved for 20 centers and 20 patients per center for a total of 400 patients. Sixty-eight percent of the 207 patients in the present database are presently in the randomized protocol and 100 percent of the data that will be presented is in the 12 and 24 month database is from the randomized protocol.

I would like to make it clear here and I will again at further points that all the material that will be presented in this presentation is from the randomized protocol.

The study was terminated early as per the Data Monitoring Safety Board, as stated by the sponsor due to a shown superiority in fusion. Two hundred and seven patients are from 16 sites with 22 investigators. Ninety-seven of the patients evaluated at the 12 month time point, 53 of them in the Novus group with 44 in the control and 38 were evaluated at 24 months; 23 in the Novus group and 15 in the control.

The goals of the study again were to demonstrate

the safety and effectiveness of the Novus LC device for anterior interbody fusion for the treatment of patients with degenerative disc disease as previously defined.

I am just going to flash the inclusion criteria. You have seen these by the sponsor. Please stop me if you have any questions.

Here we have the exclusion criteria.

The safety endpoints were, again, assessed by comparing the incidents of complications and subsequent surgical procedures associated with the devices to those associated with the control device.

Effectiveness endpoints were measured both clinically and radiographically. Radiographic endpoints were as determined by AP lateral and flexion/extension plain films with four criteria being used: evidence of trabecular bone, translational stability of less than 3 millimeters, angulation stability of less than 5 degrees and absence of radiolucency, more than 50 percent of the implant.

CT radiography was introduced by the sponsor as a supportive mechanism. A 0, 1, 2, 3, 4 system was used, which looked at the interspace and the degree of trabecular bone seen. The CT radiography was based on a seven monkey study that used a TIF device with bone morphogenic protein

without autograft or allograft and based also on a single human patient, who had the explanted device at 16 months, as previously mentioned.

The effectiveness endpoints from a clinical perspective were as shown here; maintenance of disc height, improvement in pain/disability status -- this was as measured by the Oswestry score -- maintenance or improvement in health status, as measured by the SF-36 -- both pain and health status will have some questions for the Panel after this presentation -- and maintenance or improvement in neurological status.

Looking directly at how success was measured, in the original protocol, the Oswestry score was deemed to be a success if the postoperative Oswestry score was improved over the Oswestry preoperative score by one point. I would like to point out here that an improvement in Oswestry score is reflected by a decrease in the Oswestry score and the Oswestry score is based out of a hundred points.

In response to FDA concerns whether a one point out of a hundred scale was adequately demonstrative of clinical significance, the sponsor reanalyzed the data using the Beurksen approach, as previously described, which states that a five point improvement is necessary to claim

superiority or improvement in success in pain.

Regarding neurological success, four criteria were used; sensory, motor, reflexes and straight leg raise. This was based on the Iowa protocol and as you can see here, there is a scoring system with 24 points assigned to both sensory and motor and four points to reflexes and straight leg raise.

It is important to see in the definition of success that a patient can clinically have a decrease in function and still be considered a success using these criteria.

Regarding the general health assessment, as per the SF-36, the original protocol defined success using the eight subscales stating that success would be equal to at least five and that would be any five of the eight subscales.

It was weighted so that role-physical, physical functioning and bodily pain all needed to show an improvement, while the other five would need to show a maintenance or improvement to be considered a success.

Again, two FDA concerns regarding the significance of using that protocol, a modified analysis was provided as discussed by the sponsor. This was an analysis that was

created by the creators of the SF-36. The data now is divided into a mental component summary or MCS versus a physical component summary, the PCS.

Success in PCS is stated -- the numbers that are here, the 5.42 and 6.33 represent a standard error of the means that was predetermined by the developers of this survey. It is important to note here that success in the PCS allows for a postoperative worsening of 5.42 points to be considered maintenance or improvement and success in the MCS component, as well, allows for a postoperative worsening of 6.33 points to be considered a maintenance or improvement.

Data collection time points were preoperatively, intraoperatively, prior to hospital discharge, 3, 6, 12 and 24 months.

Okay. Now, looking at the data, this is a summary of patient accountability. And, again, I would like to emphasize this includes exclusively the randomized patients. I will just leave that up for a second.

Regarding demographic information, there was no significant differences in either group, except for height, which a one inch difference was found to between the treatment groups and that was found to be statistically

significant. That was felt not to be of clinical significance.

We have an overhead. Here you will just see a distribution of the radiographic findings preoperatively and this is as per the sponsor, no significant differences between the treatment groups.

We will go back to the slides now.

Regarding surgical information, the mean blood loss in the investigational group was approximately a hundred cc's more than the control group and the average range or the range for blood loss in the investigational group was 50 to 2,650 cc's while the control group was 50 to 750 cc's. On all other aspects, there were no significant differences.

I would like to just present the surgical information as provided previously. And, again, we have an overhead, please.

We have here a summary of preoperative medications and medical conditions and there are no significant differences between the treatment groups.

I think we skipped a slide there. Okay. We are missing a slide and I will just relate to you what that slide was to say.

It is a brief description here -- actually, maybe it is an overhead. Yes, it is. I apologize.

Okay. Here we have a look at the safety information. This table encompasses, again, the randomized group and what I have summarized here is any complication that occurred in either group that had a population of more than four patients that had that complication.

I will discuss neurological vessel damage and non-union on subsequent slides.

Regarding ileus, this was exclusively a postoperatively complication and this resolved without any problems in either group.

Regarding implant breakage, this was a problem that was seen exclusively in the control group. This all occurred intraoperatively and no complications arose from that. Regarding bone graft displacement, this, again, also occurred exclusively in the control group and five of these patients required revision several days postoperatively. One patient was treated conservatively.

Regarding the neurological and disc complications, which I have grouped together, in the investigational group, there were three disc herniations that required reoperation. These occurred at an adjacent level to the device

implantation, four nerve root injuries that required reoperation, three persistent nerve root complications that are being treated conservatively and two sensory events that have since resolved.

In the control group there was one disc herniation that has not required reoperation, two persistent nerve injuries, reoperation is scheduled on one and two sensory events that have since resolved.

Looking at vessel damage, in the investigational group, there were eight intraoperative events, including a vena cable tear. Estimated blood loss was presented on seven patients and of those, there was a mean of 1,350 cc's with an operative time reaching up to 9 1/2 hours.

In the control group there were two events. One of those occurred during a revision surgery. There was blood loss reported only on one of the cases with a mean of 250 cc's. This patient also required blood transfusion and cell saver intraoperatively.

Operative time was up to 4.75 hours. Looking at non-union, two events occurred in the investigational group, both of which went on to posterior fusion. In the control group, six non-unions required revision. One non-union is being treated conservatively with bone stimulation and two

are being conservatively managed in other methods.

What I have here is safety information that encompasses complications that affected less than three patients in either group. I would like to apologize. "Ileus" does not belong on this list. I will leave this up for a minute. And more safety information.

One patient in the Novus group did have a ureteral transection and required nephrectomy one week postoperative.

And we have an overhead.

Looking at radiographic results, I have two bottom lines here. One is an overall fusion success as reported by the sponsor. And the line below that is an overall fusion success that is based upon all patients. There is missing data for quite a number of the patients, partially because of difficulty in reading bridging trabecular bone, primarily, as the sponsor says, as a result of the implant.

At 12 months, we see overall fusion success with the complete data reported as a hundred percent at both 12 and 24 months for the investigational group and at 42.9 percent for the control group at 12 months, going up to 60 percent at 24 months.

For overall fusion success, based upon all patients, the numbers drop to 74 percent at 12 months for

the investigational group and 87 percent at 24 months and overall fusion at 12 months for control being 27 percent, going up to 60 percent at 24 months.

Looking at disc height, this, again, is as reported by the sponsor. I would like to point out that in the anterior method as looked at by Holte, there is a decrease in disc height between 12 and 24 months in the Novus group from 92.3 percent to 63.6 percent. Very few patients are reported on in total due to apparent difficulty in reading radiographs due to the implant.

Looking at Oswestry results, this is a summary of the mean Oswestry results taken across the treatment groups and at 12 months in the investigational group we see a score of 33.1 and that remains the same at 24 hours, at 33.3. Again, I would like to point out that an Oswestry improvement is reflected by a decrease in score. In the control group we see at 12 months a score of 33.8 and an improvement to 20 at 24 months.

Looking at the Oswestry results on the initial method of analysis, we see that at 12 months, the investigational group has 53 patients reported and a percentage of 88.2 success. At 24 months, that success rating goes up to 91.3 percent. In the control group, the

results are 69.8 at 12 months and increasing to 92.9 percent at 24 months.

Using the alternate approach to looking at the Oswestry results with the five point improvement reflecting success, we see that at 12 months, our success rating is 80 percent in the investigational group and drops down to 74 percent at 24 months.

In the control group, we see that at 12 months we have a 67 percent success increasing to 79 percent at 24 months. If we were to look at the Oswestry results and comparing scores between 12 and 24 months, we do see that in the investigational group, 11 patients out of the group had a worsening of their results by some amount between 12 and 24 months and an improvement of 35 percent in that time point.

In the control group, we see that 50 percent had a worsening and 36 percent had an improvement. And now we have an overhead.

Looking at the SF-36 results for quality of life, we have -- this is the initial analysis based on success in five out of eight categories. There are two analyses here. One, again, is on uncompleted. Patients with completed and the second line is based upon all patients and at 12 months

we see that the Novus group has a success rate of 84.3 percent and at 24 months to 78.3 percent.

In the control group, we see with the same line, 74.4 percent success and increasing to 80 percent at 24 months and the numbers below are presented for the complete patient population.

If we look at the SF-36 results using the alternate approach with the PCS and MCS system, we see that at 12 months, the investigational group has a success in the PSC, the physical component, of 90 percent and at 24 months, 91 percent. The control group has a 91 percent success at 12 months and a hundred percent at 24 months. If you look at the MCS data at 12 months, we have success in the investigational group of 88 percent and 83 percent at 24 months and in the control group we see 79 percent at 12 months for MCS and 73 percent at 24 months.

If we combine the data together at 12 months for the investigational group, we have a 78 percent success rate and at 24 months, a 74 percent success rate. In the control group we see a 70 percent success rate at 12 months, increasing to 73 percent at 24 months.

If we look at the SF-36 results in similar time frames, we can look at it seeing how many patients did not

worsen over these time periods and in the investigational group in the physical components, 62 percent did not worsen, which is to say 38 percent did worsen in the 12 to 24 month time period, as opposed to 71 percent did not worsen in the control group and 29 percent did.

Looking at the MCS data from 12 to 24 months, 76 percent did not worsen in the investigational group and in the control group, 86 percent did not worsen. Looking at neurological data, we have, again, two lines, both complete data and based upon all patients. There was difficulty capturing straight leg raising tests in some of the patients and that accounts for some of the incomplete data.

A hundred percent success is reported across the board for both treatment groups between 12 and 24 months if complete data is used. If we look at the entire patient group, we see that at 12 months the investigational group has a success rate of 86.4 percent and 90 percent at 24 months.

The control group has a success of 88.7 and subsequently at 24 months, 93.3 percent.

Looking at overall success, this is, again, as reported by the sponsor. We have an overall success of 72.7 percent if we calculate in all the factors; 72.7 at 12 month

time point for the Novus group and a success of 27.3 in the control group. The criteria here that are used are fusion by all four criteria previously mentioned. The pain with the success being equal to one point improvement. The initial approach to analyzing the SF-36, which was the criteria that utilized five out of eight and neurological success as previously stated.

If you account for all patients together, the success rate for the investigational group drops to 38.2 percent and the control group drops to 13.7 percent. You will see that there are deltas and P values. That will be discussed by our statistician, Gary Kamer, who is now to present.

MR. KAMER: Good afternoon.

The study to evaluate the safety and effectiveness of the Novus LC Interbody Fusion System began as a randomized concurrent control study with target enrollments of 200 Novus LC patients and 200 control patients. However, due to slow patient accrual, non-randomized assignment was implemented and a total of only 207 patients have been entered into the study.

Ninety-seven of these patients have been evaluated at 12 months and 38 at 24 months.

I will comment on the early termination of the study by the Data Safety Monitoring Board and present brief analyses of primary results. My comments will not be exhaustive. There are many statistical issues raised by this PMA and many analyses provided that will focus on statistical methods, which may be useful in the evaluation of Novus LC.

As sponsor submissions have so appropriately indicated, an interim analysis resulting in a statistically significant P value should be used as a guideline for the consideration of curtailing enrollment. All available information should be evaluated by the Data Safety Monitoring Board in making a decision to stop a study before its ultimate completion.

Most often, statistical stopping rules are concerned with the primary clinical endpoints measuring effectiveness. The Novus LC study appears to have been curtailed as a result of the significant improvement in fusion, a radiographic endpoint. Also, the endpoint of pain as measured by the Oswestry Low Back Pain Disability Questionnaire was described as significantly different between the treatment groups at six months.

However, it was not indicated whether this

statistical significance was based on a traditional Type I error of 0.05 or the Type I error required for the interim analysis, 0.0006.

The data for the Oswestry submitted on 3-25-97 to FDA, indicated a six month P value of 0.061, tutale(?), and a 12 month P value of 0.097 tutale. Stopping rules were to have been applied to 12 month data. Where does this leave us? The current data must be appropriately evaluated to determine if the Novus LC is both safe and effective.

I will now discuss some methodologies, which may assist in the interpretation of this study. I will be discussing only the randomized data from the study.

First, all the 12 and 24 month follow-ups are on randomized patients. Secondly, randomization offers many advantages, which are well-known to those involved in design, implementation or evaluation of clinical studies.

Before enrollment began, the number of patients necessary to establish appropriate safety and effectiveness claims was determined. One of the most important and most elusive of the necessary parameters for this is the definition of a minimum clinically significant difference, delta or D. For this study, those D's varied from minus 5 percent -- I am going to use minus. In calculations, I

believe normally they are considered just 5 percent -- to 20 percent, depending on the 24 month study endpoint.

Statistical testing for equivalence was performed using William Blackwalter's(?) methodology. I would like to just read -- which is for inferential analysis, for any endpoint, if equivalence is established statistically with a pre-study minimum clinically significant difference of 20 percent or a D of minus 20 percent, then we can be 95 percent sure that the Novus LC success rate is no more than 20 percent lower -- that is 20 percent lower -- than the success rate for the control.

Therefore, the clinical appropriateness of the sponsor's pre-study deltas may be evaluated since large negative d's or large values require so few patients to establish equivalence. If a d is considered unacceptable for any reason, equivalence may easily be tested using a more appropriate value for d. We like to do this before the study, unfortunately. But we can analyze this. A smaller negative value, such as d minus 8 percent, is an example.

The value of d should be clinically determined and may vary across endpoint and follow-up interval.

This table shows the sponsor's pre-study minimum clinically significant difference as d's for the various

overall study success rates, rate endpoints, at 6, 12 and 24 month follow-ups. Also shown are the observed success rate differences for positive values indicate that Novus LC performed numerically better and the control and negative rates indicate the opposite, that the control performed better.

Success rates were defined by the protocol method as described in the submissions. This is the first submission or this is the original analysis we are using. The last column indicates whether a statistical test established equivalence. With the exception of the 24 month SF-36 and disc height maintenance endpoints, statistically significant differences were attained, thereby establishing treatment equivalence.

However, the large pre-study d's associated, as you notice there, associated with some of the findings of equivalence, may require clinical evaluation especially when used with earlier than 24 month follow-up data.

Let me give a couple examples. For pain at 24 months, equivalence has been established with an observed difference in success rates of 1.6 percent in favor of the control. However, the estimate variability, the variability of the estimate, the 1.6 percent, is sufficient to assure

that the success rate for the Novus LC is no more than 20 percent worse than the controls. That is what these results said.

However, as will be shown later, this finding of equivalence depends greatly on the definition of equivalence, which is really the pre-study d used in the comparison of the treatments. Now consider the 24 month SF-36 results. Again, the pre-study d is minus 20 percent and the control is 1.7 percent more successful than the Novus LC. However, this time the variability of the estimate is too large to result in a finding of equivalence.

In other words, the observed 1.7 percent difference could be greater than 20 percent. This endpoint requires no additional analysis at this time. Only additional data could possibly establish equivalence.

If one considers having the original d's for a more conservative look at the data, pain at 24 months and SF-36 and neurological status at 6 months, joined the two previous endpoints as failures for establishing equivalence. I am not suggesting halving the d's or any particular suggestion at all on this. I am only showing how a change in the definition of "equivalence" may result in changes in the outcome.

Now, for a simple descriptive analysis, this table shows the overall Oswestry success rate for each treatment, as well as the differences in rates across all follow-up intervals. This type of table may assist in the identification of any outcome trends over time. As you notice, you will see the differences are 12.1, 16.9. At 12 months, it is 18.4 and then at 24 months the control actually has a better success rate by 1.6 percent.

This table shows the same for the overall SF-36 and its components. Again, the Novus appears to do better in the earlier follow-ups than in the latter one at 24 and even 12. As seen in the previous tables, results may differ over follow-up interval. While the issue of the sufficiency of a 12 month outcome follow-up or its ability to reflect 24 month outcomes may be discussed on a clinical basis, an effort has been made in the submissions to justify the predictive abilities of 12 month data based on examining the data for trend using longitudinal analysis.

This analysis was based on such a small number of patients, 34, as to have little meaning statistically. Even if there were more patients, such an analysis would be applicable to that subset, not the entire sample of 207 patients. While clinical evaluation could support the

ability of 12 month data to reflect 24 month data, only a sufficient number of patients evaluated at appropriate follow-up times, 12 or 24 months, can provide the data required for statistical evaluation of safety and effectiveness for a particular device.

In conclusion, this panel must review current data and make a clinical evaluation based on the clinical and statistical strengths and weaknesses of the study. I hope that my comments and suggestions will be of some assistance.

Now, Paul Williams will present FDA's questions for the Panel.

Thank you.

MR. WILLIAMS: Before we go on to proposing the questions towards the Panel and getting Panel input, I would like to ask members of the Panel if they had any questions with regards to the FDA's presentation that myself, Dr. Panitch or Mr. Kamer could answer.

DR. BOYAN: Mr. Williams, I think what we will do, if it is all right with you all, is have the -- after the questions are presented -- actually, we might even wait for that to hear from our clinical, preclinical and statistical reviewers, then have you present the questions and open it up for discussion.

MR. WILLIAMS: Okay. That is fine. It was our intentions to go ahead after that, after the Panel members' presentation, to go ahead and run through the questions once and then resubmit them each individually for individual Panel discussion.

DR. BOYAN: Perfect.

MR. WILLIAMS: Thank you.

Agenda Item: Panel Discussion

DR. BOYAN: Dr. Yaszemski, would you present the Panel with your clinical review?

DR. YASZEMSKI: After listening to the presentations by the FDA -- my name is Michael Yaszemski -- after listening to the presentations by the FDA, I have three comments to make.

First, with respect to safety, it seemed that there was a difference between the control group and the Novus group with respect to intraoperative complications. The intraoperative complication rate in the Novus group was -- if I wrote it down correctly -- about 10 percent, one of which was a cava tear.

I would be interested in knowing from the FDA if we have a listing of what the other complications were in that control group.

Secondly, with respect to the clinical outcome scales, our statistical analyst demonstrated to us that they seemed to change for the worse in the treatment group between one year and two year and we have been discussing earlier whether one year results were predictive of two year follow-ups. It seems perhaps that with the small number of patients and the trend toward decrease in clinical scores that perhaps some concern exists there.

Finally, if we look toward simply fusion as the endpoint without considering the patient's clinical responses, it appeared that the one year fusion rate was as good as the two year fusion rate of the control group. Since I feel that once fused that is a good endpoint and that is unlikely to change, as has been mentioned earlier, considering fusion alone, there appears to be an adequate fusion rate of one year.

That is all the comments I have.

DR. BOYAN: Thank you.

Dr. Hale, could you give us an assessment of the preclinical component of the PMA?

DR. HALE: With regards to the mechanical testing that was performed, in the absence of any standards for functional evaluation, I didn't have any concerns or any

specific questions with the mechanical testing that was performed. The device is clearly from a mechanical point of view exhibit behavior sufficient to withstand the forces that they would be expected to experience in this application.

The one comment or concern that I guess I would like to bring out is that mechanical properties don't necessarily -- superior mechanical properties don't necessarily correspond to superior properties in the intended application. The data that was presented has shown that these devices increased the stiffness of the intact spine in an animal model. And, again, as we discussed earlier, concerns about changes that adjacent segments having an increased stiffness of this device, compared to the intact spine may not necessarily be a desirable thing.

There was also information presented that suggested that rigid immobilization increases the probability of fusion and there is literature available that suggests the opposite as well. So, increased stiffness may or may not be construed as being an advantage of this device.

I think that is the only comment I have at this point.

DR. BOYAN: Thank you.

And then our last reviewer, if we could have Dr. Janine Janosky review the statistical part of the PMA.

DR. JANOSKY: I have for the presentation by FDA and also from the sponsor today, I have a number of limitations and concerns that I have about the application. Some of these have been echoed from the staff at the FDA. But let me just run through the list and especially for those that I didn't hear today.

We see that the data were polled over investigators and over sites, but then I see no information to tell me whether those investigators and those sites were comparable. Not only are those two groups -- and given the fact that the randomized part of the phase of the study had to be stopped and a non-randomized continued, due to non-compliance from investigators and patients, I do believe that there most likely is some type of bias until I am shown otherwise among the investigators or among the sites.

So, those data remain a question to me. When we are talking about it being a randomized controlled study, we are actually talking about only approximately n equals 90, n equals 100 subjects, if I remember correctly, and I see that number of 200 being thrown around and that is actually those

two together. So, we need to step back and remember we are only talking about a hundred, 120 patients.

The follow-up rate is a very big concern. One is going into that 24 months, but then also what is the complete data up to that point? I don't see data for subjects presented all the way up to the 24 points.

The sponsor mentioned a longitudinal analyses with perhaps imputation. That was not clear and if that was done, as Mr. Kamer had suggested, that -- I would like to see more information because that is very questionable, given the number of subjects that we had. So, what was the imputation in that longitudinal analyses.

There was high neurological complication rates, if I remember correctly. We are talking about 9 out of about 80, which seems to be about 10 percent and if we expand that to thousands of patients being treated with the cages, that number is of concern.

Then two other issues: One is the stopping rule. I see no evidence to show that the P value was adjusted, which is a very big consideration and the final idea is looking at the windows on the SF-36 and on the pain scale, you definitely need a window on there because of reliability. So, the sponsor's information presented about

improvement or maintenance was only looking at a one point difference, again, are very questionable.

In summary, these are a number of considerations. Perhaps some could be addressed, if that is the approach you would like to take or I can, you know, go through what implications they might have for the conclusions.

DR. BOYAN: Thank you.

Mr. Williams, would you like to present the questions now?

MR. WILLIAMS: I would like to add that in hearing the comments from the Panel reviewers that some of those concerns will be brought up within the questions and if those are not addressed, we can bring those up at the end and further address them.

First of all, what I would like to do is go through and read the Panel questions. The slides that will project the questions are just the questions and of themselves, there is some additional information that the Panel has and should reference. Also, the sponsor has a copy.

With regards to all Panel questions, there is a slight intro that I would like to read. According to the protocol, data on the following effectiveness endpoints were

collected: pain, as measured by the Oswestry Low Back Pain Disability Questionnaire; quality of life, as measured by the SF-36 Health Assessment Questionnaire; fusion, as measured radiographically; disc height maintenance, as measured radiographically and also, neurological status.

Data on the safety endpoints, which included nature and frequency of complications and surgical interventions, were also collected.

The first question that we would like to propose to the Panel is: In response to FDA concerns regarding the clinical relevance of a difference of one point as measured by the Oswestry Disability Questionnaire, the sponsor submitted an alternative approach to analyzing the results of the Oswestry. This alternative approach as presented by Beurskens in the Journal of Pain defined a clinical significant improvement as a decrease in Oswestry score of 4 to 6 points out of a possible 100.

Again, I would like to reiterate that decrease in Oswestry score is reflective of an improvement. The question proposed to the Panel is: Do the differences of 4 to 6 points out of a possible 100 represent a clinical significant improvement within patient in the Novus LC Interbody Fusion Device?

I will go on to question 2, which deals with the SF-36 assessment. In relation to the SF-36 assessment of quality of life, the developers of the tool grouped the data into two summary scales and within your Panel packets, we provided the SF-36 Physical and Mental Health Summary Scales: A User's Manual, which described how they derived the PCS and MCS scales.

Based upon multiple evaluations of these scales, they derived a numerical measure of data dispersion. They found that scores of 5.42 and 6.33 in the PCS and MCS summary scales, respectively, equal two standard errors of measurements.

With regards to this question -- well, let me add a little bit more with regards to the SF. Assuming that this equals two standard errors of mean, a difference between two groups greater than these scores may identify a statistically significant difference between groups. With regards to this topic, we have two questions.

The first one, particularly for the statistician, we would like to hear her comments with regards to is this a valid interpretation of the use of these scales as per the explanation given in the user's manual?

And then a general question for all Panel members,

which states: Sofamor Danek used this to demonstrate changes from pre to postoperative time points within patient assessment. Is this clinically meaningful?

The third question, the IDE protocol proposed that success of the study, when comparing the investigational and control devices, be based upon attaining equivalence. In the IDE, the equivalence was defined as the investigational device performing no more than 20 percent worse, which is equal to a delta of .2, than the control device at the 24 month time point.

Sofamor Danek is asking that we consider the same definition of equivalence and this is based upon the same measurement tools as per the protocol at 12 months rather than 24 months.

The first question to this is: Is this an appropriate definition of equivalence at 12 months based on the following concerns:

The concerns are: The sponsor assessed fusion radiographically, using the following four criteria: As stated before, evidence of bridging trabecular; stability, which included translation and angulation and radiolucency around the implant.

This was done using plain film techniques, AP,

lateral, flexion/extension views. We would like to add that it was difficult for them to determine the presence or absence of trabecular bone into and through the Novus LC spinal cage, using validated radiographic techniques currently available.

At earlier time points, stability as measured on flexion/extension x-rays may not be reflective of solid fusion, but rather may be due to instrumentation or hyperdistraction.

The second concern with that first question is previous spinal cage studies have assessed back pain, function and neurological status at 24 months. We have a concern whether assessments at 12 months, utilizing these tools, provide comparable reliability.

Part b to this concern of 12 versus 24 is also based on the two items with regards to Question a, the questions are: Are 12 month values as predictive as 24 month values for a given treatment? And are 12 month differences as predictive as 24 month differences between treatment groups? And that is: Can these measures at 12 months be considered as reliable surrogate endpoints for long term outcomes and be used on the basis of market entry?

With regards to what you have listed as Question 4

and Question 5, we would like to switch the order of those two questions. So, I am going to go ahead and go to Question 5 and then come back to Question 4.

Question 5 is regarding the CT scans, the reconstructed CTs introduced by the sponsors. Are the CT scans as taken, reconstructed and interpreted in this study an appropriate methodology of assessing fusion in the presence of morselized autograft?

And b: Are reconstructed CT scans at earlier time points predictive of long term fusion results?

I would like to go back to Question 4. question 4 is -- I would like you to pay particular attention and keep this in mind before the actual voting on the approvability of this device. You have a graph in front of you also that is included on the questions, but I will go ahead and read the narrative statements.

The sponsor stated that they found that the Novus LC device had an increased rate of fusion and an equivalence for safety and other effectiveness parameters in comparison to the control, the femoral allograft ring at the 12 month follow-up time point as per the success criteria defined in the original protocol.

Assuming -- this is where I would like you to pay

particular attention -- assuming that the results at the 12 month follow-up time point are sufficient to make an evaluation of the safety and effectiveness of this PMA -- so, that is the first concern -- and the results at 12 months of the entire cohort study of randomized patients are consistent with the results that are presented to date, would the results presented support a recommendation of approval for the Novus LC device?

There are a couple of things I would like to expand on with regards to this question and that is keeping in mind that we are stating that we are looking -- we want you to consider that at the 12 month time point when all patients have reached that 12 month time point and as a result we will have more patients at 24 month time point, that is the time -- that is the period of time that this question deals with.

I guess that is in conclusion to my presentation. We would like to, I guess, go back and take each question individually, have a discussion.

DR. BOYAN: Right. Thank you.

I think to begin with we will take a few minutes, maybe five to ten minutes to have -- if it is necessary, to have a general discussion of the general issues before we go

into specific questions. I would like to give everybody on the Panel an opportunity to make a general comment and then -- or comments that may not be directly focused on these particular set of questions.

To do that, we will begin at this time with Dr. Laurencin.

Oh, yes. The questions can be directed either to the manufacturer or to the FDA. Is that -- so, you should be prepared to answer at any time if you hear your name called or they can be comments made without directed questions.

Dr. Laurencin, would you like to start?

DR. LAURENCIN: I would like to get more information in terms of the FDA reviewer and also from Sofamor Danek about their concerns regarding the fact that the 24 month data shows trends toward worsening patient outcome over the 12 month data.

DR. BOYAN: Would anybody on the FDA like to respond to Dr. Laurencin's question?

Mr. Williams.

Could you repeat the question, please?

DR. LAURENCIN: My question is regarding the 12 month versus 24 month outcome data where there seems to be a

trend toward patients, when the patients enter the 24 month outcome point, that they have -- they have poorer -- they have a poorer outcome in terms of their Oswestry scores. Is this -- this was noted. What is the FDA's view of that in terms of that outcome?

DR. WITTEN: I think we are presenting an observation, but I think that question could be addressed to the sponsor to answer.

DR. BOYAN: That was Dr. Witten, who was speaking.

DR. SOMES: I am Grant Somes.

If you look at the cohort of people that were there for 6, 12 and 24, the rates stay about the same. Some of the people that have made it to the 24 months were early failures. A larger percent of early failures on the treated group have made it to 24 months than successes.

So, the cohort, if you have this cohort of people that were there for all three time points, their rates are not decreasing. Is that clear?

DR. BOYAN: Did you want to respond or --

DR. LAURENCIN: I am not sure if I am clear on that. You are saying that if you look at the patients that were at 6 -- that were measured at 6 --

DR. SOMES: 12 and 24, that were there for all

three time points. Then the rates would not be decreasing. Some of the early failures have already made it -- a larger percent of early failures have made it to 24 months and some of the successes. We think that this was related to a learning curve on that product.

DR. LAURENCIN: Okay. So, you are saying that successes look like they are -- and they are still in the earlier time points.

DR. SOMES: That is exactly right. If you look at a cohort, the cohort of people that have been in all three time points, that cohort.

DR. BOYAN: So, to paraphrase what you have just said is that given the group size that you have, as it has turned out, as luck would have it, learning curve or luck, the bad ones have finished first.

DR. SOMES: Some -- if you look at a cohort --

DR. BOYAN: And it is hard -- at this point it would be hard to say whether it was learning curve or whether it was just the way the world works.

DR. SOMES: Exactly. That is our supposition exactly.

DR. BOYAN: Right.

Okay. Any other comments, general questions?

Yes, Dr. Cheng.

DR. CHENG: If that is true, if the bad outcomers, the patients with the bad outcome are taken out of the group to be evaluated at 24 months, then your percentage of patients with successful --

DR. SOMES: No, they weren't taken out. No. They have made it to 24 months. They weren't taken out. They have been evaluated already at 24 months. They were not taken out.

DR. CHENG: So, what is the denominator then? The denominator has a greater or lesser proportion of the patients who had a bad outcome.

DR. SOMES: The denominator -- the denominator at 24 months, I am not exactly sure what it was, but it was smaller. I am saying a larger percent of failures, early failures have made it out to 24 months than early successes.

DR. CHENG: Okay.

DR. SOMES: So, if you look at the cohort that were there at all three time points, the rates are stable.

DR. LAURENCIN: This was done at 20 -- with 22 different investigators?

DR. SOMES: Yes.

DR. LAURENCIN: Can you tell me about the site

specific results -- you said that there was a bunch of bad outcomes. Where were the bad outcomes? Were they at one site? You attribute it to a learning curve. Was this a learning curve that everyone experienced? Would it be expected that an orthopaedic spine surgeon in the community when he starts practice would have this sort of learning curve, too, for this sort of outcome?

DR. SOMES: I am not positive it is a learning curve. That is something we -- I speculate on. I speculate on -- I did not do the analysis. I was on the Data Safety Monitoring Board.

DR. LAURENCIN: On the site specific, in terms of the bad outcomes in the group, where were they and was it the same -- was it one or two centers or was it all through or where?

DR. KITCHEL: I am Scott Kitchel. Just to remind you, I am an orthopaedic surgeon from Eugene, Oregon.

The bad results were not clustered at any single center. They were spread relatively uniformly throughout the centers. And we, indeed, believe or the most reasonable interpretation of that is that, indeed, there is a learning curve. You have to remember the study was started a couple of years ago when most surgeons did not have any experience

at all with anterior interbody fusion procedures.

So, there is, indeed, a learning curve in those first few cases.

DR. BOYAN: Let's go to Dr. Janosky.

DR. JANOSKY: I hear different methodologies placed for us and I guess for the most part this question is for the sponsor.

What exactly are you looking for? Are you looking for equivalence or are you looking for difference between the investigation and a control? Then as a follow-up question to that, was an interim analysis P value used?

MR. LIPSCOMB: I am Barry Lipscomb, Sofamor Danek.

The goals as we stated this morning in the description was an equivalent study design. I think that is covered in Dr. Kamer's speech as well.

Now, in terms of the interim analyses that were performed, we can have a further discussion on that by Dr. Somes.

DR. SOMES: The P values reported were just the actual P values.

DR. JANOSKY: So, they were not adjusted due to the --

DR. SOMES: Those were the actual P values that

turned out. So, I mean, if you wanted to adjust and you said, well, we need really a .01 because it is at a certain time -- now, your -- I guess, could I make a comment on your stopping rule comment?

The stopping rule comment was a probability -- a sequential probability ratio test. So, it was automatically adjusted for how many ever tests we want to do. I mean, that was a Wald test, Wald Sequential Probability Ratio Test, and that was in a letter to Bailey Lipscomb that he forwarded to the FDA and that was to stop recruiting controls.

DR. JANOSKY: I see presented are actual P values, though. Am I not correct? In all the presentation today made by the sponsor, those were actual P values.

DR. SOMES: All those presentations today, but that doesn't -- I don't know as any of those addressed the stopping of recruiting of controls. I don't think that was addressed today.

DR. JANOSKY: Right. See, the study was designed with a certain power, given a certain equivalence.

DR. SOMES: That is right.

DR. JANOSKY: But you didn't come close to that in terms of your recruitment.

DR. SOMES: On the sample size.

DR. JANOSKY: Right. That is correct. You did not. So, what were the adjusted power values?

DR. SOMES: The power, the deltas that you saw today, those were with 80 percent power. All those deltas you saw today were with 80 percent power.

DR. JANOSKY: But those weren't adjusted to that you stopped the trial early. Those were not.

DR. SOMES: Those were the -- for the actual data. Those were the 80 percent power for the actual data. So, those were at 80 percent power, all those deltas you saw in there today. We achieved this delta on the equivalence and all those were all with 80 percent power.

The differences were greater. The differences were greater than we had proposed from the beginning. Our 12 month differences were greater than we had proposed. So, we could easily have the same amount of power.

DR. JANOSKY: But you are not looking for differences. You stated you are looking for equivalence.

DR. SOMES: Right. We are going for equivalence.

DR. JANOSKY: So, we can't talk about differences.

DR. SOMES: We have a delta to calculate the sample size we need, that equivalence delta, that delta, but

not a difference, but that delta that is used to calculate the sample size. What would be a clinically significant difference. You have to have some delta to start with sample size.

DR. JANOSKY: Right. But when you are claiming equivalence and you are not interested in difference, you are just saying --

DR. SOMES: No. Equivalence within some delta range.

DR. JANOSKY: That is right. And I hear these two things being used interchangeably, which is not what is going on.

DR. SOMES: No. No, the delta that we are talking about is the delta to achieve equivalence. That is the delta we used to calculate the -- not we, that the company used to calculate the sample sizes.

DR. JANOSKY: Right, but when you are reporting results and you are either saying they are equivalent or they are not equivalent --

DR. SOMES: Right. That is exactly right.

DR. JANOSKY: That is not what I heard today. I heard superiority statements made.

DR. SOMES: There were some -- that is exactly

right. There were some statements made on superiority. You are right about that.

DR. JANOSKY: So, that should not be an issue on the table today.

DR. BOYAN: Before you speak, are you going to substantially change the discussion?

DR. KACHARINI: No, same discussion.

DR. BOYAN: But is it a significant new piece of information you are going to add?

MR. KACHARINI: Yes.

DR. BOYAN: Okay. Then identify yourself, who you are, all that stuff.

MR. KACHARINI: I am Richard Kacharini. I am the vice president for statistical services at C. L. McIntosh & Associates. We are a medical device consulting company in the Washington area.

My only interest in this product is I am a paid consultant to Sofamor Danek.

The questions which Dr. Janosky is addressing are very good and very important. However, the issue of whether or not equivalence occurs isn't resolved by having the mean be within a specific interval and, in fact, it is resolved by having a statistical test that you, in fact, are

statistically above the lower bound of that interval, the so-called two one-sided type of approach and that is an approach that the company has used to show that they are, in fact, statistically significantly above differences of 5 percent, not 20 percent.

I further challenge anyone on the Panel to tell me how when in numerous of these measures the product is superior to or lies above the mean of 0, how you can misconstrue that to be indication that this product is somehow worse than the control.

DR. JANOSKY: Can I just restate --

DR. BOYAN: Yes.

DR. JANOSKY: Can I just please restate the question to him also?

DR. BOYAN: Yes.

DR. JANOSKY: Before we get off this point.

DR. BOYAN: Right. I think that he has a struck a chord here and I think Dr. Wilkinson -- after you, Dr. Wilkinson and then I will --

DR. JANOSKY: Let me go back to my original question.

Dr. Somes, your trial was designed to determine equivalence or not find a difference as large as -- at least

as large as 20 percent.

DR. SOMES: That is correct.

DR. JANOSKY: But then you are making statements or statements were made today by the sponsor that things were superior.

DR. SOMES: The sponsor is not claiming superiority. If, in fact, a test demonstrates that it is statistically significantly greater than the control, it is, at least, equivalent. That is the construct in which that particular test procedure was done.

DR. JANOSKY: But that is not the way that you designed the trial. Is that not true?

DR. SOMES: I understand, but the company is not claiming superiority. And, in fact, if one goes on to do the tests and those tests were done and were provided to the Agency, the test of the difference between the control and the treated group against a lower bound of 5 percent, not 20 percent across the board, it is statistically equivalent to a difference of 5 percent, not the design basis of 20 percent.

DR. JANOSKY: So, where we heard statements today that the investigation was greater than the control, we really need to discount those statements if we go back to

the hypothesis of equivalence.

DR. SOMES: The sponsor is not claiming superiority.

DR. JANOSKY: Okay.

DR. SOMES: You can take those pieces of information for what they are, as evidence that it is not worse than the control.

DR. JANOSKY: Right, but we are not claiming better.

DR. SOMES: That is correct.

DR. JANOSKY: Okay.

MR. CHAMBERS: Gary Chambers, statistician with FDA.

The submission I read and reviewed indicated deltas going from .05 or 5 percent to 20 percent. I calculated many things before I decided to just go back and look at the deltas and do a simple test, the hypothesis test as proposed by Blackwalter. If I remember, I got two of those that were not significant that were indicated to be significant or, in other words, equivalent by the sponsor.

The deltas used to reach significance had to be the ones that were a priori stated and 5 percent was for one -- I think one of the endpoints. I believe that was

neurological. The rest went from 15 percent -- I think there was one 10 percent and the majority were 20 percent.

MR. LIPSCOMB: Bailey Lipscomb.

I think there are a couple of points that need to be made here. In our approved protocol that we submitted back in 1994, we submitted to the Agency certain delta values that we thought were justifiable and appropriate for determining statistical equivalence. That protocol was approved by the FDA at that point in time.

In our amended protocol, we adjusted some of those delta values and that protocol was likewise. That is a priori. Now, it seems like to me we have kind of an a posteriori(?) activity going on here but I want to get everybody back to square one.

Square one is if you look at the 12 month results, the 12 month results show equivalence for those endpoints and when you do kind of a backward -- moving in backwards into that Blackwalter calculation to say, well, how big a difference would this be -- you know, the threshold for significance, those are below 5 percent.

Forget what we said early on with 20 and 10 and 5. We have shown after the fact here in the study that we are below 5 percent. Dr. Kamer has done us a great service. I

would like to congratulate him because even if you look at the 24 month results, there is a whole lot of "yeses" on the slide that he put up for 24 months for overall success that were "yeses."

Even if you divide those deltas in half, like Dr. Kamer did, there is a whole lot of "yeses" on there. So, I think we have enough data at 12 and according to what Dr. Kamer has shown, I think we have got a pretty good amount of data at 24 to support the claims that we are making.

DR. BOYAN: Okay. I think that -- let's go to Dr. Wilkinson and then I think we have done this subject.

Dr. Wilkinson.

DR. WILKINSON: Well, I am obviously not a statistician. I sit in the room with patients who are hurting or who are feeling better. Everybody knows that statisticians are impeccably honest and you have to be careful of statistics and everybody knows that if you put garbage into a computer, you are not going to -- well, what you will get out.

What bothers me about this study is that the control group what was chosen here had terrible results. This was not the standard of interbody fusion capability. The other devices got far better results than the control

group. If you compare your product to something that has terrible results, it is going to look good.

DR. BOYAN: I think we need to keep going and let's hear from the Panel because this is not the time to defend the results. Now, don't do that to me because I am going to -- this is fair. You are going to not like me right at this exact minute, but you will be okay in a second.

Okay. We will go around and in response to that, though, I, too, must make a comment about the control group and I will now -- I feel, after your comment, to make a statement about the control group -- the way that the study was designed, not to be positive or negative about the choice of the current state of care, but the control group cannot ever, ever, ever look as good by your measures at 12 months because what you are using for one of your measures requires the control to resorb.

You are starting off with a bone ring that has to resorb. So, much of your data -- you have an unequal amount of autologous bone graft in the control and in the experimental and the bone graft, the autologous bone graft is your magic medicine.

So, unless you can explain to me -- I am not

complaining about it. I think you have very interesting data and you all have stood up there one after another and presented to show us that your data at 12 months is better than your -- in the experimental than in the control. But what I think we are being asked to discuss here right now is is that relative to what is going to happen at the 24 month data where the picture is not quite so clear.

Okay. You are feeling compelled to speak. You have -- I am going to give you one minute.

DR. KITCHEL: Perfect. Maybe 30 seconds.

DR. BOYAN: 30 seconds.

DR. KITCHEL: I would just like to remind you that the control group for this data was selected in extensive consultation with FDA and if you look at that control group, I would challenge you to try to find a more appropriate control group.

DR. BOYAN: I am not arguing about its appropriateness. I am arguing about the comparison, only the comparison at 12 months, not the appropriateness.

DR. KITCHEL: Specifically, there was to be no other instrumentation used and other than allograft bone, there is no other appropriate model to use for that control group.

MS. NASHMAN: If you could state your name for the record.

DR. KITCHEL: I am sorry. It is Scott Kitchel again.

DR. BOYAN: Okay. So, now that we have got all that out of our systems, let's go to Dr. Yaszemski.

DR. YASZEMSKI: Michael Yaszemski.

Perhaps, Scott, you are going to be the appropriate person to answer this or one of the other surgeons. It appeared to me -- I am going to come to the issue of the complications again. When I saw the slide with 9 or 10 intraoperative complications in the investigative group and very few in the control group, it seems to me that the operation up to the point of either putting the allograft ring in or putting the instrumentation in is the same operation. So, I suspect with respect to the approach, the surgeon getting there gets there the same way in both.

Does anyone from the company or the surgeons involved know what were the other intraoperative complications, other than the vena cava tear, and any kind of an explanation? Is it an instrumentation specific thing that caused so many more in the investigation group? I would like to hear about that.

DR. KITCHEL: These two specifically related to vascular --

MS. NASHMAN: Again, could you please state your name for the record?

DR. KITCHEL: I am sorry. Scott Kitchel again.

The issue specifically related to vascular injury and the difference between the two groups. I think the first thing that you have to look at is the level of adverse reaction that was picked up or was reported. Some of those were as small, perhaps, as nicks in the vessel made during retraction and, indeed, I don't think are either implant related or control group related.

The circumstance related to the vena cava tear was also not related to the implant at all, but was merely related to increased retraction and spreading of the vessel and the double barrel device or the implantation device was not in the wound at the time that occurred.

So, I think the answer to your question is, first of all, that if you look at the available literature about anterior approaches to the lumbar spine and mobilization of those vessels, the percentages of vascular injury range from down around 1 up to about 10. So, both groups actually fall within the accepted literature values and the differences

are explainable primarily on the basis of the severity of injury and the reportability of that injury.

DR. BOYAN: Dr. Hale.

DR. HALE: There is just one point that I would ask for some clarification, either from the sponsor or from the FDA reviewers on, and not in terms of the differences between the controls and the investigational device, but just in terms of the number or percentage of patients that were followed up at the different time points. The numbers that I had led me to believe that there was less than 80 percent follow-up definitely at the 24 month time period. I am not sure about the 12 month time period.

If somebody could make a comment about that and why those numbers are so low and if there is any intention to follow-up with the rest of those patients?

MR. LIPSCOMB: This is Bailey Lipscomb from Sofamor Danek.

In the table we provided -- and I don't know whether you actually -- we submitted it to the FDA. I assume you got it. We actually presented a table that showed the follow-up of an individual at different points in time -- not of an individual, but of the different groups, investigational versus control at different time periods.

As Dr. Kitchel said this morning, at each one of those time periods, the rates were greater than or equal to 92 percent. Now, those rates were calculated on knowing something about the patient at that given point in time. You may not have totally complete data, like in other words, you might have missed an x-ray. You may have missed something, but we know something about that person to where in some shape, form or fashion that patient has been evaluated.

That is what went into those rates. I notice one of the comments -- I can't remember who it was earlier made a comment about low follow-up rates and I am a little bit miffed at that one because I think we are confusing follow-up rates, which is more of a compliance issue, with the ability to do certain things, like we acknowledge in our PMA that it was difficult to measure disc height. The x-rays were taken. The radiologist had a hard time reading it. That showed up as a missing data point. And I don't want us to get zinged for not having that data point.

I mean, we don't have the data point, but it is the fact that it was an impossible measurement to make. So, I think our compliance is good. I think our compliance is supported. I think if you start going in and counting data

points in certain things like Oswestry, SF-36, those types of items that are straightforward, our compliance is very good. It is where you start getting into issues like straight leg raise and some of those that we fun into some difficulties.

DR. HALE: That is exactly the confusion that I was trying to bring up because some of the numbers, such as you presented this morning, show 92 percent; whereas, some of the other data show considerably less than that.

Does the FDA have a response to that?

DR. WITTEN: I think Mr. Lipscomb has already explained it. That is, they had -- the follow-up rate that they presented and not complete data on all data points each time the patient was seen.

MR. LIPSCOMB: We presented the data that way in the PMA. So, I mean, I think we covered that.

DR. BOYAN: Dr. Grobler.

DR. GROBLER: I just have a couple of points.

Firstly, it seems to me that going beyond the statistics, if we can get past the fact that at 12 months we have a safe and effective device, my question to the sponsors is what just in short would they say would make it compelling to say that 24 months would be the same? Because

to me that is the only issue at this time.

Twelve months, if we have statistically -- can show -- there was a long discussion on statistics. I presume that can be solved especially if we may move further on in getting more patients involved to the endpoint. But the issue is what happens at 24 months and what do you predict would be the factors that make 12 months -- 24 months as good as 12 months?

I am sure it was said by many, but I just need --

DR. KITCHEL: Scott Kitchel.

To me, the issue of 12 versus 24 months primarily revolves around the two major issues, safety and effectiveness. The safety issues, I think, we can clearly show in the study and if you look at the review of the adverse events, there are no new adverse events that occur between 12 and 24 months. There are people who have some adjacent level disease and some other problems that have shown up between 12 and 24 months, but they are clearly not device related and don't impact on the fusion.

The effectiveness issue, I think, primarily relates again to what we believe to be the primary endpoint, which is fusion. This device after all is to aid in effusion and the points that have been made this morning, I

think, and been reiterated are that the fusion rates we are very confident of based on the CT scanning, flexion/extension, radiographs as well, at 12 months and that once a patient is fused, we are not aware of any data in the literature to show that a patient who has achieved a solid arthrodesis, that will ever change with any amount of time, whether it be another 12 months or another five years.

DR. HALE: I guess what I -- I feel in some way the same because looking at other implants out there that had been approved, as well as conventional implants, how often do you really get safety issues after 12 months. The high risk is usually at least in the first 12 months.

Would you agree with that?

DR. KITCHEL: I would agree and I think, again, based upon the publicly available information, as well as my clinical experience with other devices of this same sort, I am not aware of any new safety issues that arise after the 12 month point.

DR. GROBLER: I have one other question, actually a statement in a way.

I think we should accept that placing two cylinders side to side does need more exposure than placing a femoral graft? Would that be a fair statement?

And I think your risk must be -- even if it is slightly, it most probably is higher, just the exposure. It is not precisely the same exposure.

DR. KITCHEL: Scott Kitchel again.

I would agree with you that depending upon the size selected and the size of the femoral ring allograft, but just as a general overall statement, the exposure enabled to put two side by side Novus LC devices requires a wider side to side exposure than the majority of femoral ring allografts do.

MR. WILLIAMS: How big were those allografts? The FDA data said they ranged from 6 to 60 millimeters in size. A 6 millimeter ring allograft isn't going to do a hell of a lot.

DR. KITCHEL: Scott Kitchel.

I agree with your statement. The data is as presented. They range from 6 to 60. I believe in the case where it was a 6 millimeter graft, it was -- that that was a dimension of height that that refers to, not a diameter dimension.

MR. WILLIAMS: Sixty wasn't the height.

DR. KITCHEL: No, 60 wasn't the height. Sixty was the diameter.

DR. GROBLER: I have just one more quick -- I agree with Dr. Wilkinson as well. I must admit the results on the control group is appalling. It is unbelievable. I mean, we have seen femoral allografts presented by Dr. Saylor(?), for example, a couple of years ago and I think that study was published in the 90 fusion rate and patient satisfaction, it just seems to me, it is not a question of the selection. It may be a question of the procedure, per se, and at the time when people were becoming slightly biased, getting such good results with the implant, could that be a factor, that -- I am just not sure how you got to those bad results.

DR. KITCHEL: Scott Kitchel again.

I think that the results you are talking about in the control group, first of all, you have to separate out the fusion results versus the clinical results. If you look at the clinical results, they really are quite consistent with what is in the literature or already published, available for both clinical outcome and pain relief. If you look at the fusion results, I agree they abysmal, but I think if you compare the criterion used to establish fusion in this study to anything else that has been previously published, you will find that these fusion criteria are

quite a bit more rigorous and more strenuous and that none of that published data has been subjected to this level of analysis. And that is what that difference is.

DR. BOYAN: Okay. Dr. Kerrigan.

DR. KERRIGAN: The way I got this from both the sponsor and the FDA presentations is that the overall fusion success was much better for the investigational device than it is for the control, something like 60 percent versus a hundred percent at 24 months. It didn't matter whether it was 12 months or 24 months. It was pretty clear the fusion success whether -- for the investigational device, what it was at 12 months stayed the same at 24 months. But, meanwhile, if you look at that and say, okay, well, the control, we weren't getting the fusion perhaps. Then why did the Oswestry, the pain scores and the clinical scores gradually improve from 12 to 24 months?

If we didn't have the fusion, then what is it that is making these patients better? I mean, either their ways to detect fusion aren't valid or we are missing something there or it is not the fusion that is making the patients have less pain, that there are other things and that patients are complex and it is not just the fact that all of their pain is being -- is a cause of what is going on at

that one disc, but of other things.

So, I just want to, one, make sure that I did get this correct as far as the sort of gradual improvement from 12 to 24 months in the control group, not the investigational group and maybe you have another explanation for it.

DR. KITCHEL: Scott Kitchel.

I believe the quandary that you are raising is one that has been addressed in a number of, again, published and available scientific articles regarding fusion surgery. There is no question that there is not a one-to-one correlation between solid arthrodesis and pain relief and that has been established with any number of different fusion models, whether it is the publicly available information from BAK or from RAY TFC regarding this model or whether it is the study done under the auspices of the combined group of pedicle screws or whether it is any of the older studies done with simply bone fusions themselves.

There is not a one-to-one correlation between fusion and pain relief. There are many papers reporting patients who have excellent pain relief that don't have a solid arthrodesis. Similarly, there are many patients who get a solid arthrodesis that don't report adequate pain

relief. And, clearly, what you are pointing to is an issue that the scientific community nor the clinical community knows the answer to. But you have spoken part of it and that is that there clearly is not a direct correlation nor it isn't a sine quo non, a fusion, that if you get one, you are going to get pain relief.

That is something we all explain to our patients before we perform any of these procedures. I think your second point about improvement in fusion rates in the control group goes back to something that Dr. Boyan said earlier about biology and that is there clearly is -- even though this was the best control group we could select in consort with FDA, there clearly is a difference between incorporation of allograft into a bone and the fusion that occurs in that way and the incorporation of cancellous bone and a titanium plug into cancellous bone and the fusion that occurs that way.

DR. KERRIGAN: I heard you say in your presentation that -- and just see if this is right -- that your most objective criteria for saying that -- for success, is that there is fusion or not. Is that correct?

That was the first data that you presented.

DR. KITCHEL: I am not sure that I used the word

"objective." And I think that has a certain value to it. I would say that what I consider to be the most important clinical endpoint of the data for this study is fusion, again, based upon what we have just discussed because there is not a clear one-to-one correlation between fusion and SF-36 or fusion and Oswestry pain scores.

DR. BOYAN: Okay. Dr. Cheng.

DR. CHENG: I would like to just revisit this outcome measurement of fusion again. You brought up a very good point in your last question about the lack of correlation and you tried to address that and no one really knows the answer, but you are using nonetheless a criteria of fusion as a basis for stating that your implant or device is superior than the control value in this experimental group.

One thing I would like to look at is the patients who underwent revision operation for additional stabilization. It looks like there are five listed in the data and six -- excuse me -- nine in the control group and there are also two revisions performed in the investigational group. What were the intraoperative findings in regards to whether or not those patients were fused as the ultimate, the gold standard for fusion is the

intraoperative examination for the surgeon to look at the bone to see if it is fused, was it fused.

DR. KITCHEL: I think we would like to look at Slide C8, perhaps. I believe your numbers are correct. We will see here in just a minute, but I know that in the experimental group or the device group, there were two operations for non-union or for failure of fusion. Those were both done from a posterior approach and with supplemental fixation and autogenous bone graft.

To answer your question, it is very difficult from a posterior operation to assess the results of an anterior operation. We could tell radiographically, I think, beyond any doubt on those patients, based on their CTs, that they did not have solid arthrodesis. But to go in posteriorally, the only real judgment available to you is to somehow try to toggle the vertebra and neither one of those cases were mine. So, I am not going to make up something that happened because I wasn't there.

DR. CHENG: So, my comment -- the basis for my comment is that although there are only two non-unions, according to the definition that you used, there were five reoperations for presumed pain, I assume, without other evidence.

DR. KITCHEL: I think this slide speaks to that. You can see what the indications were. One was a disc herniation at an adjacent level. I am not sure that is really fair to implicate the device necessarily in an adjacent level disc herniation. One of them was for some antecedent trauma that occurred again after the implantation when the patient was involved in, I believe it was a motor vehicle accident, sustained a fracture, wound up having a second operation.

I don't think it is fair either to implicate the device for that and one of them was for bony fracture. In the control group you can see there were five non-unions and three implant displacements and one indication of spondylolisthesis or actually one vertebra migrating over the other.

DR. CHENG: So, do you know from the data if there are any times where the radiographic criteria indicated a solid fusion, but the intraoperative examination did not correlate that? As that does happen in other types of spinal surgery.

DR. KITCHEL: Yes, it does. I don't know the answer to that. I can tell you that in the two control cases that I personally had that were non-unions, there was

evidence intraoperatively of the fact that there was not solid arthrodesis or you could wiggle the facet joints and get gross motion between those vertebra posteriorally, but I can't speak to the other, whatever that number is, minus 2, because they weren't my cases.

DR. CHENG: So, was the data that has been presented here by the sponsor, has that been looked at with a stricter definition of "fusion" as the criteria. I know you felt that there were stricter than other studies that have been performed. However, the criteria used do show that there is motion at the motion segment.

DR. KITCHEL: The criteria that we have used again that we consider to be the current gold standard if you will is the CT and all of these patients who are listed as non-unions and/or implant displacements on clinical evaluation at the time of surgery matched with patients who had CTs that showed that they were non-unions.

DR. BOYAN: Let's go now -- we have got still Dr. Silkaitis, Dr. Holeman, Dr. Wilkinson for general questions and then we are going to start through the specific questions from FDA. Okay. So, we will do that side of the room here.

Dr. Silkaitis.

DR. SILKAITIS: My comment is is that this was a controlled, randomized, clinical trial. When you take a look at the results if you compare 24 months to 24 months, the claim is equivalence and the numbers speak to themselves. Obviously, if you look at 12 months, you have got data that appear to show much higher values for the device, but, again, we are looking at equivalence. So, from that point of view and using a rigorous, a more rigorous clinical trial than what is typically published shows the equivalence of the device.

DR. BOYAN: Okay. Dr. Holeman.

DR. HOLEMAN: I have a question and I think that is directed to Scott. You made a statement relative to the fact that when you look at effectiveness of the device that you are looking at that based on (a) to fusion and you also indicated that it has been determined that there is no correlation between fusion rate and the relief of pain.

Perhaps it will take the information from an investigator to answer this, but what kind of information, what is told to the patient -- if you already know that there is no relationship, that the patient may have this without getting any significant relief from pain, what is told to the patient in order for the patient to make that

decision to have that surgery or not to have that surgery?

DR. KITCHEL: Scott Kitchel.

If I said there was no relation, I certainly misspoke. There is not a one-to-one correlation between achieving a solid fusion and pain relief. There is a high correlation between achieving a solid fusion and pain relief. In other words, what I am trying to say is, indeed, most patients who get a solid fusion get pain relief and most patients who have had pain relief wind up having solid fusion. But there are outliers to both of those statements. It is not a pure one-to-one correlation.

And the control data provides a very good example of that but there are numerous articles in the literature, which also provide that and it is a conundrum that is unanswered. It probably has to do with the biology and a number of other things.

Is that clear or what you like to me to expound on it further? The answer is we don't know why some people get a solid fusion and don't get pain relief and we don't know why some people who don't get a solid fusion do get pain relief. But I think your panel members, who do this sort of surgery will confirm for you that we all see those patients. And there is a big hug from all of us if you can figure that

out for us waiting somewhere.

The second part of your question, could you repeat that for me, the what do I tell my patients before I offer them this procedure? Is that essentially what you are asking?

DR. HOLEMAN: As what are just the benefits.

DR. KITCHEL: First of all, you have in your packets, I think, the informed consent for this procedure, which is very extensive. What I personally tell my patients is I try to have a very one-on-one discussion with them about, first of all, what I consider the indications, why I think they are a good candidate for this procedure, but then also what I consider to be a realistic view of the results and my realistic view of the results is based upon the information that you have heard today.

It is also based on the publicly available information for the BAK device and the RAY TFC device. So, for a one level anterior procedure, I tell my patients because I would always like to err on the side of being a little bit conservative, that they have about a 90 percent chance of being judged as having a good outcome and I define for them a good outcome as being a combination of achieving a solid fusion, getting good pain relief and be able to have

an equal or higher level of function after surgery than they did before.

I then continue to go through from there a discussion of the complications that you have seen most of presented today. There are some others that I also discuss with them. Now, that is specific to one level anterior surgery. I don't tell all fusion patients 90 percent because I try to rely on what is available in the publicly reviewed literature for those success rates.

DR. BOYAN: Dr. Wilkinson.

DR. WILKINSON: Well, I had a number of concerns and questions about this PMA, some of which have been answered. The chairperson has told me that she has heard my voice a number of times today and suggested that I share my written comments with the FDA and only hit the highlights, which I will do, so we can get out of here before midnight.

Let me go through of these, though. The draft device labeling and summary of safety and effectiveness described 15.1 percent peritoneal or vascular injury and 8.8 percent neurologic injury if you include retrograde ejaculation as a neurologic problem at 12 months.

An FDA review dated 9-17-97 upgrades those to 16.4 percent and 29.5 percent. An FDA review of an amendment

dated 10-24-97 lists 12 percent peritoneal and vascular complications and a frighteningly large 40.7 percent neurologic complication rate. That is page 8 of that review, the table.

Despite this, the draft patient information brochure describes a hundred percent neurologic success with the Novus LC and the slide we were shown today says 7.7 percent neurologic loss. After six months of histologic -- the histologic evidence in the sheep said that only 37 percent had histologic evidence of fusion and that is, to me, rather disturbing especially since the bending stiffness was greater with the autografts than with the Novus device.

This, as I mentioned earlier, could become more important with passing months since all three types of operated discs showed progressive loss of height and, thus, with time, stability becomes more dependent on bone fusion and less on mechanical stabilization. Unfortunately, no longer term animal data was presented.

I had some concerns about patient selection, which I will pass over for now. I was concerned about the radiographic criteria for fusion stability, how much bridging trabecular bone and where gave evidence of fusion. How could this be quantitated in those patients who did not

have CT scans.

Most of the stabilizing fusion growth presumably had to grow through the interstices of the metal device and not just within its lumen. So, how is new bone growth in those areas documented, especially recognizing the strong likelihood of x-ray and CT interference caused by the metal device?

From the data I had, only 27 patients out of the original 150 or so had CT scans at 12 months and only two at 24 months, though, those did seem to show bone fusion. From the data I had, only four patients had bending films done at 24 months, even though their earlier radiographs had shown progressive disc space settling.

In amendment No. 6, different numbers were presented for bending films at 12 months and 24 months, but there was still a larger percentage of those who showed loss of disc space height. I wondered what criteria were used to judge the diagnostic adequacy of bending radiographs. Was there a criterion that normal motion of adequate degree must be demonstrated at adjacent interspaces to confirm that the patient actually did bend?

Amendment No. 6, the FDA reviewed 10-24-97 reports a lack of overall fusion success using the four radiographic

criteria in 6 of 31 evaluable fusions at 12 months in the original submission, but in the amended data, there were no unsuccessful fusions, in 39 evaluable fusions. Was the original data wrong or was the amended data sanitized?

The Oswestry scale is a 0 to 5 point scale for ten items. The scale at 100 comes from giving a score of 2 for each of those points on the scale. Ten items 0 to 5, maximum 50, multiply by 2 to give a percentage score. So, each item is worth 2 points or 1 point of change in severity for a value of 4 to 6 points.

This means only one out of five improvement in only two out of the ten categories. Down on the farm we don't call that very good improvement. If you can show only one out of a scale of five improvement in only two out of ten categories. So, I don't know what the statisticians would tell you about this, but I can tell you the patient sitting in my office is going to be unhappy if that is as much improvement as they see.

I will quit.

DR. BOYAN: Response.

Dr. Witten, did you have a comment to make before the response?

DR. WITTEN: Just related to one of the

discrepancies in the neuro complication. One of the amendments referred to the overall complications that are considered in two ways, prevent and cumulative, with some of the patients counted several times, if not resolved. The numbers that were presented in the presentation were not cumulative. They were per patient events.

DR. WILKINSON: I am sorry. What does that mean in terms of reality?

DR. PANITCH: In the --

MS. NASHMAN: State your name, please.

DR. PANITCH: Orlee Panitch.

In the PMA amendments, the complication rates that were presented were both in a per event and per cumulative and if there was a complication that persisted over several time points, that was reported more than once. So, the complication rate, I think, that you are referring to, which is from the reviews, but not what was presented today, may be reflective of the cumulative complication rate where patients may be counted more than once. And that may explain why that figure is so high.

DR. WILKINSON: So, at a minimum, the complications that did occur were very lasting, very long lasting complications.

DR. PANITCH: Some may have been. That number does reflect that.

DR. BOYAN: Dr. Kitchel.

DR. KITCHEL: Perhaps we can kind of go back, Dr. Wilkinson, and address your questions one at a time because I can't remember them all. I know the first one was a histological question. So, I am going to defer to Dr. Sandhu and ask him to come and answer that because he was the one who participated in the animal studies.

DR. WILKINSON: I really did not intend these to be questions. They were more comments.

DR. SANDHU: My name is Harvender(?) Sandhu. I am a practicing orthopaedic spine surgeon at the Hospital for Special Surgery and assistant professor at Cornell University Medical College.

I performed the preclinical animal studies and I am going to address just a couple of points that were raised. That is Slide 34. That is the slide on the right.

This was a sheep intervertebral fusion model that has been previously discussed. The intent of this study was to specifically evaluate the functioning of a recombinant growth factor, recombinant human BMP2, in a spine interbody fusion model in combination with the intervertebral fusion

cage.

Our intent in this model was to purposely create a very challenging environment to be able to discern differences between these two conditions. So, we did a couple of things that we normally don't do in humans. We placed only a single cage in the sheep. We placed the cage in a lateral position than an AP, which is a much more mechanically challenging environment. The sheep, obviously, could not be immobilized and had no way of splinting their pain and they were moved immediately after surgery. Finally, our criteria for fusion were extremely strict.

Can I move to Slide 37.

This is an example and as the slide previous of what we called a failed fusion. There is bone graft and there is bone graft passing through the interstices, but because some of the interstices did not have bone in them, we defined this as a failed fusion.

In contrast, move to Slide 40. When we used the growth factor, we had complete fill of all the interstices and we declared this a fusion. So, in other words, we were using a very strict criteria to determine fusion. That is where we came up with the fusion rate of roughly 33 percent.

Thanks.

DR. BOYAN: Dr. Sandhu, show us a typical one that might be applicable to this study. Show us one where you put autograft into the cage.

DR. SANDHU: Could you show Slide 36. I unfortunately, picked out only slides I declared pseudoarthroses. And this is another one, as an example. I declared this as pseudoarthrosis, although you are seeing bone passing through the cage. You would have to take my word that there were several other conditions in which we declared fusions that had bone through and through the cage. But this is one that was I would say almost equivocal.

DR. WILKINSON: This is Dr. Wilkinson again.

If this study did not address the actual conditions to be encountered in the humans, which you just said, can you share with us the animal studies which did address the conditions that the human was going to encounter? Presumably, you did animal studies of what you are going to do to the big two-legged animals.

DR. SANDHU: Sure.

Again, whenever you are doing an animal study, you are attempting to model or simulate a human condition and it is often very difficult. Even with large animals, such as sheep, which are quadrupeds, you don't have the same

mechanical environment. So, in one sense we are looking at biologic processes and our definitions for fusions and success may be entirely different than what we see clinically because we are looking at biologic processes and we need to have very strict criteria to discern differences.

I will go back to Slide 27.

In taking into account those criticisms, we also did a study in a non-human primate so that we could simulate at least the biologic process a bit more. This is a rhesus monkey. The rhesus monkey was fused at the lowest mobile segment, which is L7, S1, which is the most mechanically demanded segment. Because of anatomical differences, however, and because we also wanted, again, to create a challenging environment, we only put a single cage in, this time in the anterior to posterior direction.

The study, too, was to evaluate biologic processes and the functioning of the recombinant growth factor. I think this is an important study because it may address some of the other issues that have been raised as well.

Please go to Slide 28.

In this study, in addition to evaluating successful fusion mechanically and histologically, we had the opportunity to perform reconstructed CT scans similar to

what we performed in humans. Here is an example of a reconstructed CT in the sagittal position. As you can see, this is very different than the scans you have seen previously in that there is no fill of bone in the center of the device.

So, this gave us the opportunity to see whether this could predict whether, in fact, a fusion was present histologically and mechanically. And, in fact, what you are seeing here confirms the fact that there is no fusion. We did not see bone passing through and through on the CT scan and histologically we, again, are not seeing bone passing through and through.

In contrast, in the condition in which we are seeing bone through and through on the CT scan, you can see the differences here. This simulates what we have seen previously this morning, both in the extracted cage, explanted cage in the human that had a CT scan prior to explantation. There was bone through and through the device. And, again, you are seeing a fairly impressive biologic fusion consolidation of bone from vertebra to vertebra.

So, in this sense we were able to use the -- we were able to determine whether the sagittal reconstructed CT

scans could predict what was happening in these devices histologically. As you can see, the last CT I showed was very similar to what we saw in the explanted cage, as well in some of the clinical cases that we have been seeing.

DR. WILKINSON: Are you saying that is the same cage?

DR. SANDHU: That is the same cage that had the CT scan.

DR. WILKINSON: That is the Novus cage. Where are all the little holes?

DR. SANDHU: This is a coronal section of the cage. This cage is, again, placed in the primate, which is a much smaller animal. So, the size of this cage is much smaller than the ones we use in humans. The vertebrae of the primate is only about, I believe, 15 kilos. It is a small animal.

DR. WILKINSON: Well, it depends on which primate.

DR. SANDHU: So, the relative proportion of the holes in the metal, it cannot be exactly the same. You can't reproduce that because you are using a much smaller cage.

DR. WILKINSON: So, I will ask my question again then. Other than these six small monkeys, did you have any

human data -- any animal data that used this cage in a way equivalent to the way it was going to be used in the humans? Baboons are pretty close to human weight. It is possible.

DR. SANDHU: In my experience, we haven't come up with a good --

DR. WILKINSON: So, have you or have you not done the study? That is all I am asking.

DR. SANDHU: I don't know yet of a model that we can exactly replicate --

DR. WILKINSON: I am not asking that. I am asking did you do a model?

DR. SANDHU: Because we don't have a model available that can exactly replicate that, no, we did not.

DR. WILKINSON: The answer is "no." No animal studies of an equivalent use of this device. Okay. Thank you.

DR. GROBLER: Could I make a brief comment, if you don't --

DR. BOYAN: Okay. A comment from Dr. Grobler.

DR. GROBLER: I guess my concern is precisely the same. I don't think the animal model supports this application very strongly because what was used is not equivalent to what is presented here as a clinical model.

And, secondly, I do think there is animal models out there that can be used. I don't think -- there is a primate model. There has been previous studies done in similar circumstances. So, I think you could do it if you think it is needed. I am not saying it is needed. All I am saying is this study I don't think has bearing on this application.

As I see, this is recombinant bone, right? There is no case that bone graft was used with the implant. Is that correct? That is all I am saying. I am not saying the study should be taken as a big factor in the decision-making, but I just think it really isn't a study that supports in any way this application.

MR. BOYD: Just a quick word --

DR. BOYAN: Well, just a second, in a minute, yes. I think we do need to clarify apples and oranges because we are talking a lot of apples and oranges in this discussion.

Autograft takes awhile to resorb. It is radio opaque on a CT scan. Some of what we are looking at is, in fact, autograft that has been incorporated in as the new bone is coming in.

I think in fairness to the audience they need to know that. In the control, you are using allograft, which

has no cells in it that are alive and it has to be completely resorbed and replaced in order to be successful, but the long term result at two years could be clinically equivalent in fusion and apparently in pain is, in fact, better by the data that has been presented to us.

The cage in the way that you have used it with the bone graft in it, the autograft, I think there is nobody on this panel that won't agree that you have presented reasonably convincing evidence that you had overall better fusion than you had with the control. But that would be expected because it is -- one is apples and the other one is oranges and I understand that that is the clinical trial that you were set with. I am not punishing you because you did the thing that you were asked to do.

But I think that the audience needs to understand what each of those different experimental parameters actually means to the outcome that you get.

Now, if you are going to add something substantially new to this discussion, then I would love for you to speak. State your name, who you are, what you do. If it is not going to be a whole lot new -- because we still have one whole more PMA to do and we want to give this one a fair chance and answer the questions for the FDA.

Is it going to be -- okay, you feel like it is. All right.

MR. BOYD: My name is Larry Boyd and I am an engineer with Sofamor Danek. I think what I wanted to do just briefly, and then I will sit down if you would like, was to just talk about the fact that it is the whole preclinical picture that really gives us a picture of the likely performance of the device in humans. So, it is the combination in my mind of animal work in models that while there may be inherent compromises just due to the availability of animal models, gives us some idea of the biologic ingrowth combined with in vitro testing in explanted spines, push-out testing in human spines and a variety of other ways that we can use to draw conclusions on the likely performance of the device versus comparable or historical devices and its potential performance in the human situation.

So, I guess what I wanted to say was the animal testing is just a part of the overall evaluation of the device and I wanted to make sure that we do consider that.

The other thing I wanted to make sure we remember is we do have explant histology on a single device albeit, but in humans at 16 months, which is very telling in terms

of the performance.

DR. WILKINSON: It tells about one patient.

MR. BOYD: I understand that.

DR. GROBLER: Could I -- I just need a quick comment. I think -- if you don't mind.

DR. BOYAN: I keep promising Dr. Laurencin that --

DR. GROBLER: If you don't mind, I just want to finish -- the animal model, I think, would have been perfect if instead of using the recombinant product, you use autologous bone. And my question is why was that not used?

Why was the animal model not done similar to what you are going to do to the human? Because that is the obvious question.

MR. BOYD: Larry Boyd again.

There is an easy answer to that and the fact is that we needed to show in the primate that there was a positive effect to the recombinant growth factor and that the carrier itself had no favorable effect. So, we, in fact, did a negative control to prove that had the carrier alone been used, that spine would not fuse. We were trying to prove a benefit of the growth factor.

DR. BOYAN: But that is a study that has very little to do with the current PMA. That was a gratuitous

study related to something else. I think that we need to put it where it belongs in how we are thinking about what we are doing. Let's stick to what we are talking about here. We are talking about a cage that has autologous bone graft in it, 12 month data and some 24 month data.

Let's go to Dr. Laurencin and let him ask his questions.

DR. LAURENCIN: Just a comment and question.

This is a study of 145 of these devices of which we are -- the usual norm would be to have a two year follow-up point but we have -- that means we have less than maybe 15 percent being followed up at two years. But really when we look at it, we are asking whether we can use a 12 month follow-up point, but we have 53 or only a third of these patients followed up for 12 months.

My concern is the fact that we are looking at a study in which -- a study of 145 patients, 145 implants, in which we only have 33 percent of them at a 12 month time point, a one year time point, when the standard in terms of the literature -- if we were doing a study like this and we are going to report this in the literature, we report a two year minimum time point. We have only got one-third at a 12 month time point.

MR. LIPSCOMB: Bailey Lipscomb, Sofamor Danek.

This PMA -- you are right, your numbers are absolutely right. This PMA is supported with at this point in time the randomized cohort, which is about 97 patients. If you look at that, the percent at one year is approximately 70 percent and the percent at two years is 30 percent.

I know if you go on and you add the non-randomized patients, you are absolutely right. I just wanted to clarify that point.

DR. LAURENCIN: But you made your study non-randomized in order to get -- because you felt there was a need to get more patients in the study. But then you -- but now, I mean, it is almost like at this point you are saying that we needed more patients in the study, but let's look at only the patients that were randomized.

MR. LIPSCOMB: Well, those are the patients we have at one year right now. There are no non-randomized patients at one year.

DR. LAURENCIN: The point is that in the group of 145 in your total cohort, it is only one-third that have reached their one year point. The question is whether that is too soon to be able to -- if you have a group of 145

patients that you are demonstrating as your patient population for a study that only one-third have reached the 12 month point to be studied.

MR. LIPSCOMB: It is certainly a valid question, but I think my response to that would be is the fact you have to go back and you have to look at the numbers that Dr. Kamer put up, the numbers that we put up. It is an equivalent study. The delta values at 12 months were met and even a lot of them for 24 months were met.

So, you know, the power is there.

DR. WILKINSON: But it is equivalent only to a terrible control.

MR. LIPSCOMB: I think if you look at why that control was chosen at that time, it was a suitable regulatory control, plus there was literature supporting that control group.

Obviously, you have looked at the fusion tables because you raised a question about our fusion rates and the reporting of fusion rates and why they were different at six months. I would like to clarify that. It is the fact that the first time in the original submission that went in, the fusion rate -- there was difficulty in measuring bridging trabecular bone and if the radiologists couldn't measure

bridging trabecular bone, it was counted as a non-union or not a fusion.

DR. LAURENCIN: What is the mean follow-up right now of these 140 some patients? If you had to say how many -- the average time that that patient is out right now.

MR. LIPSCOMB: I don't have an average number for everybody in the whole study.

DR. LAURENCIN: Could you approximate it --

MR. LIPSCOMB: Based on the data that is presented to you right now, we have a certain percent, 30 percent at two years. We have 70 percent at one year.

DR. LAURENCIN: About 15 percent at two years of this 23 out of 145.

MR. LIPSCOMB: Well, we have -- if there are 207 people and you have approximately 97 at one year, that says approximately half your people at one year.

DR. LAURENCIN: I am talking about the Novus implant group.

MR. LIPSCOMB: Okay. About one-third at one year based on this present data analysis.

DR. LAURENCIN: No. I am just talking about the Novus group that has 23 out of 145.

MR. LIPSCOMB: Did you say one year or two years?

DR. LAURENCIN: I am just asking what the average for the entire group, the mean time out for the entire group.

MR. LIPSCOMB: I am sorry, sir. I don't have an average number for that. I mean, if somebody were enrolled yesterday, that would be one day and had surgery. So, I do not have that mean value.

DR. BOYAN: Maybe we can address some of these other issues as we answer the --

MR. LIPSCOMB: But I would like to clarify one point because of the way it was stated, is the fact that on that fusion rate at six months is the fact the difference between those rates were due to the fact that in the first time if the radiologist could not measure bridging trabecular bone, it was counted as a non-union; whereas, the second time, it was counted as missing data.

All you have to do is look at the denominators and see how that shakes out. I just wanted to clarify that point.

DR. BOYAN: Okay. Thank you.

To the Panel, I want to give you very briefly -- I mean, this is if there is like a burning question, a factual question that can be addressed that you need in order to

make -- to go to the next set of questions that Mr. Williams has put in front of us. If there is a burning question that should be addressed to FDA or to the Panel, raise your hand. Otherwise, let's -- I think we will probably cover most of the things we need to cover in addressing these questions.

Now that I have taken all the fun out of doing it, so, we will now go to the questions and Mr. Williams -- there has been a request on the part of certain individuals that there be a quick break. I am a professor. I am really hesitant to let anybody out of this room, but I think that it may be necessary for us to do that.

So, I am going to count exactly five minutes. If you want to be here when the questions get addressed, you will be back in five minutes.

[Brief recess.]

DR. BOYAN: We will begin with the questions on the PMA. I would like to ask Mr. Williams to come back and lead us through the questions one at a time, beginning with Question No. 1.

MR. WILLIAMS: Okay. With regards to Question 1, Question 1 is dealing with the reanalysis of the Oswestry Disability Questionnaire and it has to do with the alternative approach that is presented by Keurskens in the

Journal of Pain.

I will just go ahead and read the brief intro and then the question and then it will be followed up by discussion.

In response to FDA concerns regarding the clinical relevance of a difference of 1 point as measured by the Oswestry Disability Questionnaire, the sponsor submitted an alternative approach to analyzing the results of the Oswestry. This alternative approach as presented by Beurskens in the Journal of Pain defined a clinical significant improvement as a decrease in an Oswestry score of 4 to 6 points out of a possible 100.

Do the differences of 4 to 6 points out of a possible 100 represent a clinically significant improvement within patient in the Novus LC Interbody Fusion Device?

DR. BOYAN: For this series of questions, what we will do, we will -- with each one we will start with a different person. With this one, I would like to start with Dr. Wilkinson. If you don't have -- I thought it was a pain question -- if you don't feel obligated to make a comment, you don't have to. Just say "no comment" and we will go to the next person.

DR. WILKINSON: Well, I think I really made my

comment about this question earlier, that a 1 in 5 improvement in only 2 in 10 categories is simply not enough to satisfy my patients. And it might satisfy a statistician, but it is not going to satisfy a patient. It needs to be better than that, a lot better than that.

DR. BOYAN: I think we will go next to Dr. Holeman. We are going to go in the opposite direction.

DR. WITTEN: Excuse me.

DR. BOYAN: Dr. Witten.

DR. WITTEN: Can I be recognized?

DR. BOYAN: Right.

DR. WITTEN: This is Dr. Witten. I would just be interested in hearing what amount of improvement you would consider a clinically significant improvement.

DR. WILKINSON: I don't think a patient is going to be happy until they have at least two points of improvement in at least eight out of ten of these scores. So, if that is -- two times eight is 16 times two is 32 percent improvement. Just off the top of my head, I think a patient would consider that a pretty minimal improvement to go from intense pain to moderate pain half of the day is pretty limited improvement.

DR. BOYAN: Dr. Holeman.

DR. HOLEMAN: My comments are along the line of Dr. Wilkinson. Pain is a personal experience. Self-report is a good instrument, kind of tool to use, to solicit information relative to pain. I do feel, however, that if a patient has been significantly helped by the insertion of the device and this is self-report, that a better self-report score should be presented, as opposed to a 4 to 6 point difference out of 100.

So, I don't think that a 4 to 6 point difference would indicate clinically significant improvement; perhaps mechanical but not clinical.

DR. BOYAN: Dr. Silkaitis.

DR. SILKAITIS: For me, I guess it would be very difficult to determine whether that particular 1 point would be valuable to the patient or not. We are talking about, I think, means here and it is group data. So that if we are talking about group data and group means and moving towards a difference in means -- and I am not a statistician, but, obviously, there is something happening and that should be taken into account. But I do defer to the clinicians for that response.

DR. BOYAN: Okay. Coming around, Dr. Cheng, do you have any statements on this question?

DR. CHENG: I don't consider it a clinically significant improvement and I think it is a very lenient definition of that and no difference was seen with the control, meaning there was very likely no benefit to the device over the control, however good or bad the control is.

DR. WITTEN: Excuse me. This is Dr. Witten.

I just have the same question for Dr. Cheng as for Dr. Wilkinson. Based on your knowledge of this scale, can you characterize what you would consider a clinically significant improvement?

DR. CHENG: I would have a little difficulty answering that because I think it requires a statistician. Outcome measurements are validated and there are very specific means of doing that. So, I would defer that to the statistician.

DR. BOYAN: Dr. Kerrigan.

DR. KERRIGAN: I really don't have any substantive comments.

DR. GROBLER: I feel I would accept that for the following reasons: One, that this is a measurement we use all the time and I think if that is only one piece of the whole picture, pain is only one component of the subsections of the Oswestry and a patient coming to you with pain, if

you statistically work that out, that is going to be a very small piece. So, in the big picture, taking all the other aspects into consideration, I would really not take this one aspect as a big factor and I would accept that in the big picture.

Secondly, I presume Pain is a peer reviewed journal and, therefore, I presume the methodology used there should be acceptable. And I stand to be corrected. I did not go into that article, but I do feel this is only one point and I would in the big picture be very hesitant to take this as the only factor.

DR. BOYAN: Dr. Hale.

DR. HALE: I have no comment.

DR. BOYAN: Dr. Yaszemski.

DR. YASZEMSKI: No comments on this question.

DR. BOYAN: Dr. Janosky.

DR. JANOSKY: I think to answer the question we need to think about the two components that are going into any score and one is clearly the actual process that is going on and the second is the scale itself; namely, the measurement error associated with that scale.

If I look at the standard deviations and what I know from looking at the manual of the Oswestry, this

actually is taking into account what is standard error within any accepted score. That is a statistical consideration and then from that comes a clinical significance as to what you are measuring, whether that meaning makes any sense or not.

So, I would say "no," that this needs to be revised, really strong statistical and clinical.

DR. BOYAN: Dr. Laurencin.

DR. LAURENCIN: I can just only comment from a clinical perspective that in terms of the -- I don't know a lot about the scale, but, obviously, any scale in which one can have increased pain after the operation would have an improved score, has to be called into question as one of the measures for the outcome.

DR. WILKINSON: I think really what I am trying to say --

DR. BOYAN: This is Dr. Wilkinson.

DR. WILKINSON: This is Dr. Wilkinson again. You are probably familiar with my voice by now.

What I am really trying to say is that what I would like to hear is how many patients were significantly improved, not what the statistics looked like. If a patient has three points worse on two activities and three points

better on another activity and one point better, then he is improved. But you ask the patient, I am worse off, Doc, because I need to be able to climb a ladder to earn a living.

So, I need to know really not what the statistics say. We are talking about human beings here. How many individual patients say they are or using a scale are significantly improved.

DR. BOYAN: Is that -- FDA, does that answer your question?

MR. WILLIAMS: Yes.

DR. BOYAN: Okay. Question No. 2.

MR. WILLIAMS: Question No. 2. In relation to the SF-36 assessment of quality of life, the developers of the tool grouped the data into two summary scales, which was provided in Amendment 7 in your panel packs. Based upon multiple evaluations of these scales, they derived a numerical measure of the data dispersion. They found that scores of 5.42 and 6.33 in the PCS and MCS summary scales, respectively, equal two standard errors of the measurement.

Assuming that this equals two standard errors of the mean, a difference between the two groups greater than these scores may identify a statistically significant

difference between groups.

This question is similar to the question on the Oswestry, but we have a more specific question that we would like to get comments from the statistician and that is: Is this a valid interpretation of the use of these scales?

DR. BOYAN: So, let's just stop with that and turn to Dr. Janosky.

DR. JANOSKY: In addition to this, I see a review and a memo in there from Mr. Kamer addressing this issue. I couldn't find any information to tell me exactly what the SEM was. I am assuming that it is a standard error of measurement based upon the published scale done on a totally different population.

MR. KAMER: Yes, I believe it was, but it wasn't perfectly clear, if I remember correctly, what it was.

DR. JANOSKY: I will answer based on that assumption because I have come to the same assumption and that is most likely what it is.

If that is the case, then I would conclude that "no," this was not a valid interpretation. Typically, we use SEMs to have an estimate of an individual score variability, not in terms of a group variability. So, I would suggest one of two approaches. One is to calculate

the SEM from the population of the baseline measurement in the current study and/or use the standard deviation either from the published scale or from the current study.

So, the answer is "no," it is not a valid interpretation.

MR. WILLIAMS: Thank you.

And a more general question for the Panel is: Sofamor Danek used this to demonstrate changes from pre to postoperative time points; that is, within patient assessment. Is this clinically meaningful?

DR. BOYAN: So, let's start -- we have never started with Dr. Grobler. So, we will start with Dr. Grobler and we will go that direction.

DR. GROBLER: I guess this is anticlimax. I don't have a response to that one.

DR. BOYAN: Dr. Kerrigan.

DR. KERRIGAN: I don't either.

DR. BOYAN: Dr. Cheng.

DR. CHENG: Only if it is measured compared to a control group.

DR. BOYAN: Dr. Silkaitis.

DR. SILKAITIS: No comment.

DR. BOYAN: Dr. Holeman.

DR. HOLEMAN: No comment.

DR. BOYAN: Dr. Wilkinson.

MR. WILLIAMS: Of course. Basically, the same comments. This, again, is a scale that does not seem to weight the severity of the change. I am not familiar with the score. I don't know that we were given a copy of it. It may have been in one of those crates I left at home, but if the answers are plus and minus only, that is a very poor way of assessing patient satisfaction. If they are worse in one area and better in another, the score may come out good and the patient be dissatisfied.

So, I think all of these instruments are very interesting, but I really want to know what the individual patient feels about his or her outcome.

DR. BOYAN: Dr. Laurencin, do you have anything?

DR. LAURENCIN: I think I agree exactly with what Dr. Wilkinson said.

DR. BOYAN: Thank you.

Dr. Janosky, you have already basically answered it, right?

DR. JANOSKY: Right. I would suggest that if it was used, there was a window of error placed on there.

DR. BOYAN: Dr. Yaszemski.

DR. YASZEMSKI: No comment on this question.

DR. BOYAN: And Dr. Hale.

DR. HALE: No comment.

DR. BOYAN: No comment. Okay.

I think we have done the best we can do for you on that one. Is there anything else that you would like us to address on that?

MR. WILLIAMS: No. I think that was adequate.

Question 3, the IDE protocol proposed that success of the study when comparing the investigational and control devices, be based upon attaining equivalence. In the IDE, the equivalence was defined as the investigational device performing no more than 20 percent worse than the control devise at the 24 month time point.

The first question is, Sofamor Danek is asking that we consider the same definition of equivalence based upon the same measurement tools at 12 months rather than 24 months. Is this an appropriate definition of equivalence at 12 months based on the following concerns, which are: (1) The sponsor assessed fusion radiographically using the following four criteria, evidence of bridging trabecular; translation, angulation, radiolucency using plain film techniques.

We have a concern that it was difficult to determine the presence or absence of trabecular bone into and through the Novus LC spinal cage using the validated radiographic technique currently available.

At earlier time points, stability, as measured on flexion and extension x-rays, may not be reflective of solid fusion, but rather may be due to instrumentation and/or hyperdistraction.

And then (2) previous spinal cage studies have assessed back pain, function and neurological status at 24 months. We have a concern whether assessments at 12 months utilizing these tools provide comparable reliability.

DR. BOYAN: Dr. Hale, would you like to begin?

DR. HALE: Well, without reiterating the whole discussion we had this morning, I guess, my sense is that 12 months is probably not an adequate predictor, that there are not data available to demonstrate the 12 months that does predict the 24 month outcomes. So, I would be inclined to say that "no," it is not equivalent.

MR. WILLIAMS: Okay. I want to go ahead and clarify the question again because I think our concerns kind of overflow into each other and I would like to make them as specific as possible for the Panel.

Dr. Hale has pretty much started to answer the second part of this question and I would like to reiterate what the first part of the question is and try to clarify it a little bit more. The first part of the question is the study was originally defined as an equivalent study for 24 months and our concern is can you just take that equivalence definition and apply it to 12 months. So, that is the first part of the question.

DR. HALE: I am not sure that I can really answer that. I mean, I think that may be more of a statistical matter. It seems to me that that should be defined before the study was undertaken, though. It is not something that you would define midstream, that you were going to change your endpoint.

MR. WILLIAMS: Okay.

DR. BOYAN: We are going to go my direction because I am just bursting -- what?

DR. WITTEN: I just want to clarify, though, what Paul has already clarified.

DR. BOYAN: Dr. Witten.

DR. WITTEN: Yes. It is Dr. Witten. I am sorry -- which is that that part of the question is related to the delta, which was defined before the study was begun at 24

months. I think the question means is there a different delta that might need to be considered or done over a different time point other than 20 percent.

MR. WILLIAMS: Are deltas dependent upon the endpoint time point?

DR. BOYAN: I think the problem is the control and not the experimental because they cannot possibly be the same at 12 months, where they could be the same at 24 and that is because of the way that the body handles the two therapeutic modalities. They both have autologous bone graft. In my opinion, there was probably more autologous bone graft in the cage than there was in the allograft. The allograft also, in my opinion, was thicker than was the wall of the cage.

So that some of what appears to be bone at 12 months, new trabecular bone, may, in fact, be cancellous bone that is residual from the bone graft and may give the appearance on a plain radiograph as being bone bridging when, in fact, it may not really be bone bridging. However, I do believe that the impression you would have a fusion mass would be better at 12 months in using the cage than the impression of a fusion using the allograft.

By 24 months, the allograft will have been

remodelled and the two sets of data may begin to look very, very similar and, in fact, the equivalent. So, your criteria for equivalency at 24 months cannot be the same criteria for equivalency at 12 months.

Dr. Yaszemski.

DR. YASZEMSKI: Michael Yaszemski.

I think that in a subsequent question the issue of type of radiographic analysis is going to be raised, but the way I am going to answer this one is I would accept those data at 12 months with respect to the radiographic demonstration of union but not for the clinical outcomes.

DR. JANOSKY: Different time points are requiring different deltas. So, I would say "no."

DR. BOYAN: Dr. Laurencin.

DR. LAURENCIN: From my knowledge of statistics, I think that what Dr. Janosky is saying is what I would reiterate that at different time points there are different criteria that are there.

But from another standpoint, again, this is a study that has 145 patients in the Novus group or so, of which only one-third have actually reached the 12 month point and, so, even if we look at the 79 patient group, only 60 percent, 67 percent of those people have reached the 12

month point. So, this isn't even a 12 month study yet.

DR. BOYAN: Dr. Wilkinson.

DR. WILKINSON: Well, the question really has three parts and my answer has three parts. No, no and no. Not being a statistician makes that easy.

DR. BOYAN: That was succinct.

Dr. Holeman.

DR. HOLEMAN: No, no comment.

DR. BOYAN: Okay. Dr. Silkaitis.

DR. SILKAITIS: No comment.

DR. BOYAN: And Dr. Cheng.

DR. CHENG: I don't think I have enough knowledge of statistics to answer the first part.

DR. BOYAN: Okay. Dr. Kerrigan.

DR. KERRIGAN: I agree with your comments, Dr. Boyan, and the only other -- and my answer is also "no, no, no," but I just want to comment. Again, my concern about the 24 month, the difference in the control, whether or not -- even if the control -- we are not sure if there really was a higher success in fusion comparing the Novus group to the control group, but in any event what we -- I think that the main thing that we should be looking at is pain and it is not fusion rate and looking at these scores, definitely,

I see no evidence of improvement or I definitely see a difference in the control between the 12 month and 24 month and that is what needs to be investigated. Why are we seeing this difference from 12 months to 24 months? And I know the numbers are small, but clearly the -- it is necessary to look at the 24 month.

And then just with my other concerns of what is the -- what are sort of the things that are difficult to measure? What is the -- how the -- the fusion occurring at one level, how does that impact on other levels and other functioning and was that the problem in the first place, the longer time period, maybe that is why we are seeing some of the difference between the 12 month and the 24 month.

So, I think 24 months is the answer.

DR. GROBLER: I would once again -- we are dealing with patients and the fusion rate to me is only one step. I would accept the 12 months if more patients came to the 12 month stage. I think there is not enough patients that reach 12 months. If at 12 months, there is statistically enough patients, I think you can assess radiologically if they are fused.

The question to me then which I would say a "no" for is are they clinically between one and two years, are

you getting the appropriate response, but I think what this device intends to do is to fuse two segments. We are not saying that the clinician made the right decision, that fusion is the treatment for this disease. What this device needs to do is to fuse the segments.

We assume that fusion is going to treat the pain, but I would just urge that we should be very careful that we don't think that the metal is the reason to treat the pain. The metal -- the whole construct is only to fuse the segment. We as clinicians think that if that segment is fused, the pain will be better.

The issue at stake is if it is fused, I think you can see it at 12 months. I am not sure about the second year's clinical outcome and on that basis I would say "no" for the clinical assessment.

MR. WILLIAMS: Okay.

DR. BOYAN: Have we adequately addressed all the parts of 3 by accident. We didn't do Part B yet.

MR. WILLIAMS: I believe the majority of the Panel members have touched on all parts in some shape or form, but I would like to go ahead and just read Part B and if any of the Panel members have any additional comments with regards to those questions, we would definitely like to hear them.

Part B states: Are 12 month values as predictive as 24 month values for a given treatment? Are 12 month differences as predictive as 24 month differences between treatment groups?

DR. WITTEN: Excuse me.

DR. BOYAN: Yes, Dr. Witten.

DR. WITTEN: I believe that you have already touched on these, but if anybody has any additional comments, we would certainly be glad to hear them.

DR. BOYAN: Are there any additional comments?

Okay.

MR. WILLIAMS: Okay. As stated at the end of our presentation in the initial reading of the questions, that we were going to change the order. So, I am going to jump to Question 5.

The sponsor introduced reconstructed CT as supportive evidence of fusion. Are CT scans as taken, reconstructed and interpreted in this study an appropriate methodology for assessing fusion in the presence of morselized autograft? And are reconstructed CT scans at earlier time points predictive of long term fusion?

DR. BOYAN: Let's begin with Dr. Yaszemski.

DR. YASZEMSKI: We heard before, I believe, from

Dr. Heim, if he is still here, that he has been doing this for awhile and that with the advent of the 1 millimeter cuts, he felt that in those patients he declared fused that this was a reasonable assumption.

I have looked and I have not found this in the literature. I would like to find it or if someone is aware of it, I would like somebody to refer me to the reference. I think that from a scientific perspective if it is not yet in the peer reviewed literature, I am going to have to answer "no" to that, although just from my perspective as a clinician, I think it is a very reasonable thing and I anticipate it will be okay, but from a purely scientific viewpoint without being aware of a study being published as of this point, I am going to have to answer "no."

With respect to the second question, I am going to answer that two ways. Of course, if my answer to the first one is "no," then the answer to the second one has to be "no," but if we assume that the answer to the first one were "yes" -- let me comment on that -- just so that I can speak to the point of the earlier time points -- I think that once a fusion happens, the fusion is there.

So that if by whatever means you had an adequate way of saying that a fusion existed at an earlier time

point, then, yes, that would be predictive of that fusion being present at a later time point.

DR. BOYAN: Dr. Janosky. I was busy looking at the CT scans.

Dr. Laurencin.

DR. LAURENCIN: I think we have to use what we have available. This may be the best thing that we have available right now in terms of CT scans, reconstructions, the small cuts. Are reconstructive scans, earlier time points predictive of long term fusion? Probably they are if they can -- but, of course, the one question we have is what sort of integrity is present in that fusion, what sort of strength characteristics that fusion has. We may not be able to tell from CT scans alone.

DR. WILKINSON: Well, I do think good CT scans are very important and very helpful. However, I think that validation of this as a predictor of bone fusion is necessary; some correlation with the strength of fusion, I think, is extremely important.

The criteria used to evaluate these need to be looked at very carefully. Bridging bone that is not anchored at both ends doesn't do much good. So, you really have to have very careful criteria for these CT scans. I do

not think CT scans should supplant flexion/extension films, which give a different kind and supplemental type of data.

So, I think you need both things, but I really believe the CT scan is going to be very, very helpful in seeing the anatomy, at least, of bone. Are reconstructive CT scans predictive of long term fusion? No. We have certainly seen instances where the fusion was not stable at one year and became stable at two years.

Will we find instances where fusion bone resorbs between one year and two years? We certainly see bone resorb around other metal implants. That is expected in the literature of spinal implants. So, I want somebody to prove it to me before I am willing to accept it now.

DR. BOYAN: Dr. Holeman.

DR. HOLEMAN: No comments.

DR. BOYAN: Dr. Silkaitis.

DR. SILKAITIS: Yes. I would like for industry a little bit of a clarification. Are the Panel members saying that if you are using a newer technology, diagnostic technology, that it has to be published in the peer reviewed journal first before it can be utilized in clinical studies?

DR. BOYAN: Dr. Yaszemski, I think you are the one that made the comment.

DR. YASZEMSKI: I can just speak for myself. I would like -- if somebody is going to propose to me that they have something new and it works, it would help me if those who perform peer review have looked at it ahead of time, somebody other than the person proposing the technique has said, yes, I have looked over their scientific method. I have looked over their presentation and I think it is reasonable.

So, for me, personally, that would help.

DR. WILKINSON: I think it has to be validated. If it is validated in the peer review literature, that takes a lot of work off of the manufacturer. If it is not validated in the peer review literature, then the manufacturer has to do the validation.

DR. SILKAITIS: Okay. I am glad you added that because if the manufacturer validates it, then it is perfectly acceptable is what I hear.

DR. BOYAN: And I think that is directly apropos to the question at hand is that as you look at the CT scans, they are quite convincing. But it would have been helpful to have had at least in an animal model, recognizing that the human wouldn't have been accessible to us to validate against -- it would have been nice in an animal model to

have CT scans side by side with conventional histology to compare what one looks like and the other one looks like in this situation and also next to the radiograph, the plain radiograph.

Dr. Cheng.

DR. CHENG: No, I don't think it has been validated in the literature yet as to the benefit of CT scans in their reconstruction. It may be a promising tool, but I don't think it has been validated.

As an example, if you take a cavity of bone and you pack it with bone autogenous graft or allograft for that matter and you take a CT scan in that area, you are going to see cancellous bone throughout that area, but I don't think anyone would say it is fused.

DR. BOYAN: Dr. Kerrigan.

DR. KERRIGAN: No comment.

DR. BOYAN: Dr. Grobler.

DR. GROBLER: I really think that CT scans is a new technology. We are all using it. To reduplicate a study to show that CT scan is a new technology and gives you more information in a certain field, I think it is going to most probably take a couple of years. Secondly, I don't think Harrington(?) rods, if we have a study with Harrington

rods on the table now, we most probably won't be able to approve that because there is no study that will show or we won't accept CT scans as a methodology.

I think CT scans is a methodology with the correct cuts, the correct interpretation and I would say "yes." I would take that as an interpretation of fusion.

The second part of the question, I truly believe if a fusion is solid at 12 months, it is solid at 24 months. If it is not solid at 12 months, you most probably have not had a methodology to adequately assess it.

DR. CHENG: Dr. Grobler, when is a fusion not solid?

DR. GROBLER: A fusion is not solid if, one, on flexion/extension views, there is motion, as we indicated. Secondly, if there is not continuous trabecular bone throughout the construct of fusion. I mean, how do we all -- I mean, how do we decide a patient can go back to work. I mean, I need to make a decision at some time any patient that I do a fusion on. I need to say, Mr. Joe Blow, you can go back to work now and I will not make that decision if I am not sure the patient is fused.

So, everyday in the clinic we make that decision. I use CT scans and I think CT scans is there to be used. It

is perfect if you have a controlled study and I am sure one of us in academic position now will go back and try and code one of our clinical residents -- our residents to do it, but the point of the matter is I think that is a state of the art to assess fusion and I personally would accept that.

DR. BOYAN: Dr. Hale.

DR. HALE: I would agree that this is a promising methodology but perhaps one that has not been adequately validated at this point. I also would echo Dr. Wilkinson's comments about the scoring scheme that was used and take exception to the way that some of these CT scans were categorized in terms of non-adjacent areas being counted as a higher level of fusion.

With regards to the second part of the question, I am not sure that I feel qualified to answer that.

DR. BOYAN: Whether they are predictive of long term fusion?

DR. HALE: I am just -- I mean, I think if you accept that CTs are an appropriate methodology, then I don't see any reason why they wouldn't be, but I don't have the clinical experience really to comment on that.

DR. BOYAN: Okay. So, to paraphrase what I heard here is that in general the opinion is that CT is a

promising technology, that it has applicability in this particular clinical situation, that a fusion, once it occurs, is unlikely to go away and that it is -- am I stating this correctly, Dr. Grobler -- and that if there was evidence of a fusion early, it would probably still be there later. Is that -- did I cover the general feeling of the group, that we are not making the statement -- we are not going the next step with this.

DR. WILKINSON: With the caveat that you really have convincing proof that the methodology shows the fusion is there and is stable.

DR. BOYAN: Correct. The validation, I think, was an important concept that we got through.

Okay. Question No. 5.

MR. WILLIAMS: Okay. Well, it is actually reversed. Question 4, this question we would like you to pay particular attention to. We would like general comments and keep it in mind in the final voting on the product.

The question is: The sponsor stated that they found that the Novus LC device had an increased rate of fusion and an equivalence for safety and other effectiveness parameters in comparison to the control, femoral allograft ring, at the 12 month follow-up time point as per the

success criteria defined in the original protocol.

Assuming that the results at the 12 month follow-up time point are sufficient to make an evaluation on the safety and effectiveness of this PMA and the results at 12 months of the entire cohort of randomized patients, meaning that all patients in the randomized part of the study have reached their 12 month follow-up are consistent with the results presented to date, would the results presented support a recommendation of approval for the Novus LC device?

And I would like to just add to that that if you have any comments regarding this question, please elaborate and be as specific as possible to support your comments.

DR. BOYAN: I need a point of clarification from either Dr. Witten or Ms. Nashman. If we address this question, we are not yet voting. Is this --

MR. WILLIAMS: No. We are hoping this is going to be --

DR. BOYAN: This is informational.

MR. WILLIAMS: This is just an informational and then an official voting to follow.

DR. BOYAN: Okay. So, Panel, please, how you state your answer to this, keep this informational. We are

not at the -- this is penultimate. We are not at the ultimate yet. Okay.

Dr. Laurencin, why don't you start us off.

DR. LAURENCIN: I am not quite sure because I think that there are some questions that have been raised in terms of the statistical group that is needed. I think the question is whether the initial power analysis called for 79 patients in the randomized group to be studied or 140 patients to be studied.

So, I think that is a very important question. Obviously, there is thought that if we are shifting from two years to a one year time point, then there are greater statistical concerns that have to be met and the question is whether that non-randomized cohort that said 79 right now is valid or does one need to have an entire group?

I think that is the first question. The second is whether the types of questions that have been asked of the patients in terms of pain, in terms of outcome are actually going to be -- are compelling enough to say that at the 12 month point, that the outcome is -- the outcome for these patients is good.

I think these are some of the questions that are there that I have about it.

DR. BOYAN: Dr. Wilkinson.

MR. WILLIAMS: Even if I can be sure that my two year old grandson is going to Harvard, which is what you are asking us to do here, I would say "no." I think I have serious questions about the data that has been presented today at the 12 month level. I think even if the data is pristine at the 12 month level, you need to show 24 month data.

There are too many questions that have been raised for the data between 12 months and 24 months. The animal studies leave open-ended questions. The clinical data leaves open-ended questions, that clearly the mechanics are changing. The disc spaces are narrowing, that clearly the fusion patients are doing -- the control patients are doing better at two years and the treatment patients are doing worse. There are too many questions that have been raised.

So, no matter what data you show me at the end of 12 months, it is not going to be adequate.

DR. BOYAN: Dr. Holeman, do you have a comment?

DR. HOLEMAN: Not really, other than just to say that 12 months probably would not provide enough opportunity for the physician to make an assessment of what activities can and to what extent a patient can participate or the

surgery has provided an opportunity for the patient to participate in activities of daily living.

So, I don't think 12 months.

DR. BOYAN: Dr. Silkaitis.

DR. SILKAITIS: At this time I don't have a comment.

DR. BOYAN: Dr. Cheng.

DR. CHENG: I am not sure why the FDA is asking this question, but I am going to decline to answer as I don't feel I can give you a valid opinion on that now.

DR. BOYAN: Dr. Kerrigan.

DR. KERRIGAN: I agree with everybody so far and, yes, I find that question a little bit tricky because I don't go along with the assumption. So, it is kind of hard to answer that question.

DR. BOYAN: This is one of those questions. That is why I asked if this was informational or if we were there.

DR. GROBLER: I will be prepared to answer the question if my assumption is correct. If my assumption is that if I have these data points and these data points were validated by a full follow-up of all the patients and this database is transferred to 24 months and we are sitting in

this room a year later and the same figures pop up; fusion rate a hundred percent for Novus, 42 percent for control; Oswestry, 88, 69, 84, 74, neurological hundred percent, hundred percent; overall 72 and 27. And all these statistical issues have been cleared, then I would say "yes," that we don't have that --

DR. WILKINSON: 24 months or 12 months?

DR. GROBLER: 24 months. I mean, we are asking to assume that the 12 month data is transferred to 24 months. Is that correct? Or what is the question? I thought the question was if we have this table that you gave us on this page, if these data points are identical at 24 months and all the patients followed up correctly to the time that they need to be assessed, would we accept these results? Is that what the question is?

DR. WITTEN: Actually, the question actually says if all the patients reach the 12 month time and we got these numbers -- because I think it is already been pointed out that it is not the majority of the patients who have reached 12 months. So, the question actually is if all the patients reach 12 months and the 12 and 24 month data for those patients who reach 24 months still look like this, would this be approvable? So, it is really addressing that issue.

But I think that the question that you thought we were asking is a question we would like to ask, too. So, you can answer that one as well. We would be interested in that answer, too.

DR. GROBLER: Well, if I have these figures at 12 months and all the patients came to the endpoint and this is totally reproduced as it is to 24 months, then I would accept it.

DR. BOYAN: Dr. Hale.

DR. HALE: Based on Dr. Witten's clarification, if that is the assumption we are to make, I guess I would say I would accept this as well.

DR. BOYAN: Just to make sure that I understood what you just accepted, what I heard Dr. Witten say was that if all -- when all the patients reach the 12 month time point, the data still look like this, that that would be sufficient.

DR. GROBLER: No. What I understand --

DR. BOYAN: Is that what I heard you say?

DR. WITTEN: That is what I actually asked, yes.

DR. BOYAN: Okay. Now, let them answer again because I --

DR. GROBLER: Well, then I need to answer again.

If these data, if everyone comes to 12 months and this is the data that we have and it is 12 months is completed, are you asking if the same data is seen at 24 months --

MR. WILLIAMS: No.

DR. GROBLER: You are not asking that?

MR. WILLIAMS: No. We are saying once all the --

DR. GROBLER: So, I can forget 24 months because -- I am not sure why this was put in here.

MR. WILLIAMS: Once all the patients have reached -- well, specifically this question is asking once all the randomized patients that have been enrolled in the study has reached 12 months, if the rates that you see are the same rates, is the device -- would it support a recommendation of approval?

DR. GROBLER: I would defer my comment then.

DR. HALE: I think I misunderstood that then. I would expect if they were at 24 months, that that would be acceptable, but because of the other issues that we have already talked about, I don't know that I would accept them at 12.

DR. BOYAN: Is there anybody else that -- now that that has been clarified wants to make any more comment?

[There was no response.]

Okay. Now, have we addressed the issues that were important to FDA for us to address?

DR. WITTEN: Yes, thank you.

DR. BOYAN: Okay. So, we are there.

I would now like to ask Ms. Jodi Nashman to explain to us how we are to vote, what our instructions are.

MS. NASHMAN: Jodi Nashman.

The Panel is about to make a recommendation for the premarket approval application put before us and before they do that, I am going to read the options that the Panel has in making a recommendation.

The Medical Device Amendments to the Federal Food, Drug and Cosmetic Act require that the Food and Drug Administration obtain a recommendation from an outside expert advisory panel on designated medical device premarket approval applications that are filed with the Agency.

The PMA must stand on its own merit and your recommendation must be supported by safety and effectiveness data in the application or by applicable publicly available information. Safety is defined in the Act as reasonable assurance, based on valid scientific evidence that the probable benefits to health under conditions of use outweigh any probable risks.

Effectiveness is defined as reasonable assurance that in a significant portion of the population, the use of the device for its intended use and conditions of use when labeled will provide clinically significant results.

Your recommendation options for the vote are as follows:

One, approval. There are no conditions attached.

Two, approvable with conditions. You may recommend that the PMA be found approvable subject to specified conditions, such as resolution of clearly identified deficiencies, which have been cited by you or by FDA staff.

Prior to voting, all of the conditions are discussed by the Panel and listed by the Panel chair.

You may specify what type of follow-up to the applicant's response to the conditions of your approval recommendation you want; that is, FDA or Panel. Panel follow-up is usually done through homework assignments to the primary reviewers of the application or to other specified members of the Panel.

A formal discussion of the application at a future Panel meeting is not usually held.

If you recommend post approval requirements to be

imposed as a condition of approval, then your recommendation should address the following points.

(a) The purpose of the requirement.

(b) The number of subjects to be evaluated.

(c) The reports that should be required to be submitted.

The third option is to recommend disapproval or not approvable.

The Act specifies for denial of approval, the following provisions are applicable to Panel deliberations:

(a) If the data do not provide reasonable assurance that the device is safe under the conditions of these prescribed, recommended or suggested in the labeling.

(b) Reasonable assurance has not been given that a device is effective under the conditions of these prescribed, recommended or suggested in the labeling.

(c) There has been a fair evaluation, a lot of material facts in the discussion, which determined the application to be false or misleading.

If you recommend that the application is not approvable for any of these stated reasons, then we ask that you identify the measures you think are necessary for the application to be placed in an approvable form.

So, again, your options are to vote or to recommend that the application be approved, be approvable with conditions or not approvable. Please note that following the voting, the chair will ask each Panel member to present a brief statement outlining the reasons for their vote.

At this point I will return to Barbara Boyan.

DR. BOYAN: Before beginning the voting process, I would like to ask the sponsor if there are any additional comments that you would like to make before we undertake the vote?

MR. LIPSCOMB: I guess the only comment we would like to make is we wouldn't want to lose sight of the fact that we did a prospective randomized control study.

The other thing, that we had an issue that came up about the number of patients and the number of people at one year versus whatever. It was a difficult procedure to run a prospective randomized control study and that is why we went to the non-randomized, was to foster that enrollment.

That was a major setback, trying to conduct a study of that order.

The other thing, too, when you get into the issues of the No. 1 and No. 2, which has probably already been

discussed to the hilt, I don't want to lose sight of the fact that those references that were chosen came out of literature articles in which they had done their own studies and they had come up with those points.

Of course, you know, your opinions are your opinions, but that wasn't -- those weren't numbers pulled out of the air. So, I just want to make sight of those facts.

DR. BOYAN: Thank you very much. This was Dr. Lipscomb that was just speaking to us.

I would like to mention both for the Panel's benefit and for the record that the votes taken are votes in favor of or against the motion made by the Panel. These are not votes for or against the product.

I would like now to entertain a motion. Is there a motion?

DR. YASZEMSKI: Dr. Boyan, this is Dr. Yaszemski.

I would like to make a motion.

DR. BOYAN: One moment on the motion. Okay. I am ready.

DR. YASZEMSKI: Dr. Boyan, I present to the Panel a motion of approvable with conditions. I would like to offer the following conditions:

No. 1, as Dr. Laurencin has brought up, I believe that the sponsor should continue to follow the cohort to increase the percent follow-up of the 145 patients.

No. 2, the vascular injuries, although many of them were minor, concerned me and I believe that the sponsor need to continue to collect and monitor data for intraoperative adverse events and postoperative complications.

No. 3, I believe the sponsor should report two year follow-up on all the patients as they become available and then present that follow-up to the FDA for their review.

Finally, No. 4, in line with Dr. Janosky's recommendations, I believe they should calculate the standard error of measurements for the persons in this cohort.

DR. BOYAN: Thank you very much, Dr. Yaszemski.

Is there a second to the motion? And I would like to remind the Panel that it is possible to amend the motion as necessary. We can adjust the motion right now if we need to address the motion to satisfy all parties concerned.

Are there -- may I ask you, are you making as a condition that it be approvable, but that the two year follow-up be part of that -- that it not be a postapproval

follow-up, that it be a preapproval follow-up?

DR. YASZEMSKI: That would be a condition of approval.

DR. BOYAN: Okay. Any other -- I need a second then.

DR. CHENG: Could you clarify what you just said?

DR. BOYAN: Well, maybe I didn't say it clearly, but what I was trying to say was that the data for the two year follow-up be presented -- what I was asking was was Dr. Yaszemski saying that the data for the two year follow-up had to be presented to FDA before they would consider making an actual approval.

DR. CHENG: Then I would like to amend that motion.

DR. BOYAN: Okay.

DR. CHENG: I would actually present the motion as not approved and the reasons for that would be the reasons as just outlined by Dr. Yaszemski. And I don't see a reason to approve it when the data has not yet been available for review. Why are we approving data, which has not --

DR. BOYAN: No, we are not approving. We are recommending -- the recommendation on the table is not to recommend approval. It is to recommend approval with

conditions. So the conditions that we are saying is it would be approvable if these conditions were met.

DR. CHENG: Well, I believe every PMA is approved as long as all the conditions for success are met. I don't see the reason to approve this one as it stands.

DR. BOYAN: Okay. Do I have a second?

Okay. What would you want to see before it could be approvable? All the things that Dr. Yaszemski just stated plus?

DR. CHENG: Basically, that is correct. I would like to see that data at 48 months. I think --

PARTICIPANT: 24.

DR. CHENG: Excuse me -- no, 24 months so that the complication rate will become a little bit better known at that time, as well as the results of the fusion and the assessment, the outcome assessment of the Oswestry pain scale can be looked at in terms of validation from a statistical point of view. There are two ways of measuring it and this way, one point, five points, that can be straightened out as well.

So, I would like to see those questions addressed before -- and reviewing that data before recommending an approval.

DR. BOYAN: Okay. I think we need to clarify this because I became -- I think what just happened here is something I am not quite sure about. So, I want to go back before we -- so that we handle this in the right parliamentary way. We had a motion. I asked for a second. I didn't hear a second. I heard silence. I waited a little while.

Now, you have come in with what is effectively a new motion. So, if there was a second for the first motion, they need to second now because we need to get this done right. We had a first motion.

Let me just review the first motion and then let me hear is, in fact, a second for it. The first motion is approvable with conditions. The sponsor should continue to follow the cohort all the way to completion, that vascular injuries need to be continued to be monitored and reported to FDA, that there needs to be a report of a two year follow-up, that it would be -- the report of the two year follow-up would need to be reviewed by FDA before the conditions would be met and that the sponsor would need to tabulate the standard error of the measurement.

Did I state what you wanted, Michael?

DR. YASZEMSKI: Yes.

DR. BOYAN: Okay. Now, that is the motion on the floor and is there a second for that motion?

DR. GROBLER: Am I permitted to ask a question before?

DR. BOYAN: Yes, a clarification.

DR. GROBLER: If it is approved with these points, that implies that the sponsor does not need to change their methodology. They just follow the patients on to a further point. Is that what you had in mind?

DR. YASZEMSKI: Yes.

DR. GROBLER: I would second that.

DR. BOYAN: Okay. So, now we have a second and now we can amend that and maybe at this point, Dr. Cheng, you would like to amend that motion with some of your concerns or are you still -- you are holding non-approvable recommendation in the wings here, but is there something with the current motion that is on the floor that you could amend such that your needs would be met?

DR. CHENG: For the vote of approval with recommendations, no.

DR. BOYAN: Yes, Dr. Janosky.

DR. JANOSKY: Point of clarification. Am I not correct in saying that if we are doing approvable with

conditions, we must have some assurance that the current application provides both safety and effectiveness data? And I think that is actually where we need to start because given we are putting these conditions on here, the total outcome from 24 months, we have no way of knowing what it is going to be with only 30 some percent of the patients followed up to even 12 months.

So, you are saying once you get 24 data, irrespective of whatever turns out, it is approved. I think we are really getting ourselves into a corner we don't want to be in.

DR. GROBLER: Excuse me. I didn't understand that. I understood that the FDA -- we will review it again and if --

DR. JANOSKY: No, no, no.

DR. LAURENCIN: No. We would approve it.

DR. JANOSKY: No, that is not what we are saying.

DR. BOYAN: What the motion presently is, just -- let me say it another way so that I can be sure that I am getting it right as well. We can open up this to discussion. This is not a concern. We have plenty of time to discuss this so we are all comfortable with the motion.

-- is that the cohort that currently exists be

continued to be followed to completion, that the question of vascular injuries is one of concern and needs to be carefully considered and reported adequately, that the two year follow-up needs to be done and the results reported to FDA and then FDA would come to some conclusion based on that data as to whether or not they wanted to approve or not approve. And, finally, that the sponsor tabulate the standard error of the measurement.

MS. NASHMAN: Just at this point -- am I coming across on the microphone -- now I hope I am -- just for the record what -- the process that we are going to follow, a motion has been made. We are going to have discussion on the motion. After the motion has been discussed to an extent that is applicable, the motion can either be amended or it can be withdrawn and another motion can be put forth.

DR. BOYAN: Okay. Dr. Grobler.

DR. GROBLER: If this application is denied, does that imply that the methodology that this process was started off as is wrong and it should be done all over again or what does that really mean. To me -- can you answer that?

DR. WITTEN: This is Dr. Witten and maybe I can clarify that.

Not approvable means the data as it stands are not enough for safety and effectiveness, but it doesn't mean that the process would necessarily start all over again.

Approvable means that the safety and effectiveness have been demonstrated. Approvable with conditions means the safety and effectiveness have been demonstrated but some additional information is requested and that information would have to be -- you would have to explain what that information was and why we need it for approvability.

DR. BOYAN: Okay. Mr. Dillard, do you want to add anything to that?

MR. DILLARD: No.

DR. BOYAN: Does everybody understand what our three motions actually represent?

Dr. Witten.

DR. WITTEN: No.

DR. BOYAN: No. Okay.

So, may I ask for a point of clarification? Is the level of conditions that we are placing now in this motion, are these within the limits of what you would consider to be an approvable recommendation?

DR. WITTEN: I think you would need to clarify what you were hoping to get out of the conditions. If you

think the safety and effectiveness have been demonstrated but there is some additional information that you feel is needed, then that is one situation.

If you feel that the safety and effectiveness has not been demonstrated and the additional information is what you would need in order to demonstrate and effectiveness, then that would be another situation entirely.

So, I think it is really up to you all as a panel to tell us what you think that this other information would give you or why you need it to decide which category this fits into.

DR. BOYAN: Okay. So, I would like to then go back to Dr. Janosky's comments because I think those are the appropriate comments at this time.

DR. JANOSKY: I know that there is a motion on the floor. Perhaps if we would start with deciding whether we are convinced of safety and effectiveness, at least for myself, I know that would be a very good starting point.

Chair, would you --

DR. BOYAN: Well --

DR. JANOSKY: Because the issue is first has it been -- are we convinced or are we satisfied that it is safe and effective. From that then follows our motion. Is that

not true? Because we have to establish safety and effectiveness for approval or approvable with conditions.

I would venture to say "no," I am not convinced of the safety or the effectiveness data based on the application, based on the data presented in the application to us.

DR. BOYAN: Right.

Mr. Dillard.

MR. DILLARD: Yes. This is Jim Dillard from the FDA. I would like to maybe help and see if I can't clarify this a little bit.

Jim Dillard from the Food and Drug Administration. I would like to try to help clarify a little bit if I can the fact that what you are talking about also is reasonable assurance. Remember the standard. There are a couple of things you need to remember when you are thinking about this and discussing it. One is valid scientific evidence. You need to take a look at the preclinical and the clinical information and you need to decide if you think it is valid scientific evidence.

If you believe it is valid scientific evidence, then I believe you need to take a look at, from the standpoint of safety and effectiveness, does it provide

reasonable assurance of safety? Are the risks minimized and are the benefits in the right direction in terms of a risk/benefit ratio? And do you believe that there is reasonable assurance that the device as it is labeled and as it is intended to be used, is there data to provide you the reasonable assurance that the device can be used effectively for that intended use?

I think those are the kind of concepts you need to keep in your mind, based on the data that you are looking at. So, I think in terms of an approvable recommendation or an approval recommendation, that data ought to say to you that, yes, we are looking at provides reasonable assurance of safety and effectiveness and the types of conditions then that you might recommend to us as the advisory committee are additional conditions that you believe from the standpoint of either preclinical or clinical input, we need to look at to ultimately move the application to an approved application.

If you don't believe that reasonable assurance of safety and effectiveness has been met, then that tends to make your decision more towards the not approvable situation. So, the break point you ought to really keep in your mind is that reasonable assurance of safety and

effectiveness.

I don't know if that helps, Dr. Janosky, at all.

DR. BOYAN: Dr. Wilkinson, you had a comment that you were ready to make.

DR. WILKINSON: I was just going to say that my understanding is that if we are not convinced now that there is at least a minimum -- enough data to have proved safety and effectiveness, that we shouldn't be approving this thing and that is what I think I just heard.

DR. KERRIGAN: I agree. I don't have reasonable assurance.

DR. BOYAN: Okay. We have an option to withdraw the motion or to amend the motion. We can do either one or the other. We can handle this several different ways. I think I am getting ready to have another message handed to me. This is another one -- nope, she is not all the way here yet.

DR. CHENG: Can I amend the motion as I did before?

DR. BOYAN: Yes -- wait. Let me -- I am getting instructions. One moment. I am sorry. Robert's Rules aren't my forte. I am learning as we go.

You would think that we would all know these

things. This is an unusual situation, as you can imagine.

Once the motion is made and seconded, we do have to vote on it. So, at this point, I would say that we need to get the best motion we can get and then vote our conscience and then if we can't agree, then we need to move to a new motion.

DR. GROBLER: Can I have a minute?

If we accept that the study is completed, that the methodology need not be changed and that the FDA, this panel or the FDA would reassess this at 24 months, will that be seen as a denial now and to be reconsidered or can that be an approval with these points in mind?

DR. BOYAN: Dr. Witten.

DR. WITTEN: Well, I would just you to look at that question in light of what we have already said, which is a product is approvable if reasonable assurance of safety and effectiveness has been demonstrated.

DR. GROBLER: At this time.

DR. WITTEN: Yes.

DR. BOYAN: Dr. Wilkinson.

DR. WILKINSON: What I heard this morning in the training session reiterated that a denial of approval at this point does not stop the device. The manufacturer can

certainly come back and reapply when they do have sufficient data to make their point.

DR. LAURENCIN: Call the question.

DR. BOYAN: Okay. Let's start the voting. We will start --

PARTICIPANT: Could you hold on one second?

DR. BOYAN: Yes.

I am going to restate the motion in its current form as I understand it.

The current motion on the table is that this application is approvable with conditions. The conditions are that the sponsor continue to follow the cohort to completion, that the question of vascular injuries be investigated in greater detail and that it continue to be monitored and the results reported back to the FDA, that the sponsor needs to report the two year follow-up data and that this be -- that FDA find the two year follow-up data be consistent enough with the data that we saw today to render approval.

Finally, that in the data set that they submit to the FDA that it must include a standard error of the measurement analysis.

That is the motion. I think that Dr. Yaszemski is

not correcting me. So, I think that I got it right.

And we will start the vote. The vote is either to recommend approval -- I mean to -- pardon me -- to vote "yes," "no" or you have the option of abstaining. We will start with Dr. Cheng and we will just go around the table.

DR. CHENG: I would vote "no."

DR. KERRIGAN: No.

DR. GROBLER: Yes.

DR. HALE: No.

DR. BOYAN: Dr. Yaszemski?

DR. YASZEMSKI: Yes.

DR. LAURENCIN: No.

DR. WILKINSON: No.

DR. BOYAN: Dr. Holeman doesn't vote and Dr. Silkaitis doesn't.

DR. WITTEN: We need to have everyone state their reasons.

DR. BOYAN: Yes. Right. I know.

Now we need to go around and have everyone state the reason for the vote that they made and we will start again with Dr. Cheng.

DR. CHENG: I believe I have probably already stated them. I have many of the same concerns as Dr.

Yaszemski. However, I will address that if we have a chance to redo the motion.

DR. BOYAN: Dr. Kerrigan.

DR. KERRIGAN: I have addressed all my concerns previously. In summary, though, I don't have reasonable assurance that the device is safe or effective.

DR. BOYAN: Dr. Grobler.

DR. GROBLER: I will state once again that if what is meant it in the proposal is the conditions outlined in the proposal is met at 24 months, I feel that -- I feel comfortable as to the safety and effectiveness of this implant.

DR. BOYAN: Dr. Hale.

DR. HALE: Although I don't have objections to the methods that were employed in this study, I don't feel like the data at this point demonstrate reasonable assurance of safety and effectiveness.

DR. BOYAN: Dr. Yaszemski.

DR. YASZEMSKI: I feel if the 24 month data were as the data today is with the more complete follow-up with respect to the percentage of patients followed, then it would demonstrate safety and effectiveness and that the safety valve was built into the proposal, that if those

conditions were not met, then the approval would not be granted.

DR. BOYAN: Dr. Janosky.

DR. JANOSKY: Not assurance of safety and effectiveness for the data as presented.

DR. BOYAN: Dr. Laurencin.

DR. LAURENCIN: For the reasons I have previously stated.

DR. BOYAN: Dr. Wilkinson.

MR. WILLIAMS: I also am not convinced that the data are sufficient to demonstrate safety and effectiveness. We have discussed a number of current problems in the study and we clearly have to have 24 hour data on the table before we can consider it. I can't guess at what pig is hidden in this sack.

DR. BOYAN: Thank you very much.

Now, for the next thing that we can do is we have -- the vote was 6 for not approving the motion and 2 for approving the motion. So, the motion did not carry. We can now entertain a new motion. Do I hear a new motion?

DR. CHENG: Yes.

DR. BOYAN: Dr. Cheng.

DR. CHENG: I would like to put forward a motion

that this not be approved at the present time for the following reasons, and that, perhaps, these questions be re-addressed. I would like to make a few comments, though.

A prospective randomized study is extremely difficult to perform well. As an investigator myself, I am familiar with those difficulties and I sympathize with the sponsor in terms of the difficulties in doing a study. I think they have made great inroads towards that end and I congratulate them on that aspect of it.

However, as the data stands at present, I don't think the data are enough to approve, even with conditions. So, therefore, the following issues I would like to be dealt with before making a motion to approve it at some future date.

So, my motion is not to approve it in its present form. And the issues to address are as we have stated before, the two year follow-up and knowing what is the success and failure at a time point. I would also like to see a little stricter definition of "fusion." I think the data is there and they have collected it. It does require reanalysis, but that is a matter of running your computer again. But I think a stricter definition of "solid fusion" is no motion at the fused site.

In addition to that, the statistical validity of the outcome measurement assessment, I think, should be addressed, as Dr. Janosky has stated earlier.

The vascular neurologic complication rate will be better assessed as Dr. Yaszemski had that concern earlier.

DR. BOYAN: Do I have a second for the motion to disapprove with the concerns as expressed by Dr. Cheng?

DR. WILKINSON: Second.

DR. BOYAN: Okay. I have a motion. I have a second. Any discussion of the current motion?

Yes, Dr. Wilkinson.

DR. WILKINSON: One thing. We need to be careful that a lack of motion -- the demonstration of a lack of motion, while a construct is still solid, does not prove fusion. You need more than just the motion films.

DR. CHENG: Why don't you amend that as to how you see fusion should be defined, which will be helpful for the sponsor?

DR. WILKINSON: Well, they are doing the science at this point. I would just say they have to prove that there is solid fusion and that it persists at 24 months.

DR. BOYAN: Okay.

Dr. Grobler.

DR. GROBLER: I guess it is fair to give some form of format of what we see is that what do we need to prove to obtain fusion as a general for any -- the present study on the table and even future studies, what do we -- where is our science now to prove fusion? Do we have any recommendations.

Dr. Wilkinson, do you have any thoughts on that? What should we do to get nearer to the answer?

DR. WILKINSON: Well, I think they need to show that there is neither motion nor gap in the bone -- demonstrable bone visible at 24 months. There can be absolutely no motion at 6 months and at 24 months, the disc space is settled. The device is lucent and now you can have tremendous motion or you can now, as the disc space collapses, now you can see gaps around the device that you couldn't see at 12 months.

I don't know that, but for me to be satisfied I would want to know those things. Bony continuity over a sufficiently large area, lack of motion and time stability.

DR. GROBLER: I am not sure what our role is, but I do think we should have a role to give standards on that. I mean, do we do that now or --

DR. BOYAN: What I think we do now is we vote and

then after we vote, we will go around the table one more time and get a chance to make a comment about the vote.

For this one then, why don't we begin the vote on the opposite side, with Dr. Wilkinson.

DR. WILKINSON: I vote "yes."

DR. BOYAN: Dr. Laurencin.

DR. LAURENCIN: I vote "yes."

DR. BOYAN: Dr. Janosky.

DR. JANOSKY: Yes.

DR. BOYAN: Dr. Yaszemski.

DR. YASZEMSKI: No.

DR. BOYAN: Dr. Hale.

DR. HALE: Yes.

DR. BOYAN: Dr. Grobler.

DR. GROBLER: No.

DR. BOYAN: Dr. Kerrigan.

DR. KERRIGAN: Yes.

DR. BOYAN: Dr. Cheng.

DR. CHENG: Yes.

DR. BOYAN: Okay. Now, Dr. Wilkinson, do you have some final comments that you would like to make as to why you voted for the motion and any additional comments that might be helpful to the sponsor that you would like to add?

DR. WITTEN: Excuse me. Yes, additional -- why he voted and also any additional comments you have --

DR. WILKINSON: I think my only last comment was that a tremendous amount of work has gone into this project already. I, too, have done scientific data collection and clinical studies. I am quite sympathetic with the amount of work that has gone into this study. This device may turn out to be very, very good. Prove it to us and come back.

DR. BOYAN: Dr. Laurencin.

DR. LAURENCIN: I want to reiterate that I believe a lot of good work has been done with a number of excellent scientists involved in this project and I do believe that the study population should be -- it should be carried out to the 24 month point. I think that right now we see only one-third of the study population at 12 months, which I believe by any standard is -- would be suspect.

I believe the fact that there are some signs of worsening of patients in the study from the 12 month point on is very, very concerning. It is very difficult to make an argument to have the -- to judge it in a 12 month point when you have got some patients and maybe the learning curve factor as one person from the company said, but still that is happening and it is certainly very concerning in terms of

that.

Also, in terms of the FDA, I would just say I would consider making a draft guidance document for these implants. Certainly if there was a draft document made at this point, which would clearly specify the two year time point, then, perhaps, they would have moved to a two year time point. Certainly we have uncovered some of the questions that may come in in terms of a draft document.

In particular, I would be interested in questions that have some simple answers. One, do you feel improved by this procedure or not? And maybe a "yes" or "no" answer. And, two, are you satisfied with this procedure? "Yes" or "no" answer. I think that would be a great part of a draft document and I think this may be just the time when a number of PMAs in this area coming out to get a draft guidance document together.

DR. BOYAN: Dr. Janosky.

DR. JANOSKY: I voted "yes" because I do not have reasonable assurance about safety and effectiveness, given the current PMA. And the suggestions that I made today, I think those incorporation would clearly strengthen the PMA if it were to return.

DR. BOYAN: Dr. Yaszemski.

DR. YASZEMSKI: My comments for strengthening it would be two. No. 1, to increase the percentage of patients followed up, as Dr. Laurencin has suggested. And No. 2, to show a validation of the imaging technique for showing fusion.

DR. BOYAN: Dr. Hale.

DR. HALE: I would basically agree with the comments from the other Panel members. I think there was just a lack of data at the 12 month time period, as well as the 24 month time period for a variety of reasons and I think the study, the application would be greatly improved by increasing those numbers.

DR. BOYAN: Dr. Grobler.

DR. GROBLER: I think basically the same, that the study should be completed as far as getting patients back. The x-rays, the CT scans should -- at the time of full completion of the study, hopefully, there will be more CT scans done and I think we need to in some way give some recommendations as to what we see as the gold standard for assessing a basic issue like fusion. This may not be the forum but I do feel that should be addressed as soon as possible.

DR. BOYAN: Dr. Kerrigan.

DR. KERRIGAN: Again, I don't have reasonable assurance that safety or effectiveness -- efficacy were demonstrated and my recommendations are that pain should be sort of the most important outcome method used for this study and that 24 months is necessary.

DR. BOYAN: And Dr. Cheng. I think you have already -- do you have something else that you would like to add that --

DR. CHENG: No.

DR. WILKINSON: Could I just raise one point of order, that under this format the chairperson also has to vote?

DR. BOYAN: No, she doesn't. Isn't that wonderful.

DR. WILKINSON: And give her reasons.

DR. BOYAN: No.

DR. WILKINSON: I have always had to vote.

DR. BOYAN: No. They made you vote. They don't let me vote. And I think that -- one thing I would like to say is that this was, obviously, a very difficult decision and I think the comments from the Panel reflect that and actually everything that I think has been reflected by things that people have said.

It is difficult, given the nature of the control that was perhaps the correct control at the time -- it is difficult to draw an assessment at 12 months, that you need to let the control have its chance or it is not an appropriate control. But I think a lot of really good things were shown and I am -- it is definitely a learning experience for all of us.

So, based on that, we need to make a decision and -- I would like to thank everybody that participated, the company -- this is very difficult for you -- all the members of the Panel, the FDA staff that did the review. I think that everyone put their best in and I think we are going to have a happy outcome at the end of all of this as well.

But we do have another PMA and we want the second PMA to have a chance on its own merits, not to be colored in any way by what has gone on. It needs to stand on its own. So, I am going to propose that we take a break for dinner, that we have -- you don't think so -- this is what I think we need to do because I think we need a rest. I think we need a 45 minute dinner break, not a minute longer than 45 minutes because we need to kind of clear our brains and then come back ready to work. 6 o'clock on the dot.

[Where upon, at 5:15 p.m., the meeting was

recessed, to reconvene at 6:00 p.m., the same day, Thursday, December 11, 1997.]

E V E N I N G S E S S I O N (6:06 p.m.)

DR. BOYAN: We are ready to begin the review of the premarket approval application submitted by Acromed, Inc., in the Brantigan Cage with VSP fixation.

I would like to remind the public observers at this meeting that, while this portion of the meeting is open to public observation, public attendees may not participate except at the specific request of the panel.

We are now ready to begin with the sponsor's presentation. I would like to ask, again, that each speaker state his/her name and affiliation to the firm before beginning the presentation.

I would also like to remind the panel as well as the sponsor that it is a new PMA and that it stands on its own merit.

The previous discussions of the day, with the exception of the general discussion of this morning, as if it never happened.

MR. CHRISTIANSON: Good evening. My name is Bill Christianson. We are here today to present the results of an IDE study for the Brantigan IF cage with the VSP spinal fixation system.

I am vice president of regulatory affairs for Acromed Corporation of Cleveland, Ohio, the sponsor of this

PMA application.

We studied the Brantigan IF cage with the VSP spinal fixation system. The IF cage is an interbody fusion device that is made up of polymer common fiber composite material that is strong and has a modiolus of elasticity close to cortical bone.

The cage is radiolucent so that bone graft can be assessed by planar radiographs. It is designed with an open architecture to maximize the volume available for packing with autologous bone graft, and to maximize the area of contact with the vertebral body end plate.

We studied the use of the IF cage with the VSP system, a stainless steel fixation device, that consists of plates and screws with optional transverse connectors and washers.

The VSP system is currently cleared for commercial distribution in the United States, labeled for severe spondylolisthesis.

The FDA approved this IDE study in November of 1991. The first patient was enrolled the next month. The last patient was enrolled into this IDE study in October of 1994.

The PMA was submitted to FDA in July of 1996. The FDA filed the application six weeks later and granted

expedited review status because the indications represented in this application are for serious patient conditions for which no device has received FDA approval in the United States.

The data being presented today are from a data base closure on April 18, 1997, and represent the reformatting of the clinical data as requested by FDA.

Acromed's IF cages have been commercially available in other countries for a number of years. IF cages have been approved in 15 countries in Europe, in Israel, Australia, Korea, Japan and, based on the same data being presented here today, in Canada earlier this year.

At the very least, over 16,000 disc phases have been treated worldwide with Acromed IF cages over this period of time.

Clinical reports from the rest of the world are uniformly positive, and complications and complaints are low in number and minor in severity, as will be documented later in our presentation.

We request that at the completion of the review of data presented today, the panel recommend to FDA that the Brantigan IF cage with the VSP spine system should be approved for marketing in the United States as well.

The Acromed presentation of preclinical and

clinical data about this spinal fusion device will be delivered by the individuals listed on these slides.

Now I would like to present Dr. Martin Persenaire to present the design rationale for the Brantigan IF cage with the VSP spine system.

DR. PERSENAIRE: Good evening. My name is Dr. Martin Persenaire, vice president of clinical affairs at Acromed.

I will present the design rationale for the IF cage with VSP system.

Vertebral lumbar interbody fusion was pioneered by Dr. Clowerd in the 1940s to treat painful intervertebral discs.

With his technique, he removed the entire nuclear portion of the disc and replaced it with multiple blocks of transplant bone.

Pliff has a number of advantages, such as restoration of disc height and lumbar alignment. A major disadvantage is the inconsistent quality of bone grafts.

For Pliff to be successful, the graft needs to address both a mechanical and a biological function. It must support physiological loads while the bone is being incorporated by creeping substitution.

The published success rates vary and, thus, pliff

with bone grafts, in spite of its theoretical advantages over most other types of fusion, never gained widespread acceptance.

Alternative fusion techniques continued to be developed. From 1986 onward, pedicle screws and plates with fusion were increasingly used.

In this technique, the disc itself is not removed. Pedicle screws and plates are intended to function as temporary replacements for deficient posterior elements.

They allow positional control of the spinal segments, and provide the immobilization necessary for optimal healing of the fusion grafts.

In some patients, clinical failures were caused by breakage of screws.

Engineers and surgeons came to realize that screws could never be made strong enough to adequately support the spine, if fusion failed to mature, or if the interior column did not share the loads.

To overcome the limitations of both the Clowerd and pedicle screw methods, the next logical step was to combine pedicle screw fixation with pliff, using interbody fusion devices that eliminated the limitations of the bone grafts.

This resulted in the development of the IF with

VSP system, which combines and enhances the advantages of both techniques.

The IF cage itself consists of a whole rectangular implant made of carbon fiber reinforced polymer. It has an open architecture that maximizes the volume available for autologous bone and large contact areas of the graft with the end plates. The cage has ridges to resist migration.

The dimensions of the cage have been chosen such that they, by themselves, can withstand all in vivo loads.

The material itself has two unique characteristics that greatly enhance the functional qualities of the cage.

First, in contrast to metals that are 10 times as stiff, the modiolus of elasticity is very close to that of bone, allowing compressive loads through the interior column to be passed on to the bone graft contained within the cages.

In other words, the cages do not stress shield the bone graft packed in the hollow area.

Once the graft has matured, the interior column support function of the IF cage should decrease as the interbody fusion bone increasingly assumes the support function.

Second, the material is radiolucent, allowing direct visualization of the graft and assessment of the

fusion status with routinely available planar X-rays.

Two cages are combined with VSP pedicle screws and plates which replace or supplement the posterior elements and position and immobilize the motion section during the time that it takes for the autologous graft to heal.

The IF cage with VSP system has been designed not only to treat simple diagnoses, such as single level degenerative disc disease, but also to address the needs of patients with much more complex biomechanical abnormalities and clinical conditions.

This includes patients with spondylolisthesis, patients with degenerative disc disease at multiple levels, and the most notoriously difficult group, patients who suffer from the results of previously failed fusion procedures.

In these patients, the restoration of the spine requires the combination of interior column support and posterior fixation for realignment and immobilization, whereby long-term, the interior column support will be taken over by the fusion bone.

The literature unanimously reports that the more surgeries the patient has undergone, the lower the clinical success rates.

This is especially true for patients with failed

fusion procedures.

It is, therefore, extremely important that once a fusion is indicated, the first attempt be successful.

The IF cage with VSP system creates the conditions that are most conducive to obtaining a successful fusion.

With that, I would like to introduce Dr. Hassan Sirhan, who will present the biomechanical test results with this system.

DR. SIRHAN: Good evening. My name is Hassan Sirhan. I am the research and testing manager at Acromed Corporation.

The IF cage was designed to meet the shortcomings of posterior lumbar interbody fusion, namely, to support the weight bearing and to resist the pull out strength or vertebral propulsion.

Based on the available published, we have set the following design criteria necessary for a safe performance for the IF cage.

First, the cage will sustain the maximum in vivo loading conditions. Under high static loading conditions, the cage strength should be comparable to the strength of the vertebral body.

The endurance limit or the infinite life load of the cage in a worst case scenario should exceed the average

in vivo loading conditions.

Furthermore, the pull out strength, or vertebral propulsion resistance of the cage should exceed the pull out strength of interbody bone graft of the same size and shape subjected to the same loading conditions.

White and Penjarbe(?) summarized several in vivo studies by Shose and Mattheson(?). They reported that the loads encountered during average daily activities are between 300 and 1200 newtons.

However, substantially high loads of up to 3,400 newtons can be encountered during heavy activities such as bending or lifting heavy objects.

Furthermore, the maximum loads that a vertebral body can sustain before a vertebral body or end plate fracture ranges between 5,000 and 8,200 newtons.

Static and fatigue testing were conducted using a 45-degree compressive sheer loading condition, which represents an L5,S1 tilt angle of 45 degrees, as shown on the left-hand slide.

As a matter of fact, the ASTM, as well as the ISO standard committees are considering adopting this procedure to characterize all interbody fusion devices.

These tables summarize in vivo loading conditions and the biomechanical test results of the smallest and the

largest IF cages.

The static test results showed that the maximum load that two small cages can sustain without fracture is 19,000-plus newtons, and 35,000-plus for two large cages.

These values are more than two times the reported maximum vertebral strength of 8,200, and by far exceed the average daily activity loading of 3,400 newtons reported in the literature.

In compressive sheer, the lowest test result was for the largest cages of 10,000-plus newtons. Therefore, in the worst case loading condition, the vertebral body is more likely to fracture before the device would.

As for the fatigue strength, the endurance limit over the load without failure after five million cycles for two large cages, subjected to worst case 45-degrees compressive sheer loading is 5,600 newtons, which is almost five times that required for a safe performance level of 1,200 newtons, and one-and-a-half times the maximum loads encountered during heavy activities such as bending while lifting heavy objects.

I would like to bring to your attention here that all the in vivo loads reported in the literature are the axial compression loads, whereas our fatigue strengths were established using a compressive sheer loading, which is a

more severe loading condition.

The pull out strength of the IF cage was evaluated biomechanically in a human cadaverous spine and compared to donor bone graft.

Tests were performed at the Cleveland Institute and published in Spine in 1991.

The pull-out strength or the vertebral propulsion resistance of 553 newtons for the cages is almost three times the pull-out strength of 126 newtons for interbody bond graft of the same size and shape as the cage.

Several investigators have documented the importance of posterior instrumentation on the clinical success of lumbar spine fusion.

The VSP system, or the fourth generation pedicle screws, was used with the IF cage in our clinical studies.

In 1995, Acromed launched the fifth generation pedicle screws. Testing was conducted by Brian Cunningham, of the orthopedic biomechanical lab at Union Memorial Hospital, comparing the strength, stiffness and the fatigue performance of the fifth generation VSP screws, to previous test results published in Spine in 1993.

The following bar charts illustrate the superior performance of the VSP systems when compared to other commercially available systems.

Looking at the fatigue performance, VSP system with the fifth generation pedicle screws has out-performed all other systems, including the VSP system with the fourth generation pedicle screws.

In conclusion, we believe that the IF cage alone has met and surpassed the mechanical requirements for interior column support, and the VSP system has met the static and the fatigue requirements for posterior spinal fixation.

Together, this interbody fusion system will correct and maintain the anatomic relationship of the spine, and provide optimum stability for fusion to occur.

At this time, I turn the podium over to Dr. John Brannigan, who will present the preclinical studies.

MR. BRANNIGAN: Good evening. My name is John Brannigan. I have a financial interest in both the IF cage project and in Acromed, the sponsor of this PFA.

I was principal investigator in this IDE study, and did my clinical work as associate professor of surgery and chief of spinal reconstructive surgery at Creighton University in Omaha, Nebraska. Currently, I am in private practice in San Antonio, Texas.

I will now present the preclinical testing of the cages and the polymer material.

In 1990, a battery of short-term and long-term biocompatability studies were undertaken on the polymer used to make the IF cage.

The tests performed during compliance with the ISO-10993 standards, as being adequate to demonstrate that a new implant material is suitable for human use.

The results of all the tests indicate that the polymer, or extracts thereof, is non-cytotoxic, non-mutagenic, not systemically toxic, non-pyrogenic, neither an irritant nor a sensitizer, and is compatible with blood. In short, the material is biocompatable.

A long-term test was conducted to evaluate if the material contained chemical leachates that might cause cancer in a rodent model.

Fibrobundles of the polymer, or ultra-high molecular weight polyethylene control material were implanted in the paravertebral muscles of rats.

The animals were sacrificed and analyzed at six months, one year and two years after implantation.

Multiple organ systems and the implantation site itself were histologically analyzed from each sacrifice point.

The results from animals implanted with the polymer were compared with results from animals implanted

with the polyethylene controls.

The conclusion of this test was that the response to the polymer was similar in profile to the expected rodent response to the polyethylene control.

The carbon cage device itself was tested in an experimental surgical model in the Spanish goat. Spanish goats had interbody lumbar fusion surgery in a random protocol.

Study goats were implanted with the carbon fiber reinforced polymer cage packed with autologous bone, and control goats were implanted with ethylene oxide sterilized allograft bone.

The objectives of this study were to determine fusion success of the cage compared with the control allograft, and to evaluate possible local or systemic toxicity, as well as the effect of the host tissue on the composite IF cage.

This study was conducted in compliance with the Public Health Service policy on the humane care and use of laboratory animals.

I performed the surgeries at the Turtle Creek Animal Hospital in San Antonio, Texas. This study was conducted in conjunction with the departments of orthopedic surgery and neuroradiology at Johns Hopkins University, and

the department of pathology at Union Memorial Hospital in Baltimore.

The results of this study were published in Spine in 1994.

Goats were sacrificed at six, 12 and 24 months. Full body autopsies were performed, and spines were studied by plane X-ray, three-dimensionally reformatted CT scans and by histology.

There were several differences in cage surgery in the goat as compared with the human IDE study. In the goat, one cage was used instead of two.

The single cage was inserted laterally in the coronal plane instead of posteriorly in the sagittal plane. This difference was required by anatomical differences between goats and humans.

Standard reamer broaches were used in both. No posterior fixation was used in the goats, effectively creating a worst case study of cages in the goat.

The right slide shows the allograft bone and cage as used in the goat.

The next slides show the appearance of allograft and cage as placed surgically.

Goats were kept indoors until wounds healed and then kept in an outdoor facility. X-rays showed the post-op

appearance of allograft on the left and cage on the right, in the disc space in goats.

At the time of implantation, the radiographic density at the cage level is lower because the cage is filled with cancellus bone.

At six and 12 months, five of five cage specimens were healed, whereas two of three allograft specimens were healed. At 24 months, all specimens were healed.

Three-dimensionally reformatted CT scans assigned specific colors to specific densities. In these slides, a light or bone color is assigned to bone density and the color red is assigned to cage density.

These slides show coronal and mid-coronal images of the six-month cage specimen, demonstrating bony fusion through and around the cage.

Three-D CT at 12 months demonstrated bone bridge in the interspace, indicating fusion. Sequential CT cuts in the axial plane demonstrated solid bone in the cage.

Sequential CT images in the sagittal and coronal planes demonstrated solid cage fusion.

In contrast, CT images of an allograft specimen that was not completely fused showed that the bone had been resorbed.

In coronal views of an allograft specimen, partial

healing has occurred by ossification of the annulus after bone resorption.

At 12 months, histology demonstrated living bone through the center of the cage. The red material is marrow content, which appears somewhat amorphous because of the thickness of the slide.

Close up of the cage strut shows fibrous materials surrounding the strut. There was some particulate carbon and polymer material in the local area.

The results of the gross and microscopic studies indicated that there was no local toxicity to the carbon or polymer material. There was no systemic toxicity.

There was localized carbon and polymer, but this material did not cause immune reaction, and did not cause osteolysis.

There was no migration of carbon or polymer particles to the local lymphatics or distant filter organs.

Particulate debris has been reported to occur around all total joint arthroplasty implants, and can cause clinical problems due to immune reaction with lymphocytic infiltration and secondary degradation of the device.

Larger volumes of particulate debris can cause osteolysis, loosening of the devices and clinical failure.

Distant migration of debris can be troublesome.

Because these events did not occur in this worst case animal study and have not been reported in any human cases to date, we do not believe that the amount of particles in these goat specimens represents a clinically significant concern.

In conclusion, interbody fusion with the carbon cage implants achieved fusion in the goat model, which occurred earlier than with allograft.

At the end of the study, plane radiographs, reformatted CT, and histology showed complete agreement in all animals, and confirmed that bone healing occurs through the cage.

There was no systemic or local toxicity to the implant material, no osteolysis and no migration of debris.

The results confirm that there was excellent biocompatability of the carbon polymer implants in this large animal in vivo study.

Now I would like to introduce Dr. Brad Hall, who will present the clinical study data.

DR. HALL: Good evening. I am Brad Hall, vice president of quality of the Bandess Health System, and an associate clinical professor of orthopedics at the University of Texas at San Antonio.

I was also a clinical investigator in this IDE study. I hold stock options in Acromed, a privately held

company.

I will be discussing the results of the IDE clinical trial to evaluate the safety and efficacy of the IF cage system.

There were 247 patients enrolled in this IDE study. Two hundred twenty-one were treated with the IF cage with the VSP instrumentation.

Twenty-six patients were treated with ETL donor bone and VSP instrumentation.

The distribution of patients among investigational centers is shown on the right.

The study was originally designed with four diagnostic categories: recurrent disc disease, spondylolisthesis, multiple level disease and multiple previous failed surgery or failed fusion.

The first two categories were randomized studies with ETL donor bone as the control group. The number was low in these categories because the information began to emerge early on that the ETL donor bone was suboptimal and patients refused to participate.

The latter two categories were more difficult patients, and no randomization was done in those groups.

In agreement with the FDA, and for analytical purposes, we combined patients with like diagnoses, and

created three diagnostic groups.

Our new diagnostic groups are degenerative disc disease, failed fusion and spondylolisthesis, all treated with the IF cage and the VSP system.

Since the number of patients treated with ETL and DSP is small, only patients treated with the IF cage and VSP will be presented here this evening.

Pool-ability of data for those new groups was confirmed across clinical sites, number of levels treated, number of previous surgeries by standard statistical methods.

There are 110 patients with DDD, 60 patients with a previous failed fusion, and 51 patients with spondylolisthesis. All were treated using the IF cage system.

Based on previous literature studies, approved controls, a 95 percent significance level, and at least 80 percent power, the relative risk of 1.5 can be detected for DDD and the failed fusion categories, and a relative risk of 2.0 can be detected in the spondylolisthesis category.

Therefore, there is sufficient sample size in each diagnostic category to provide adequate power for control comparisons to be made for the primary end points.

I will now describe the IF cage patients.

Combining all three diagnostic categories, there were 57 percent male and an average age of 44 years.

All the patients in the DDD and failed fusion groups had failed at least one previous surgery. Only failed fusion patients had a previous fusion attempt.

Patients in the spondylolisthesis group had not had previous surgery.

There were 47 percent receiving Worker's Compensation and/or litigation pending. Sixty-seven percent had at least one other medical condition and 60 percent of these patients had a positive smoking history.

The average duration of back and leg pain prior to the study were eight years and five years, respectively.

The surgeries were conducted from December 1991 through October of 1994. There were 356 levels treated in 221 patients.

Forty-eight percent had single-level surgery, 44 percent had two-level surgery, and 8 percent had either three or four levels performed.

The majority of the cases involved L4,5 or L5,S1.

Follow-up rates for one, three, six and twelve months were above 95 percent for all three diagnostic groups.

Follow-up rates at 24 months were 88 percent for

DDD, 85 percent for failed fusion, and 72 percent for spondylolisthesis.

There were seven deaths, six in the DDD group and one in the spondylolisthesis group. No death was related to the device.

At the time of the data base closure, 95 patients had 48 months follow up.

These prospectively collected data were compared to appropriately selected literature controls, an identical retrospective cohort of DDD and spondylolisthesis patients, and the clinical data published in the package inserts from two commercially-available devices recommended for approval by this panel in May of 1996.

A separate literature search was done for each diagnostic category. English language articles were identified by a Medline search using the search terms listed.

Only articles describing non-pedicle instrumented patients were considered for comparison.

These articles were reviewed for fusion and clinical success rates. Few articles permitted calculation of overall success as defined by FDA.

In general, articles tend to report on either safety or efficacy, but rarely on both.

Additional articles were used to evaluate complications and further surgical procedures.

Supplemental articles involving pedicle screw implementation were included for comparison of screw breakage rates.

The sponsor also collected retrospective data on 170 patients from eight spine centers in Europe and the United States.

These patients consisted of a consecutive series of DDD and spondylolisthesis patients, treated with identical methods and using the same selection criteria, patients treated with either IF cage or donor bone.

Both donor bone and IF cage arms will be used in comparisons for DDD and spondylolisthesis.

Prospective studies conducted at a similar time and under predefined protocols are considered a higher quality control than retrospective studies and literature controls.

The clinical data published in the package inserts for the two commercially available devices -- namely BAK and TFC -- were used for comparison to our DDD group because they are prospectively collected data, conducted concurrently and under similar IDE conditions.

The end points are clearly defined and are

comparable to ours.

The DDD groups of the three studies have similar characteristics. However, our DDD category is a more complex population due to the higher number of patients with previous surgeries.

The primary clinical end points for this study are fusion success, clinical success, overall success, complications and further surgeries.

Longitudinal results will be shown in graphs and all results presented, and comparison to controls will be made at the 24-month follow-up.

Fusion success is the radiographic assessment of the interbody fusion reported by the surgeon, according to seven descriptors, developed in mutual agreement with the FDA when the study was designed in 1990.

Fusion results were collected on a per level basis, and a descriptor of five, six or seven at 24 months is a fusion success.

We chose this definition of fusion success in consultation with clinicians, radiologists and the FDA.

When multiple levels in an individual patient were treated, all levels treated must be fused in order for the patient to be considered a fusion success.

Later, FDA will ask you if seven is the only

descriptor that should be included or interpreted as fusion.

From the beginning of this IDE study, the investigators were instructed that five, six and seven were descriptors of fusion, and would be interpreted as fusion at the 24-month end point.

FDA reviewed and accepted this definition when this study was designed in 1990.

We believe that limiting the interpretation of fusion to seven would be incorrect clinically, radiographically, and biologically.

Due to the radiolucent property of the IF cage, the fusion mass within the disc space can be clearly seen on planar X-ray.

At 24 months, bone ridging the fusion mass was no evidence of lucency or graft resorption or collapse of the construct, and indicate that fusion has occurred.

Therefore, at 24 months, patients with five, six or seven were fused.

As seen with the goat study, there is a high agreement between direct observation and radiographic assessment of fusion.

This has been confirmed in the IDE patients who underwent further surgery and re-expiration of their fusion areas.

The comparison of radiographic fusion success to surgical re-expiration yielded a positive predictive value of 94 percent, and an overall accuracy of 93 percent.

Following the clinical presentation, Dr. Petersilge will present to us the descriptors and talk about this in further detail. She is from University Hospitals in Cleveland.

At 24 months, the fusion rates for patients are 100 percent for DDD, 96 percent for failed fusion, and 100 percent for spondylolisthesis.

The graft shows fusion success rates over time for each category, along with 95 percent confidence intervals for each point.

The white line indicates the DDD group. The yellow line is the spondylolisthesis group, and the blue line is the failed fusion group.

Fusion rates have a sharp increase between one and three months and, again, between six and 12 months.

There was almost 95 percent fusion at 12 months in all three groups.

This is evidence that the IF cage with VSP is effective in achieving fusion.

Clinical success is defined by the FDA in the May 1996 panel meeting as a three-component scale consisting of

pain improvement, maintenance or improvement in function, and maintenance or improvement in muscle strength.

All three conditions must be met in order for a patient to be considered a clinical success. We will now consider each of these components individually.

Overall pain was collected using a five-point scale preoperatively and at each follow-up.

Success was defined as an increase in one or more points from the baseline assessment.

In addition to overall pain, separate pain evaluations were made for back, leg and donor site pain.

We felt the pain measurement most comparable and consistent with literature reports was overall pain, because this was most indicative of the patient's pain status as a whole.

At the 24-month interval, the pain success rates are 88 percent for DDD, 75 percent for failed fusion, and 85 percent for spondylolisthesis.

As can be seen, the percentage of patients considered successes using pain relief as their end point starts out at 70 percent at one month for the worst patients, failed fusion. The other groups were almost 80 percent at one month.

This indicates that the surgery itself provided

rapid improvement in pain status.

Function was collected using a five point scale preoperatively and at each follow up. Success was defined as no decrease in function score between post-operative evaluation and the baseline assessment.

At the 24-month interval, the percentage of patients maintaining or improving in functional status were 97 percent for DDD, 94 percent for failed fusion, and 97 percent for spondylolisthesis.

All three groups climbed significantly, to above 90 percent at three months, and remained in the high 90 range throughout the post-operative period.

Muscle strength was collected for 10 muscle groups bilaterally on a 10-point scale and are predicted here. This was collected preoperatively and at each follow-up.

Success was defined as no decrease in strength in any of the tested muscle groups.

At 24 month interval, the percentage of patients maintaining or improving muscle strength was greater than 95 percent in all three groups.

Almost 90 percent in each diagnostic category maintained or improved muscle strength immediately following the surgery.

At the 24-month interval, the percentage of

patients meeting all three conditions, and considered a clinical success, were 84 percent for DDD, 71 percent for failed fusion, and 79 percent for spondylolisthesis.

The graph shows that as early as one month, over 30 percent of the patients treated with the IF cage for any of the three diagnostic indications achieved clinical success, meaning that they had met all three criteria.

This percentage increased significantly between one and three months, and even at 24 months, the most difficult group -- failed fusion -- had a 71 percent clinical success.

Overall success was defined as fusion success and clinical success. These criteria which were used upon the FDA's request, made it possible to compare the results against other interbody fusion devices recently approved.

DR. WITTEN: Excuse me. Before you present that, I would like to say that data from a package insert or a summary of safety and effectiveness on another PMA product isn't available for use for comparison. It can't be considered by the panel in their approval process or deliberations regarding your PMA and can't be presented here.

So, the other comparisons that we would look at would be the comparisons to the literature.

DR. HALL: Can we get a ruling on that? Is that an absolute true fact?

DR. BOYAN: That is the ruling, yes.

DR. HALL: Something published and publicly available that has been reviewed by the panel, I would --

DR. BRANNIGAN: Dr. Hall, when we get to those slides, then, where we compared our data to the package inserts, we will just skip right over them.

DR. BOYAN: That would be the best thing.

DR. HALL: That is what I will do, then. You may have to remind me, though. I am reading and I might kind of forget here.

DR. BOYAN: I will remind you.

DR. HALL: Thanks. I appreciate that. Okay, let's see here. Overall success. At the 24-month interval, the percentage of patients meeting all four criteria and considered an overall success were 85 percent for DDD, 67 percent for failed fusion and 79 percent for spondylolisthesis.

The graphic shows that overall success increased over time from one month through 12 months post surgery and leveled off at 24 months.

By 12 months, all three groups were close to 80 percent overall success.

All complications were reported, regardless of the severity or the duration. This resulted in a comprehensive and exhaustive list, unlike that ordinarily found in the literature.

The literature ranges will be presented, but should be interpreted cautiously, in light of the thoroughness of the reporting in the IDE study.

Complications were divided into device and non-device related.

Device-related complications were subdivided into IF cage of VSP. Non-device-related complications were subdivided into major, minor and insignificant.

The major complications were defined as an untoward event which resulted in further surgery, or results in a permanent impairment of the patient's clinical condition. An example would be a permanent foot drop.

Minor complications were defined as an untoward event which required additional treatment, or may prolong the duration of the hospitalization or complicate treatment, which results in only a temporary impairment of the patient's condition.

Dural tears successfully repaired during the index surgery would be an example of a minor complication.

Insignificant complications were defined as events

which may require some additional treatments, but it has only minor effects on the patient's condition.

An uncomplicated urinary tract infection is an example of an insignificant complication.

Complications will be presented in total for all IDE patients treated with the IF cage system.

All device related complications were minor. There were three complications involving the IF cage. Two cages broke and one cage was displaced during insertion without clinical sequelae.

There were no cage complications reported post-surgically.

For VSP, 6.3 percent of the patients experienced a broken screw. Only 18 screws out of 1,270 screws implanted with the IF cage patients broke, for a screw breakage rate of 1.4 percent.

2.7 percent of the patients had screw loosening and one patient had screw pull-out. These are comparable with the literature results.

Major non-device-related complications included 2.7 percent dural tears requiring further surgery for repair, and 1.4 percent pedicle fracture.

There was one person each with deep wound infection, myocardial infarction and foot drop. These rates

are also comparable to the literature.

Minor non-device complications included 16.7 dural tears repaired during the index surgery, and 3.2 percent wound infection, and 1.4 percent RSD.

Other non-device minor and insignificant complications were as listed.

Further surgery was subdivided into reoperation, revision and removal as defined by the FDA.

8.1 percent of the 221 cage patients experienced a reoperation. Reoperations included wound abrievement, retrieval of broken drains and secondary repair of dural tears.

4.1 percent of the patients had a revision which included extending the VSP fixation for stabilization, pseudoarthrosis repair, or treatment of an additional level not in the study.

Two revisions involved the cage. The first patient had a pseudoarthrosis at a level treated with a single cage.

The second patient had a wedge osteotomy at a solidly fused level. The osteotomy was performed to restore lordosis during which a portion of the cage was removed.

In this study, no IF cages were removed, only VSP components.

At the time the study was devised, it was considered acceptable not to remove internal fixation implants in all cases where the bone was healed.

Only 36 percent of the patients had removal of the VSP components. Many VSP removals were elective removals, although we did not collect the actual reason for the removal.

Removal does preclude the development of any potential side effects which can be caused by the presence of stainless steel implants over long periods of time.

Such removals as conservative measures are very similar to the removal of bone plates or IM rods after a fracture has healed.

Removal also follows the recommendation in the labeling of the investigational VSP used in this study, as well as the removal language required by the FDA in pedicle screw systems available today under a 510(k).

There are two secondary clinical end points to consider in this study, the maintenance of the disc height and patient satisfaction.

The restoration and maintenance of disc height is important to open the neural foramina and to decompress the nerve roots.

At the 24-month interval, the average increase in

disc height over baseline was at least four millimeters. Each box plot represents the distribution of disc heights for a particular diagnostic category at a specific time point.

The box itself represents the middle 50 percent of the data.

All three diagnostic groups showed a marked increase immediately after the surgery, which was maintained throughout the entire post-operative period.

At each follow-up visit the patient was asked about their level of satisfaction with the results of the surgery.

There were six statements that the patients could choose from. Only patients who responded to number four -- my surgery gave me some improvement and I would do it again -- or number five -- I am greatly improved and I am completely satisfied with the results of my surgery -- were considered to be satisfied patients.

At the 24-month visit, 84 percent of the DDD patients, and 67 percent of the failed fusion patients and 86 percent of the spondylolisthesis patients were satisfied with the results of their surgery.

The satisfaction percentages, with confidence intervals over time, are presented in the graph.

Now we consider how the results just presented compare to control studies, what time trends can be identified, and determine which covariates may be significant.

Comparisons to control studies were done for each of the primary end points using the chi square analysis.

Generalized estimating equation, GEE, methodology identified potential covariates for each of the primary end points.

Covariate analysis and longitudinal analyses were presented to the FDA.

For the DDD category, fusion success was compared to eight articles, two retrospective arms, and something else. (Laughter.) I am sorry, I don't know how else to handle that.

The results are shown on the left and the corresponding control studies in matching colors are on the right.

Studies shown in the white favor the IF cage IDE study. Those in the green show no difference, and those in the yellow favor the control study.

For fusion, the IF cage system was statistically better than 11 of 12 controls. One study, the retrospective cohort, using IF cage, was not different.

In 12 controls, 12 articles, in both arms of the retrospective study, the IF cage system was statistically better than six studies, trended to be better than one study, and was no different in six studies.

One study in 14 trended to be better than the IF cage system.

For overall success, there were five control studies. One article, both arms of the retrospective cohort, the IF cage system was better than four of the five studies, and no different than the retrospective IF cage arm.

For degenerative disc disease, the 24-month fusion rate was 100 percent. The clinical success rate was 84 percent, and the overall success rate was 85 percent.

After the study comparisons, we conclude that the IF cage system is safe and effective in the treatment of degenerative disc disease.

For the failed fusion category, there were only three articles which discussed fusion success of previously failed fusion patients.

Additionally, only four articles discussed clinical success and only one article discussed overall success.

Our prospective series of 60 patients with a previously failed fusion is almost unique with regard to the

published literature.

Fusion success was compared to three articles. For fusion, the IF cage system was better than all three studies.

Of four articles, the IF cage system was better than one study, and no different from the other remaining, for clinical success.

For overall success, there was only one article. The IF cage trended to be better than this article.

For failed fusion, the 24-month fusion rate was 96 percent. The clinical success rate was 71 percent, and the overall success rate was 67 percent.

After these study comparisons, and considering the difficult nature of this group, we conclude that the IF cage system is safe and effective in the treatment of failed fusion.

For the spondylolisthesis category, fusion success was compared to 11 articles in both treatment arms of the retrospective cohort.

For fusion, the IF cage system was better than 10 studies and no different than the remaining three studies.

For clinical success, there were 10 articles in both treatment arms of the retrospective cohort. The IF cage system was better than one study, trended to be better

than one study, and was no different from five studies.

The remaining five studies have better clinical success rates than the IF cage study.

For overall success, there were four articles in the retrospective cohort. The IF cage system was better than two studies, trended to be better than two, and was no different than two studies.

For spondylolisthesis, the 24-month transfusion rate was 100 percent. The clinical success rate and overall success rates were 79 percent.

After these study comparisons, we conclude that the IF cage system is safe and effective in the treatment of spondylolisthesis.

Against the backdrop of over 16,000 levels treated outside the United States since 1991, Acromed has received only 11 complaint reports regarding the IF cage.

Six of these were for breakage of the cage during insertion, all of which were removed without sequelae and new cages were inserted.

There was one case that had a disc space infection, pseudoarthrosis and a broken cage.

At this time there are four open reports that are pending further investigation. The details of these events are currently unknown.

Next, clinical reports in the use of IF cage outside the United States will be discussed.

To seek approval for the IF cage from the Japanese Ministry of Health and Welfare, a clinical study was conducted.

This study had 66 patients at two sites. 100 percent of the patients rated their outcome as very good or good.

The clinicians who conducted this study rated the IF cage as 100 percent safe because there were no device-related complications.

The one non-device related complication is a type C hepatitis infection attributed to a blood transfusion.

The 100 percent useful rating combines the fusion rate with the Japanese Orthopedic Association low back score.

In 1997, Yamamoto reported on a series of 63 IF cage patients followed for a minimum of three years with an average follow up of over four years.

Excellent and good results were reported in 93 percent, fusion rate was 100 percent. No breakage of implants, migration or radionuclides were seen.

The Japanese Health Ministry also requires post-market surveillance, to look at potential complications from

the use of the device in Japan.

The spring of 1997 report had two complications out of 152 cases treated. One complication was low back pain due to a lumbar compression fracture secondary to osteopenia.

The second complication was a foot drop resulting from a L5 nerve decompression. Neither complication was attributed to the IF cage.

Tyco Tolborg of Stockholm reported on the first 51 consecutive cages of IF cage implementation in six Swedish hospitals.

DR. BOYAN: Dr. Hall, is this information that was presented in the PMA?

DR. CHRISTIANSON: Yes, absolutely.

DR. HALL: Yes, we gave you all that information.

DR. BOYAN: Okay, excuse me.

DR. HALL: No problem. Thank you. Those patients were maintained post-operatively in all but one patient, who had osteoarthrosis.

Minor complications were three dural tears and one post-operative infection, all unrelated to the cage. There were no broken screws in this series of patients.

In 1997, Sabato reported on 40 cases treated in Israel with IF cages without internal fixation. Follow up

was from 1.5 to six years.

Good or excellent results were obtained in 78 percent. A large portion of the patients showed an improvement in pain, walking distance and patient satisfaction. The fusion rate was 100 percent.

Complications included three superficial wound infections, one patient with a one degree increase in spondylolisthesis, and three temporary nerve deficits.

The retrospective cohort described previously and used in comparisons for the DDD and spondylolisthesis categories was a series of patients from Europe and the United States which matched the IDE eligibility criteria, in total, 56 patients treated for DDD and 114 patients for spondylolisthesis.

All patients had internal fixation. There was a minimum of 12 months of follow-up.

For DDD, the group of patients treated with the IF cage had a statistically higher fusion and overall success rate than those treated with donor bone.

There was one dural leak and one dislodged pliff graft in the donor bone patients.

In the cage patients, there was one case of nerve root damage and one wound infection.

For spondylolisthesis, the group of patients

treated with IF cage had a statistically higher fusion, pain and overall success rate than those treated with donor bone.

Major complications for the donor bone group included three broken screws, one fractured pedicle and one dural leak.

In the cage patients, there were two broken screws and two pedicle fractures.

In conclusion, our IDE data, supported by other worldwide experience has shown that the IF cage with VSP restores and maintains disc height, has high patient satisfaction, has a reasonable complication and further surgery rates, has high fusion and clinical success rate, as well as high overall success rate for degenerative disc disease, failed fusion and spondylolisthesis.

We believe these data constitute valid scientific evidence that provides a reasonable assurance of safety and effectiveness when this device system is used to treat these three conditions.

I would now like to introduce Dr. Petersilge, who will present a radiographic assessment of fusion.

DR. BOYAN: I just wanted to share with you all that we have got 10 minutes to go.

DR. CHRISTIANSON: We will be there.

DR. BOYAN: You will be there?

DR. CHRISTIANSON: Yes.

DR. BOYAN: Okay.

DR. PETERSILGE: Good evening. My name is Cheryl Petersilge. I am head of musculoskeletal and emergency radiology at University Hospitals of Cleveland, and I am an assistant professor of radiology and orthopedic surgery at Case Western Reserve University.

I have no financial interest in Acromed Corporation nor in the IF cage. My purpose today is to discuss the radiologic appearance of fusion.

To date, there are no standard, widely accepted radiographic criteria for determination of fusion and pseudoarthrosis, in part due to the lack of any radiologic, histologic cohort of study in humans.

Many of the clinical studies of fusion rates fail to define their radiographic criteria. Instead, these studies state that radiographs were assessed for the presence or absence of fusion.

The accepted gold standard for fusion is surgical exploration, which is not easily incorporated into a clinical protocol, such as this study.

There are two general approaches to radiographic assessment of fusion. One is to evaluate structural or anatomic fusion. This evaluation is typically performed

with radiographs, CT or tomography, and implies a solid osseus union.

Functional integrity of the graft may be determined using flexion and tension radiographs. This assessment does not distinguish between solid osseus and solid fibrous union.

In this study, the functional assessment of fusion integrity cannot be determined due to the presence of posterior pedicle screw instrumentation.

With no standard radiographic criteria for fusion assessment, it is reasonable to correlate the process of graft incorporation with the radiographic appearance of other known osseus processes such as fracture healing.

The initial processes of autogenous cancellus graft incorporation are vascular in growth, and the invasion of primitive mesenchymal cells, which will continue to produce new bone.

These processes will not be expected to produce any visible change on the immediate post-operative films, nor on the subsequent three-month films.

The second phase of graft incorporation is the deposition of unmineralized osteoid along the existing trabeculae of the graft.

The dead trabeculae are subsequently resorbed, and

finally, the unmineralized osteoid is mineralized.

Radiographically, these changes will be accompanied by a decrease in density within the grafted area, when compared to immediate post-operative films.

This relative lucency is a result of the absorption of trabeculae. The grafted area may have a fuzzy or ill-defined appearance as unmineralized osteoid is produced.

In the later stages, the graft will again increase in density, as osteoid is mineralized. The grafted area will continue to increase in density as more osteoid is formed and mineralized.

As the mineralized osteoid matures in response to mechanical stresses, a mature pattern of trabecular bone forms.

Radiographically, this maturation is accompanied by the appearance of trabecular bone within the grafted area.

I would now like to review the radiographic appearance of fusion incorporation. On these AP views of an L4,L5 fusion, we can see the relative density within the fusion mass at three months in this film on the left.

Here we can see the carbon cages, which are radiolucent, and which provide a standard density reference

from film to film.

On this film, obtained at six months, we can see that the material within the cage is increasing in density, when we compare it to the constant density of the cage.

This is the six month film, again, reproduced. Now we have a 12-month film and we can see that there is continuing increased density within the graft, making the cage more visible.

In addition, this sharp border with the adjacent vertebral body has been obliterated, indicating bridging bone in the fusion site.

These lateral radiographs at an L5,S1 fusion also demonstrate the changes we would expect. Here is the immediate post-operative film and we can see the graft material which is speckled throughout the cage.

On the three-month film there has been decreased density within this grafted area. At 12 months, we can see that there is increasing density within the graft area. There is formation of bone, and the fusion area is actually extending posteriorly within the disc space.

Further maturation is identified at 24 months, and trabecular bone formation is visible, as well as continued increasing density within the sites.

In summary, the radiographic descriptors which we

can apply to fusion success include bone bridging the fusion area, increased density of the fusion bone, and continuous trabecular bone bridging the fusion site.

Any one of these three descriptors is an indicator of successful radiographic fusion.

I would now like to reintroduce Dr. John Brannigan.

DR. BRANNIGAN: I will now present a typical case from our study. Patient JJ was 46-year-old male who presented on June 15, 1992 with a five-year history of back and right leg pain.

He had been injured at work while carrying a heavy load of beef on his shoulder. After extensive conservative treatment he had lumbar discectomy surgery in 1988. He received no benefit from this surgery.

When evaluated in 1992 for the cage study, he reported his pain level as one, excruciating or unbearable, and his function was two, limitation to non-strenuous activities.

He was unable to participate in any activities outside the home. Any activity caused increased pain. JJ had not worked in five years and was receiving disability benefits.

A lateral X-ray taken on 6-15-92 showed a mild

decrease in disc space height at L4,5. An MRI scan taken on 6-29-92 showed extensive degenerative changes at L4,5 and L5,S1.

JJ had surgery on 7-22-92, including cage pliff with VSP at L4,5 and L5,S1. Two weeks after surgery, JJ was found to have a superficial wound infection that responded to local treatment.

At three months follow up, on October 29, 1992, JJ reported that the pain level in his back was mild. He had no leg pain. Functional status was improving and he was able to participate in moderate activities outside the home. He expressed complete satisfaction with his surgical treatment.

X-rays showed bone density in the interbody fusion area and normal position of the pedicle screws and plates.

I would like to direct your attention to the bone density in the L4,5 pliff fusion area, and the changes in this bone density over time.

At the three-month interval, the bone graft inside the cages is visible and has a somewhat greater density than the four struts of each cage, which are also visible.

At six month follow up on 2-4-93, JJ reported that his pain was mild, function was increasing. He routinely walked two to three miles a day. X-rays showed increased

bone density in the pliff fusion areas.

On the left side, the cage struts are more clearly visible than before, indicating that the fusion bone density has increased.

The bone density inside the cage on the right is less, indicating some degree of bone resorption. The bone density at L5,S1 is increased, and the struts are visible in the lateral view.

At a one-year follow up on 8-19-93, JJ reported mild pain. He was able to do all activities except heavy lifting, and he returned to work in a light duty capacity at a food production company.

X-rays at one year showed that the L5,S1 pliff is radiologically solid on the lateral view. At L4,5, the cage end plates are visible as areas of decreased density.

The bone has partially resorbed from inside the right side of the cage. The left side of the cage demonstrates solid bone with increasing density, indicating solid fusion.

At two years follow up, on 8-29-94, JJ reported virtually no pain. He had returned to full functional capacity without any limitations.

He had been working for the past year doing heavy maintenance in the food production company. X-rays

increased bone density in all fusion areas.

At there years post op, on 12-12-96, JJ reported that he had no pain, and normal function. He was taking no medication and he had continued working at a heavy work capacity for the food production company. He expressed complete satisfaction with the surgical result.

X-rays showed solid fusion at all cage levels with marked increase in bone density throughout the fusion areas.

The area of previous bone resorption on the left, L4,5 cage, had consolidated completely.

I would now like to reintroduce Bill Christianson, who will summarize this presentation.

MR. CHRISTIANSON: In conclusion, Acromed believes that the data presented today constitute valid, scientific evidence that indicates that there is a reasonable assurance of safety and effectiveness of the Brantigan IF cage with the VSP spine system, as used to treat patients with degenerative disc disease, previous failed fusion and spondylolisthesis.

We have demonstrated that the results of our clinical investigation are comparable to or better than the literature, even in patient populations with complex and difficult-to-treat diagnoses.

These devices are approved for use in multiple

countries around the world. Thousands of patients have been treated with this device, with documented positive clinical outcomes, and negligible rates of untoward effects.

We ask the Orthopedic and Rehabilitation Devices Advisory Panel to recommend to the Food and Drug Administration that these devices should be approved for patient care in the United States as well.

DR. BOYAN: Thank you very much. We are going to go right to the FDA presentation. The lead reviewer, Janine Morris, will begin.

MS. MORRIS: If anyone is interested at this time, we have a construct here to look at this system. I will pass it along.

Good evening. We realize it has been a very long day, but we are on the home stretch. So, we hope you will bear with us. We will attempt to make our presentation as short as possible.

The review team for this PMA includes myself, Janine Morris, and I will briefly present the device description and mechanical testing of the subject device.

Following my presentation, Dr. David Berkowitz will briefly touch on the preclinical tests performed to support the biocompatability of the cage material.

Dr. Jerilynn Glass will then present a clinical

summary and Dr. Ponnapelli will conclude with his presentation of statistical analyses.

Considering that the sponsor has already spoken in detail on the device under review and the mechanical testing performed, I intend to just clarify what device is the subject of this PMA, and what mechanical testing has been provided in support of its approval.

The device that is the subject of this PMA is a dual system, consisting of the Brantigan IF cage system and the VSP spinal system.

It is important to keep in mind that these two systems were studied together as a combined system. Therefore, it is the combined system that is being considered for approval. However, for simplicity, I will describe each system separately.

The Brantigan cage system for which the sponsor is seeking approval is fabricated from a composite material, consisting of 70 percent ultrapec, the polymer matrix, and 30 percent carbon fiber.

For manufacturing of this cage, the sponsor is using an injection molded process which I have described in earlier memos as the chopped fiber process or the chopped fiber material.

Each cage is rectangular in shape and supplied in

13 sizes. It should be noted that the wedge shaped cages that you may have identified during your review are not the subject of this PMA.

However, after panel discussion and recommendation, we will ask a general question with respect to the wedge-shaped design.

Each rectangular cage has a small tantalum bead, which is used as a radiopaque marker.

The VSP spinal system consists of pedicle screws, plates, washers and a transverse connector, all of which are fabricated from stainless steel in accordance with ASTM.

Mechanical testing performed on this device were conducted on each system separately. It is assumed that the data from separate testing would adequately demonstrate the mechanical performance of the systems combined.

The sponsor has conducted numerous mechanical tests as design changes were made over the course of the study.

I will only highlight the final tests that were supplied in support of the final device design, for which we are seeking approval.

The mechanical testing was performed on the Brantigan cage which I just described. It consisted of static and fatigue testing.

The sponsor established that the largest sized cage was the worst case for this cage design under static compressive sheer testing.

The static compressive sheer strength was determined to be 5,000 newtons.

Based on this data, the sponsor proceeded with a dynamic test consisting of compressive sheer, and the largest case size was determined under a load of 2,800 newtons, that the cage underwent five million cycles without fracture.

These results differ from what the sponsor presented, by a factor of two, since the sponsor's results are based on the use of two cages per level.

The test construct was designed for a single cage. Therefore, this is the data that I have presented.

These results were compared to earlier tests performed, and these tests were found to demonstrate a greater fatigue resistance over earlier designs.

The mechanical testing that was provided in support of the VSP spinal system consisted of static and fatigue bending of the system construct.

The results showed that the static limit was approximately 1,200 newtons and the fatigue load at 600 newtons ran out to one million cycles.

Typically, to meet consensus standards, such as ASTM, these tests are usually recommended to be performed out to five million cycles. That is the normal number of cycles that would be expected to be seen in two years.

Other than a few inconsistencies in the tests in general, the bench data appears to support the mechanical performance for this combined system.

I would like to now introduce Dr. Berkowitz, who will discuss the remaining preclinical testing.

DR. BERKOWITZ: I am going to discuss the toxicity testing, biocompatability testing. I have a few remarks on the goat experiments and then I will discuss the carcinogenicity testing.

Cytotoxicity was tested on mouse fibroblast lines using both the agar overlay and the MEM elution method. There was absolutely no indication of cytotoxicity.

Systemic toxicity was tested by extracting the device into four different solvents of varying polarity, and then injecting this, depending on the solvent, either interperitoneally or intravenously. Again, there were no examples of toxicity.

Finally, there was a test for leachates that was impressive. They used a standard gas chromatography mass spec system, which could detect 66 compounds, most at the 10

parts per billion level. No leachates were detected in the system.

This is an important consideration, both in terms of the toxicity and the carcinogenicity, that other compounds aren't leaching out of the device.

For biocompatability, the pyrogenicity was well below acceptable levels. Hemolysis was tested by direct contact of the device with blood, and there was no hemolysis.

Again, for the sensitization assay, there was no indication of sensitization at all.

The long-term implantation study in the goat was a little bit difficult to interpret. One is, there was a somewhat more rapid fusion with the use of the cage, but the fact that they used an allograft in the control, but ollografts in the controls may confound that. It is difficult to know how to interpret.

Also, there were carbon particles found around the struts in the cage, and a little bit of bone resorption as well.

It is also difficult to know what exactly the causes of these were. Some of the carbon particles may have come because of the cleaning system, which is actually an ultrasonic step that has been added to the cleaning of the

device, so that may eliminate some of it.

There was also a guess that both the carbon particles and the bone resorption may have been caused by the fact that the VSP spinal system wasn't used, which would be used in humans, and it made the motion itself, which caused some of the carbon particles and the resorption.

Carcinogenicity testing was tested in bacterial and mammalian cell assays and there was no mutagenicity. Likewise, for unscheduled DNA synthesis and cell transformation assay, also no indication of unscheduled DNA or cell transformation.

There was also a two-year carcinogenicity study done and rats were implanted. One hundred rats were implanted with a polyethylene control and another 100 implanted with the pekeet component of the device.

The observations were made at six, 12 and 24 months. There was hematology, blood chemistry, necropsy observation, body and organ weights and histopathological findings.

Except for the tumors, there were no substantive differences. The results were, at the final testing, that there were 22 tumors at the implantation site in the control animals and 20 in the control animals.

All of these tumors were of mesenchymal origin.

So, there is really no reason at all to consider these anything other than just foreign body carcinogens, which is seen in rodents, and it has a very unclear relationship to what is seen in humans.

If there are no questions, I will just ask Dr. Jerilynn Glass to continue with the clinical review.

DR. GLASS: In my clinical review, I will focus only on the main features of the prospective U.S. data that were submitted in support of this PMA.

The sponsor has reviewed the original study arms, but a few points bear repeating.

The first arm was recurrent disc pathology at one level. The second arm was spondylolisthesis patients at one level.

Both recurrent disc and spondylolisthesis patients that had multiple level disease were included in the third arm.

Finally, the fourth arm, multiple failed surgery or failed fusion, was retained as a distinct category, because the first and third arms permitted only one previous surgery and the second arm permitted no previous surgery.

Because of insufficient sample size in the various arms, the sponsor was given the option to pool data, if it could be shown that no bias would be introduced and it could

be justified statistically.

The statistical review to follow will discuss data pool-ability.

The sponsor did reorganize the data, essentially combining multi-level with single level patients. The new groups are presented at the bottom of this slide.

The A-3 group is multi-level DDD with one previous surgery. A-6 is multi-level DDD with one or more previous surgeries.

It is important to note here that the A-3 group is a subset of the A-6 group. This will come up repeatedly as we go through the results.

The B-3 group is multi-level spondylolisthesis with no previous surgeries. The final group is C-4, failed fusion. Note that this group had been part of an arm previously and is now retained as a separate group.

This basically summarizes the original study arms. Please note that the first two arms involved randomized subjects, who received either the device or ETO donor bone.

The third and fourth arms involved a comparison of the device with literature controls.

In the reorganization of the data, a mixing occurred within the A-3, A-6 and B-3 groups, such that each group now had both originally randomized subjects and

originally non-randomized subjects. Again, this is going to become an issue that will be discussed in the statistical review.

On the same slide, look down at the B-3 category, spondylolisthesis. Please note that it was made up of both lytic and degenerative spondylolisthesis patients.

In fact, there were approximately two times as many degenerative spondy patients as lytic spondy patients.

Also, note that the numbers in this particular group are less than found in the other groups and the issue of the smaller size here will be brought up to you a little bit later as part of a panel question.

Also, note that with the revised groups, all the comparisons now were with the literature controls.

As already mentioned, there were 221 subjects who received the device at eight sites. Most of the inclusion criteria are given here for you to peruse.

Basically, all patients had to have back or leg pain and most, in fact, had both. They could have up to three levels of spinal revolvement.

As it turned out, there was as sprinkling of patients who had up to four levels of involvement.

Exclusion were infection, tumor, significant osteoporosis, metabolic disease and pregnancy. One of the

panel questions will be regarding the labeling and whether severe osteoporosis should be considered a contraindication for this device.

Finally, you have the radiographic criteria for each of the spondylolisthesis and DDD subjects for entry into the study.

The nature of the surgical procedure has already been described. One of the panel questions will deal with whether surgical training should be required prior to use of the device.

If you look at the middle of the slide, you will note that all sites were encouraged to follow a given set of procedures as part of a patient post-surgical rehabilitation program. So, there was an effort to have consistency there.

The evaluations were at pre-op, one, three, six, 12 and 24 months, and the end points are listed below, and they have already been discussed.

Please note that none of the end points were measured in blinded fashion.

The fusion scoring system has already been presented and it should be noted that standard radiographs were taken, but that did not include flexion extension films.

The sponsor already noted that fusion was defined

as a six or seven at six months, and a five, six or seven at 24 months and 12 months.

How do define fusion success obviously is going to be an important question of the panel, especially in the absence of flexion extension films. This will be one of the panel questions.

Another important end point already mentioned was the disc height. That was measured by the Prolon method.

These are the one to five point scales used for measuring pain. Pain was distinguished into the categories of leg pain, back pain and overall pain. Also, you see the scale for function.

The sponsor already described the five-point standard scale for measuring muscle strength.

As you look at these scales, please note that the higher the score, the better the outcome.

Already presented has been the demographics. I would only call your attention to the next-to-the-last column, where you see that very few patients had three or four spinal levels involved.

A panel question will deal with whether there is enough data to support the use of this device in more than two spinal levels.

In terms of patient accountability, at 12 months,

the percentages of patients available for evaluation was quite high, greater than 95 percent across all groups. Again, note that A-3 is a subset of A-6.

At 24 months, it was not quite as high, but still we have more than 80 percent follow up at 24 months, with the exception of the spondylolisthesis patients, B-3, where it was only 71 percent.

It is going to be important to keep these percentages in mind, because the results, as I present them, are all based on evaluable subjects; in other words, the number that actually were evaluated at 24 months, for example.

I would also make a comment now that there was rather complete data on all of these patients that were evaluable at 24 months.

In other words, when you go through the data, the denominator for the various end points for the A-3 category pretty consistently is 51. For 86, it pretty consistently is 92. So, all those patients were able to render complete data sets.

This first is one of several slides dealing with effectiveness end points. As you have already seen, we have it across all the groups, an increase in disc height. That averaged 3.5 to 4 to 4.5 millimeters.

The last column presents the percentages of patients in each group that were able to increase and maintain their disc height at 24 months.

Now, the first percentage in the larger type is what you get when you look at the evaluable subjects.

Underneath it in smaller type is what you would get in an intent-to-treat analysis. In other words, if all the missing patients were considered to have not increased and maintained their disc height, that would bring the percentages down somewhat.

This is just a reference point. So, you have the percentage for evaluable and then, underneath it, what it would be like in an intent to treat, where you pretty much look at the worst case scenarios as far as the missing subjects.

Of course, fusion rates are going to all depend on the criteria used. To the left you see the results using the criteria that were prospectively set by the sponsor in their IDE.

In general, at 12 months, 90 percent of subjects were fused according to that criteria, and that percentage was even higher at 24 months. Again, this is all on evaluable subjects.

To the right, the rates are a little bit lower,

and in this case we have used a more conservative approach. In other words, this is what the data would look like if a fusion success was defined as a seven.

We asked the sponsor, in fact, to calculate these percentages for us. They are obviously lower. At 12 months, it is between 50 and 65 percent, and at 24 months, 77 to 78 percent.

We obviously are going to ask this question of the panel. What is the most reasonable way of setting the criteria for fusion success.

This table presents the percentages of patients who got better, stayed the same, or got worse as compared to baseline, in terms of their leg pain and back pain.

For each group there are two rows of numbers. The first row provides the percentages of patients at 12 months, and the second row, right underneath that, refers to the data at 24 months.

For example, in the A-3 group, 84.7 percent of subjects were better at 12 months, and at 24 months, 90.2 percent of subjects were better.

The column to the extreme left is just a reminder that we are dealing with evaluable subjects, and the numbers there are the ones that participated in generating this data.

For the B-3 category, the first number, 50, refers to leg pain. So, there were 50 patients evaluated for leg pain at 12 months, and then the 49 refers to the back pain at 12 months.

If you look at these results, in general DDD subjects seemed to make the most improvement in their pain at 24 months. The results aren't quite as good in the spondylolisthesis and the failed fusion patients.

Also note that, not throughout all the cells, but through most of them you will note that there are some patients who, in fact, worsen. There are not very many, but there are some.

The same format applies here in this table, when we consider overall pain. Again, the percentages at 24 months were higher for the DDD subjects.

A panel question is going to relate to what part of pain we should be focusing on. Should it be overall pain? Should it be leg pain? Should it be back pain?

What is most reasonable in light of the fact that the mechanisms perhaps are different for leg pain and back pain.

Most patients come with both and we need some help in determining whether we should keep them distinct in our look at the data.

Again, the same format for these end points of function and muscle strength. At the left it can be seen that generally the percentages of patients getting better aren't as high as seen in the preceding slides on pain.

On the right, as far as muscle strength, muscle strength was one component of the neurological exam. But, if you will, it was considered as a marker or an indicator for neurological status. So, the focus has been on muscle strength.

The vast majority of subjects, in fact, maintained or improved their muscle strength throughout the study.

As far as safety, the post-surgical data are of note. The mean blood loss was 1600 mls. The surgical procedure averaged five hours and the average hospital stay was seven days.

To some, these may be considered increased rates and we are going to be addressing this in one of the panel questions.

This table highlights some of the major complications that we want to draw your attention to. The first number is the number of events and in parenthesis are the patients.

For example, there were two broken cages in two patients. One occurred interoperatively and was removed

immediately and replaced. The one at one month was not, because it was felt to be functionally stable as it was.

If you look down, there were 18 broken screws in 14 patients. These generally occurred six months or beyond.

Broken pedicles, there were six of those, four of them occurring interoperatively.

We would like to call your attention to the dural tears, because this will be coming up in a panel question. There were 44 dural tears in 43 subjects.

That calculates to 90 percent of patients had dural tears, although it needs to be recognized that all but six did not need to go on to further surgery to repair the tear.

Then near the bottom you see 3RSD. It reflects sympathetic dystrophy. There were two foot drops, one was permanent and one was transient.

Then the final category was excessive blood loss. That occurred in two patients. One patient developed DIC and the other patient expired.

Again, these are going to be brought up, or the focus of a panel question.

The next slide shows some other complications. I bring your attention to the wound infections. They occurred anywhere from interoperatively to 24 months, but most of

them occurred within the first month.

A big issue is going to be VSP component removal. The data gathered here is very limited, in that the clinicians did not indicate on the case report forms the reason for the VSP component removal.

It is not clear how many were done because patients were symptomatic, or how many were done maybe for preventive reasons. So, that is a real limitation here.

The data does demonstrate, though, that 87 subjects -- which is 39 percent of the total subjects -- had VSP component removal.

Seventy-nine of these had component removal only, while eight of them had component removal, but that was part of a revision.

Of these 87 subjects, 70 had pre-removal and post-removal pain scores. So, there could be a determination of how they fared as a result of having these components removed.

Now, I should point out that, after the removal, the next scheduled follow up was the point at which they gathered the post-removal pain scores.

That interval from post-removal to the assessment of the post-removal pain varied across these subjects.

The findings were 26 percent of the subjects

improved in pain from pre-removal to post-removal. Fifty-seven percent stayed the same, 17 percent worsened.

Therefore, 74 percent of subjects did not improve in pain following component removal. It needs to be recognized that how many of these subjects had pain going into this removal is not known.

A panel question is going to deal with how we should regard these VSP component removals and how it should be considered in our equation of what is considered a success.

The issue of overall success, again, it depends all on the criteria used. The criteria at the top are the ones prospectively used by the sponsor.

It can be approached a number of ways. We really want panel input on this. If you took it from the patient's perspective, in addition to the criteria above that they might have some other expectations going into surgery, in fact, they may want to see improvement in whatever pain brought them to surgery in the first place.

In other words, if they had leg pain, they want to see that better. If they had back pain, they want to see that better. If they had both, they want to see both of those better.

They certainly expect not to have any serious or

permanent complication. Maybe they would not feel they were a success if they had to go back for another surgery. Again, we want your input on this.

These are the overall success rates. Those on the left are per the prospective criteria set up by the sponsor, and those on the right are using a more conservative approach.

On the left, the overall success rates averaged 67 to 85 percent, if you look at the A-6, B-3 and C-4 groups. Again, this is all on evaluable subjects.

If you look over at the right -- the sponsor was actually asked to calculate these. They are recalculated in a more conservative manner, considering all types of pain that the patient had coming into the study. So, they would have to improve in all aspects of their pain.

Serious complications were weighed into the equation and that was according to the definition the sponsor has already presented to you.

In terms of that, the only patient that was excluded from the success category because of a serious complication was the permanent foot drop.

Obviously the rates on the right are lower. You see generally 45 to 65 percent at 24 months.

It should be noted that even with the more

conservative approach that was taken on the right, not all device removals and revisions were included.

As was noted in the footnote, a removal was only excluded from the success category if, at the time of this removal surgery, it was noted that a pseudoarthrosis was present.

So, there was some effort to remove those who had removals from this overall success category, but it had to be those who actually demonstrated pseudoarthrosis on this second surgery.

Also, it needs to be noted that these 24 month results represent a real mixture of patients. In other words, some of these patients were those who had original surgery and then were followed 24 months.

Some others of these patients had the original surgery. They went a while, had a second surgery, and then they were followed for a variable amount of time, until at which point they reached 24 months from the original surgery.

So, the time course of these various patients and their route to the 24-month end point varied, and that needs to be recognized as you view these percentages.

In effect -- this is a really important point, because in effect no distinction here has been made between

a device implanted in a single surgical procedure from a device that had to involve a two-surgery approach.

We are going to want some guidelines or guidance from the panel regarding how we should view this notion that some patients got by with one surgery and other patients required two.

The sponsor already presented the literature controls. I am not going to go through all the criteria they used. It is a very difficult process to find literature controls.

They tried to pick articles that had consistency, especially in the end points to be measured, so they could be compared across articles.

If you note at the bottom of the slide, the data sets that were ultimately selected are given there. In other words, the A-6 or DDD subjects were compared to DDD subjects in the articles who were treated with non-fusion surgery.

The B-3 or spondy patients, were compared to spondy patients in the literature who were treated with either non-instrumented or instrumented fusion.

The C-4, or the failed fusion, were compared to articles where the patients had failed fusion treated with either non-instrumented or instrumented fusion.

Then for a complications comparison, articles were

selected where the subjects varied in diagnoses, but all had non-instrumented pliff.

As mentioned before, literature selection is very difficult and there are always some limitations. I want to point out just a few of them.

The comparisons were not necessarily at the same time points. In other words, in this study we know that patients were evaluated at six months, 12 months, 24 months.

That is not the case with the literature, because often the literature doesn't tell you that.

It just says, well, on average the main follow up was, or they say the range of follow up time periods was. You do not have a direct comparison of a given result at a given time point.

There are also differences in terms of diagnostic categories and how they were defined, differences in inclusion/exclusion criteria, how fusion success and clinical success were defined. That varied.

Most studies, as I already mentioned, had any idea in there of overall success.

Additionally, different graft material was used in the various studies and, while some of the studies plainly said that non-instrumented fusion was the procedure, often you don't know what specific surgical procedure was used.

In the statistical review that will follow, there will be a treatment of the comparison of this device with the literature, as far as effectiveness parameters.

When it comes to safety, I wanted to highlight a few things. This slide points out the percentages of patients who had complications, and then what the range was in the literature.

If you look in the second row for dura related complications, they were higher across the board. So, all three major groups with this device reported higher dura related complications than found in the literature.

If you look at the bottom row on removals, again, it is higher, but if you noted from my previous slide, not all those articles involved patients who got instrumentation. You really can't make a fair comparison there.

In terms of overall findings, it appears that patients who received the Brantigan cage with VSP spinal system had an increased disc height that was maintained at 24 months.

The device appears to have provided adequate fusion, although the fusion rate will depend on the specific radiographic criteria used.

Likewise, an overall success rate will depend

ultimately on the criteria selected.

For over one-third of subjects, the use of this device involved a second surgical procedure for device removal, which in almost all cases was removal of the VSP components.

This component removal, however, did not improve pain in the majority of subjects, but the data are not available as to how many of those removals were because of pain.

The safety data indicated what, for some, may be considered a large blood loss, a high number of dural tears, a long hospital stay associated with implantation of the Brantigan cage with VSP spinal system.

Dr. Ponnapelli will now provide a statistical review.

DR. PONNAPELLI: I am going to present the following three statistical analysis, pooling the diagnostic groups, how justifiable is it, covariate analysis to find the influential covariates, longitudinal analysis to see how success rates changed over the time period.

Let us recall the three diagnostic groups which were introduced with the sponsor. Within this group were referred to as A-6, spondylolisthesis patients were referred to as B-3, previous failed fusion, we refer to it as C-4.

Dr. Glass already gave you the sponsor definitions of fusion, clinical and overall successes.

A more conservative definition of fusion such as includes only seven in the radiographic assessment of fusion, where the sponsor includes five, six or seven.

A more conservative definition of clinical successes also includes no further surgery and no serious complication besides improvement in pain, maintenance of the improvement in function and muscle strength.

Since there is extensive regrouping of the original randomized and non-randomized groups, there is potential for bias if the groups are not pool-able.

This gives a pool-ability analysis for all the groups and for all the types of successes. The test used here is chi square. They use both the exact chi square and also the asontrotic(?) test.

Notice that p values are quite high and there are some stars in some cells. Where there are stars, there is 100 percent fusion. That means they are completely pool-able.

The other cells are pool-able because the p value is quite high. This is true in the A-6 as well as B-3. I don't know about C-4 here because C-4 is a sort of poor(?) group because it is already a mixture.

That was according to the sponsor's definition of success. This slide shows the situation with the conservative definition of success.

Again, we consider both the A-6 and B-3. For all types of success you see that the p values are quite high and they are pool-able except for possibly A-6 and fusion. It is marginal but it is still pool-able.

The next slide, I answer the question, what are the covariates that influence success. The answer to this, the largest group was fitted with site, level of disease, smoking status, pain at baseline, function and Worker's Compensation as covariates. The results are given in this slide.

For A-6, the only covariates that influence the result are pain for clinical success, side B pain and side B for overall successes. For B-3, there is no significant covariate. For C-4, Worker's Compensation and function and pain, Worker's Compensation and function for clinical success, Worker's Compensation, pain and function for overall success.

This is with the sponsor's definition of success.

The influential covariates for the conservative definitions of success are given in this slide. You see that with the A-6, site GA(?) and pain are the influential

covariates.

With B-3, pain is the influential covariate. With C-4, Worker's Compensation, site GA(?) are the influential covariates.

Of course, it depends on which covariates had influence on the result, it depends on which successes we are talking about.

My next slide, this slide shows the comparison with literature controls, with sponsor's definition of success.

Also, I indicate, the number of studies in the literature, because to me, it is important. In the A-6, for fusion success, I will read a few of them and you will get the flavor.

For A-6, for fusion success, it appears that it is 100 percent. In the literature it is 87. With B-3, it is 100 percent, in the literature it was 82. The first one was based on 10 studies, the second one on 11 studies.

The only case in which the PMA device failed worse than the literature is for B-3 and for clinical success. The PMA device has had 79 percent, the literature had 86.3.

It turns out that this is not statistically significant. The p value turns out to be .24.

For overall success, it turns out that there are

very few historical studies from which the percentages of successes with the more conservative definitions can be gathered.

All the above data that I talked about is at the 24-month time period. To examine how success rates change over time, we presented numbers obtained from longitudinal data analysis over one, three, six, 12 and 24 months.

This is the longitudinal results with the sponsor's definition.

You see that over time, either there is an improvement or it stabilizes, the person stabilizes.

It turns out -- I am referring to this slide -- the same is true with conservative definitions of success.

Now, I summarize my observations. With the sponsor's definition the data are pool-able and they do exist at least as efficacious. The weighted average of literature study for A-6 and B-3, there is weighted evidence and we also did a meta analysis.

The same conclusion holds for C-4 also, but with considerably fewer literature studies to compare.

With the more conservative definitions of success, there are no comparable data.

There is not a large enough number of subjects with multi-level disease to judge the efficacy of the device

for greater than two levels of treatment.

That is going to be one of our panel questions, whether statistically they seem to be inadequate, whether they are adequate otherwise.

In general, the success rate improves with time.

MS. MORRIS: That concludes our presentation. At this time I can review the panel questions, if that is what you would like.

DR. BOYAN: What I would like to do right now is open the meeting up to some general discussion, where we can ask you and the sponsor -- I just was reminded of the lead reviewers.

I think before we open it up to general discussion, we will have the lead reviewers for the panel give their reviews. Then we will open it up where the panel can ask questions of the sponsor or of the FDA.

After the lead reviewers do their reviews, I think we need a five-minute break. So, we will get all the reviews out of the way and then take the break and come back for the questions.

The person who has been identified as our lead clinical reviewer is Dr. Leon Grobler.

DR. GROBLER: I think the issues that come up in this study have been clearly identified. I just thought I

would review a couple of thoughts that, in essence, would just lead our discussion through this application.

I think in the first place the clinical groups that we need to consider, I think it is important for us to realize what groups we are dealing with.

I see one of the reviewers actually brought the A-3 group as a separate group and talked about recurrent discs.

I think the three groups as we discussed is the DDD, the spondylolisthesis and the failed fusion. I do feel we need to discuss the middle group especially, the spondylolisthesis.

I think the subsections need to be discussed, as I don't think we can categorize as lytic and degenerative under the same category as indication.

Secondly, I think the factors we need to consider are pretty clear. Firstly, we need to look at the correct indication for the procedure.

We need to bring in the question of the osteoporosis and all, and refer that to the panel discussion.

Once again, the goal of the procedure needs to be clearly stated. The goal, in my mind, is that the object here is to get solid fusion at a functional spinal unit.

We should be very clear that bringing in this as a treatment for pain can muddy the discussion.

Thirdly, as far as the adequate training of surgeons involved in this procedure, I think that is extremely important.

It doesn't take much to realize that this approach is not the first step in a spine surgeon's armamentarium, as soon as you finish spinal training.

First, interbody fusion is not a procedure that should be seen as likely, and the complication rate should be assessed.

I think extreme emphasis needs to be placed on training of surgeons.

That will bring into discussion the question of safety of the procedure. It will bring into the discussion the question of is this procedure effective in obtaining these goals.

As far as assessment of the safety of the procedure, I have some concerns, and we will discuss that, regarding adequate information that needs to be given to surgeon and patient alike.

The one concern I think we all have -- I, at least, have -- is working around a non-prospective controlled randomized study.

That is probably really the only way that we can assess safety and efficacy. Therefore, I think a fair amount of discussion needs to be centered at that point.

As far as outcome, patient satisfaction, both pain and functional, and then most probably what is one of the biggest factors that will dictate success in these cases -- bony fusion -- how do we assess that.

There is some controversy in this case, what categories we need to assess, and the fact that certain, at least accepted ways of assessing fusion were elected not to be done.

The acceptable complication rate, going back to the size of this procedure, I think we should be very clear that this complication rate should be seen in context with the size of this procedure, the graveness of this procedure.

That brings us back to the indication. Having this procedure being done with relatively minor cases, where going all the way is really not our goal, will increase our complication rate.

Then the question of second and third procedures, I think that is a major issue. I still have a problem as far as what needs to be done and dictated to surgeons and patients as far as removal of this implant.

I do not have a clear handle on the status of

elective removal of implants. It seemed to me initially, in the initial application, it was clearly stated that this had to be done to assess fusion.

Later on it seemed to be more on the basis of preference from the surgeon and from the patient, I guess.

Then last but not the least -- that was initially discussed, but I think needs further discussion -- is are we talking about -- we need to realize we are talking about discussing designing a system where two unapproved implants are assessed together.

I think that is somewhat of an important issue. Especially, how often are one of these procedures being done alone?

Are we going to set the stage that it is the choice to use one or the other, or is this a package deal? I think that is extremely important, at least from a clinical point of view.

I don't think I am going to go any further, but I think those are the factors that came out of the presentation, and need to be gone into in discussion. Thank you.

DR. BOYAN: Thank you. Our lead preclinical reviewer is Dr. Joseph Hale.

DR. HALE: My comments for the preclinical testing

are going to be restricted pretty much to the mechanical testing that was performed.

I think the sponsor's choice of testing in the absence of any performance-based standards for these devices was, for the most part, appropriate.

The sponsor or the FDA reviewer can correct me if this is incorrect, but I don't recall seeing any data, though, that was ever presented for the construct being implanted either in a cadaver specimen or in an animal specimen and any mechanical testing done.

All the mechanical data presented were for either the cage tested mechanically by itself or the VSP system by itself.

There was never any test data, that I am aware of, where the entire construct was implanted and evaluated in that form.

I also have a couple of other concerns, and this may just be some confusion on my part in terms of the data for two of the tests.

Within the panel pack there was a static torsion test that was reported, and the values that were obtained for that were for a single cage, where they were below the projected torsional loads on the L4,L5 motion segment.

The rationale was that if two cages were used, it

would exceed the anticipated loads, but for a single cage, a torsional strength of 8.07 newton meters was reported with an anticipated load of 10 newton meters. That was somewhat concerning.

Also, the data that was presented for the compressive sheer fatigue test, at least within the FDA information that was presented, listed that as being 2,800 newtons with a run-out of five million cycles.

Again, the anticipated load in the body at L3,L4, in this case, was 3,000 newtons. Again, there is a sub-physiologic loading strength of the device.

The one other concern I had, or question, had to do with the one cage that failed post-operatively. I was just curious whether there had been any type of failure analysis done to determine what the exact mechanism of failure was for that particular cage.

DR. BOYAN: Finally, our statistical reviewer, Dr. Janoski.

DR. JANOSKY: I have a couple of issues that are still remaining. One is the safety information and the complication rates, and sort of a dissection about what those are.

Another is an analysis of the potential covariates. If I look at the GEE results that were

presented as follow-up to questions raised by FDA statisticians, it seems that there are many covariates that are actually impacting quite significantly on the findings, these being things like Workman's Compensation, body size in terms of obesity, et cetera, et cetera.

So, you are seeing very differential effects depending on the level of these covariates, which the sponsor didn't address today.

Finally, the comparability of the sites and the physicians, I am not convinced, if we look at the conservative definitions and if we look at some of the secondary end points, whether that pool-ability is actually appropriate or not.

DR. BOYAN: Why don't we have, right before the break, one rendition of the questions, like in toto, but without the lights turned down.

MS. MORRIS: I just want to bring to your attention that I made some slight revisions on the questions.

If you refer to the list of questions in the packet of slides -- it is at the very end -- you can follow along with me and it will look like this. Are we all on the same page?

I will begin with the clinical questions, the

first one being, the FDA has traditionally defined fusion success as solid fusion.

Typically, determination of fusion involves an assessment at 24 months, looking at a degree of translation and rotation from flexion extension radiographs, the presence of halo, evidence of trabecular bone bridging, or any combination of the three.

In the PMA, the sponsor used the following criteria for assessment of fusion from plane radiographs: category zero, no fusion done, category one collapse of construct, category two resorption of graft, category three large lucency around the entire graft, category four small lucency partially around graft, category five bone bridging in fusion area, category six increased density of fusion bone, finally, seven has a continuous trabecular bone bridging fusion.

Based on this assessment, what is the minimum category that should be used in determining success in fusion.

The second question, the sponsor collected data on pre and post-operative pain using a five-point scale for the following pain locations: back, right leg, left leg, graft donor site, and overall.

For the assessment of improvement in pain, when

determining clinical success, the sponsor used the data collected on overall pain.

Considering that in some patients there was an improvement in some, but not all, pain locations recorded -- that is, a patient with an initial complaint that included both back and leg pain may not have shown improvement in both locations.

Overall pain may represent a composite of a patient's pain, including donor site pain.

Should the assessment of pain be based on the data collected on overall pain, or should pain be assessed at a specific site at the original complaint of pain.

What is the best method of assessing patient pain to determine clinical success?

What amount of improvement in pain would warrant a success for this outcome parameter?

The third question -- and some numbers have changed, if you have reviewed these questions earlier, to reflect what Dr. Glass had presented.

Over the course of the study -- that is, out to 48 months -- approximately 43.4 percent -- that is, 96 out of 221 patients -- of the subjects underwent a second surgery.

Thirty-nine percent of the subjects underwent a second surgery that involved a removal of the VSP

components.

The sponsor reports that, based on the data as recorded on the case report forms, that it was not possible to stratify the various causes for removal.

The sponsor states that the removal of the VSP components were not considered treatment failures, since a great number of these were presented as elective removals.

Therefore, do you consider the removal of VSP components as an expected second surgical procedure?

Do you consider the removal of the VSP components as a recommended second surgical procedure for the indicated population for the PMA?

Finally, when should the removal of the VSP components be considered as -- it is written device failure, but I mean treatment failure.

The fourth question, overall success was defined by the sponsor as success in fusion as defined by category 5, 6 and 7, improvement in overall pain, maintenance or improvement in function, maintenance or improvement in muscle pain, no serious or permanent complication as defined by the sponsor, no further surgeries related to the revision or removal of the cage only.

Considering your responses to the above-stated questions one, two and three, and review of the patient

complications, is the overall success definition used by the sponsor appropriate, or would you recommend any modifications?

If modifications are recommended, please specify these modifications.

The fifth question, with regard to the safety of the device, are the number of dural tears, amount of blood loss and length of hospital stay, as specified -- a total of 44 tears, 600 mls on average, and seven-day hospital stay on average -- as assessed at 24 months considered within the realm of clinical expectation for a device and procedure of this type.

Based on the final data presented, the sponsor is seeking approval for the indication for the Brantigan IF cage with the VSP spinal fusion system -- that is, the IF cage, pedicle screws, plates, washer, transverse connector -- for an open posterior approach, using otogenous bone graft in the following patient populations: Degenerative disc disease with or without one or more prior surgical interventions at one or two levels from L2 to S1.

Degenerative disc disease is back pain with degeneration of the disc confirmed by history and radiographic studies, as defined by previous panel input.

B, spondylolisthesis without prior surgical

intervention at one or two levels from L2 to S1.

Finally, previous failed fusion with one or more surgical interventions at one or two levels, from L2 to S1.

Does the clinical data support the safety and effectiveness of the subject device for each of the above patient populations for one and two levels.

With respect to the indication for spondylolisthesis, if there is sufficient data to support the approval of this indication, should the type of spondylolisthesis be specified to match the types represented in the study cohort, that is, degenerative or lytic.

If the data does not support the approval of the subject device for one or more of the stated indications, what data is necessary to demonstrate the safety and effectiveness of the subject device for those indications for use.

Does the clinical data from one to two levels, and the limited data from three to four levels, support the safety and effectiveness of the subject device for use in treating greater than two levels for each of the above-stated indications.

If not, what data is necessary to demonstrate the safety and effectiveness of the subject device for treating

greater than two levels for the respective indication.

We have the following labeling questions.

With respect to component removals addressed in question three, can you recommend appropriate device labeling to address this issue.

The sponsor has included a precaution in their package insert, which states: Posterior lumbar interbody fusion with Brantigan IF cage and VSP spinal system should be undertaken only after surgeon has had hands-on training in this method of spinal fixation and has become thoroughly knowledgeable with spinal anatomy and biomechanics.

Should surgeon training be a requirement for this device and procedure?

The sponsor has included severe osteoporosis as a contraindication in their package insert.

Is severe osteoporosis a contraindication for this device and procedure?

Finally, what other labeling recommendations can you provide with respect to the package insert, patient information brochure and the surgical technique manual.

I will reserve question 12 after panel discussion and recommendation.

DR. BOYAN: Thank you.

(Brief recess.)

DR. BOYAN: What we will do, we will have a round of questions where we will go around. Each individual can ask one, two or a short series of questions that are directed either to individuals from Acromed or individuals from the FDA or each other, or wherever they feel they can get the answer they need to clarify issues.

At the end of that time, we will then take up individually the questions that the FDA has put to us.

Finally, when we feel that we are ready, we will come to a vote.

It is my intention that this whole section, for those of us who are starting to get a little bit nervous about what time it is, this section of the meeting should not exceed about an hour.

I want to make sure that the device gets a fair hearing. If we are onto a subject that needs a little bit of extra time, we will go ahead and we will do that.

So, let's begin. There are more people on this side sitting, so why don't we begin with Dr. Cheng.

DR. CHENG: Thank you for presenting that data, both the FDA and the sponsor. One of the panel questions -- actually the very first one that we were asked -- dealt again with this issue of definition of fusion.

Perhaps there is some information that you have,

actually, in your data which might be helpful.

That is, in determining the appropriate level, where to place the cut off mark for success of fusion, do you have information on the number of patients who underwent reoperation for removal of the device?

At that time I believe some patients who were in the conservative analysis of success were removed because they had presence of a pseudoarthrosis.

Can you tell us how many of those patients actually had pseudoarthrosis and what was their radiographic reading before they underwent reoperation?

In other words, do you have some data in which to correlate or validate the radiographic exam with the subsequent gold standard of interoperative examination?

DR. BRANNIGAN: Yes. I would like to address that. We did explore 115 levels with surgery. That finding at surgery was correlated with the radiographic findings that were recorded in the immediate previous radiographic analysis.

This analysis was not prejudiced by the surgical findings. In this, if we include 5, 6 and 7 as the fusion criteria, we had 94 percent accuracy in determining presence or absence of fusion.

As you said, the surgical exploration is the gold

standard of fusion historically, and we had a very accurate assessment confirmed by this surgical exploration.

Furthermore, I would like to point out that in this fusion assessment surgically, we had an opportunity to be very accurate.

When you removed the pedicle screws, you could place probe devices in the pedicles, and then you have mechanical control in those segments, and you can explore very subtle degrees of motion.

I think that data is very accurate regarding the presence or absence of fusion.

DR. CHENG: Did you submit that to the FDA? I believe most tests are evaluated in terms of the positive predictive value. Was that done?

DR. BRANNIGAN: Yes, we did. The positive predictive value was calculated including 5, 6 and 7 in the fusion category as 94.39 percent.

DR. CHENG: How many actually had pseudoarthrosis? I didn't hear you say that.

DR. BRANNIGAN: Excuse me a minute. In this study, this included exploration not only of the cage levels but some of the ETO levels. So, there were more pseudoarthrosis because of the ETO. We had 12 total pseudoarthroses identified.

DR. CHENG: Am I allowed to ask another question?

DR. BOYAN: Yes.

DR. CHENG: You started off doing a randomized study of the allograft or the donor bone compared to the cage.

Perhaps you said this and I missed it. Why did you change the methodology of the study from a randomized study to the longitudinal prospective study?

DR. BRANNIGAN: I would like to explain that. In my own surgical practice, I was obviously involved in this from the beginning.

I was doing ETO sterilized allograft pliff in this same group of patients before the IDE was approved.

At the time the IDE was approved, I looked back over my personal series of donor bone pliff operations and determined at one year I had a fusion success rate of 56 percent and a clinical success rate of 60 percent.

If you are willing to do those surgeries just routinely, you should be willing to do those in a study as well.

I was, but the first thing a patient asks you when you suggest the randomized control is what do you know about both arms of the study.

Of course, I felt obligated to tell them what I

knew about the donor bone pliff. When I did, the patients routinely refused to participate in the study.

So, in my practice I enrolled just one patient in the randomized study. Then after 10 or 15 of the patients just refused after the recitation, I did not attempt any further.

Now, I know the other investigators were informed of this and they did enroll some patients, because there was a different point of view.

We had some other data that suggested that maybe the fusion rate with the donor bone was as high as 85 percent, and some of the doctors felt it was 85 percent. But one by one, they dropped out.

A retrospective cohort study in Europe came up with a donor bone fusion rate in the range of about 72 percent. So, in fact, it was some place between those two limits.

One you explained to the patients what relative information you have, they were not willing to participate in the randomization.

DR. CHENG: If that was the case, how was a randomized study ever approved by an IRB board, given those figures you just mentioned?

In other words, given that history, most IRB

boards would not approve a randomized study or, if one was ongoing and an interim analysis showed the results you just said, they would have been stopped.

DR. BRANNIGAN: I think it is a matter of the timing. One of the questions the IRB asked is, what type of surgery did you do in these patients before this IDE came around. Honestly, I said, I did donor bone pliff.

They determined that, if that is what you did for those patients before, then it should be a reasonable operation to continue during a study.

DR. CHENG: Thank you.

DR. BOYAN: Thank you. Dr. Kerrigan?

DR. KERRIGAN: I have some concern about the use of the Leichart scale for pain. I wondered with this covariate thing, your definition of improvement, I believe is just an increase by one or more points in the Leichart scale.

My question is -- that is difficult, I think, when you are comparing this type of data to literature. I am sure in the literature they were not all using the Leichart scale.

I just wanted to know whether you attempted to use any other kind of pain scale or looked at the improved and defined clinical improvement in another way, like a change

in two points or something like that.

DR. BRANNIGAN: I think the closest comparison of literature is the Prolo scale, which is a 10-point scale of pain and function, five points for pain and five points for function.

What we originally outlined in the study was very similar to the Prolo. We collected a little bit more data.

Our initial evaluation of this data group that was submitted to FDA basically gave that evaluation. So, I think it is very exactly similar to the Prolo scale which is commonly used for other interbody fusion studies.

In fact, the regrouping with one point of pain improvement and the other parts of overall improvement really produced exactly the same data as the Prolo scale.

When I say exactly, the numbers were the same. Some patients went from one category to the other, but the overall results were the same as the Prolo scale, and that follows what was in the literature.

DR. GROBLER: Dr. Brannigan, I have a question on the animal study. If the fusion rate was five out of five, you had 100 percent fusion rate in the animal study in the goats; right?

DR. BRANNIGAN: Yes.

DR. GROBLER: Now, that was done without

instrumentation.

DR. BRANNIGAN: That is correct.

DR. GROBLER: What made you decide, with that information available, to add instrumentation in the human?

DR. BRANNIGAN: I think the biggest reason is that this technology came together. Pedicle screws were used clinically before the cages were made, and the pedicle screws had a number of failures due to broken screws.

It appeared that there was something that was necessary in some of those pedicle screw cases to gain mechanical support.

Additionally, the pliff that was done originally by Clowerd had good results in his hands and those of Maugh and Lynne. Yet, most other surgeons really had difficulty reproducing the quality of results that they had.

Furthermore, a failed pliff has a very horrible clinical outcome. So, at the same time, those of us who were working with pliff at that time felt that for donor bone pliff, you really ought to have some added posterior fixation.

So, the two technologies appeared to solve the limitations of each prior surgery. That is really the reason why they came together.

I guess the follow-up question is to ask, do you

really need both together.

What I can tell you is that there are a number of surgeons throughout the world who are doing our cages without the VSP.

I can think specifically of one surgeon in New Zealand who did a series. I traveled to New Zealand and I operated with him and I saw a number of his patients.

He had on the order of 90 percent plus fusion success without the pedicle screws. However, he had close to 100 percent with the pedicle screws.

Since then, he has discontinued doing the cage pliff without pedicle screws.

The one other experience that I am aware of is Dr. Sabato and his associate, Dr. Handel, in Jerusalem. I visited his hospital and worked with him.

I was not able to get a real handle on the quality of his results and how they compared to the others. What I can tell you is that worldwide, most of the surgeons who do these cage pliffs, when they have the elective of doing it with and without VSP, are doing it with the pedicle screws. That is what we very sincerely recommend as getting the best results.

DR. GROBLER: Can I ask one more question?

DR. BOYAN: Sure.

DR. GROBLER: The question about surgeon training I alluded to, and I think we all agree this is not a first surgical procedure.

What is the sponsor planning to do to make sure that putting this out there to be used will be safe for the patient?

I guess that is a pretty extensive procedure. We can discuss that at some time, the question of the blood loss, the time on surgery. It is a huge procedure. How are you going to train the surgeon?

MR. CHRISTIANSON: We have submitted to FDA in our PMA a proposed three different methods for training surgeons.

We are prepared to hold regional workshops, so that we will go to a geographic location and invite surgeons to come in.

We have a draft agenda of the program put together which will present the results from our clinical studies, the results from our biomechanical studies and our preclinical studies on the animals.

We will have a surgical technique video for the surgeons to view, so they can see some of the techniques and the specific steps that are gone through, and then an opportunity to practice on a saw bones model.

Alternatively, we have obviously trained centers from our IDE investigation sites. We will provide opportunities for individual surgeons to visit trained surgeons and operate with them at their sites.

Thirdly, there are a number of surgeons who are already experienced in doing a posterior lumbar interbody fusion using Acromed's pliff brote(?) system, which has been commercially available for years in this country.

For those surgeons, we feel that a trained Acromed representative can go through the surgical technique manual for the Brantigan cage and adequately train them, since they are already accustomed to the approach and the delivery of a rectangular donor bone block to the interbody space.

The addition of delivering a pliff cage is a minor additional step in the procedure, for those very experienced users.

DR. PERSENAIRE: I might add that in Europe we have 70 years experience of a visiting surgeon training program. We have a training program for surgeons by visiting experienced surgeons, scrubbing in and getting a lecture on not only the technique but also the pitfalls, which works very well.

DR. HALE: I would like to return to some of the biomechanical concerns that I brought up earlier and get

some clarification or answers about those.

With regard to testing the cage with the VSP system, do you have any data where you have looked at that construct in vitro?

DR. SIRHAN: Actually, we did not test them together, no. We tested them separately in a worst case scenario.

If we were to test them together, the implant or the VSP system would be stress shield, because most of the axial load will go through the cage.

As we demonstrated, our axial compressive strength of the cage is higher than the compressive sheer. So, we elected to test the cages under sheer and the VSP system in a model where all the load is transferred through the VSP implant.

DR. HALE: I think that may depend on some of the -- if you were looking at other loading modes, bending modes, maybe that wouldn't be as clear cut as all of that, but that answers the questions.

As far as the two test modes where there seem to be sub-physiologic loading levels, could you comment about that, the static torsion test data and also the compressive sheer fatigue numbers, which are seemingly less than the physiologic loads that are expected?

DR. SIRHAN: One thing first, we recommend using two cages per level. So, for the torsion it is 2.7 per cage. You are talking about two cages. Aside from that, you have the VSP system in conjunction with the cages, as well as you have a part of the annulus ring. During surgery they don't remove the disc totally.

So, these three will address the in vivo loading of 10 newton meters.

As for the endurance limit and compressive sheer, it is 2,800. That is the worst case scenario for one cage. For two cages, it is 5,600 newtons, which is almost five times the average daily activity of 1,200 newtons.

DR. HALE: My concern specifically is in reading through the material, it was my understanding that there are some instances where it would be clinically feasible for the surgeon to -- or clinically necessary for the surgeon to -- implant only one cage.

If there are always going to be two cages, I would agree that it is not a problem. In those instances where there is only going to be one cage, it seems that could potentially be a failure point.

MR. CHRISTIANSON: I will answer as a non-clinician first. Our labeling recommends the implantation of two cages per level. That would always be our

recommendation.

We discovered, during the course of our IDE study, that there were cases where surgeons were confronted with an anatomic situation where they were not able to safely place two cages, and they elected to place one cage.

In those cases, we have included all the clinical data in our presentation from those single cage level uses.

The surgical technique, in fact, anticipates that and recommends that if the surgeon is confronted with that situation, that they place the cage as close to the midline as possible.

Recall, as Dr. Sirhan told you just a minute ago, that our implantation technique also recommends implantation with VSP plates and screws.

So, we do have this belts and suspenders approach, if you will, with solid anterior column support provided by the cage with a supplemental posterior fixation.

DR. HALE: That is where we having the in vitro data of having the whole construct implanted, though, would be useful in demonstrating that those values were in excess of the expected load, but that answers the question.

DR. BOYAN: I have a question to ask you about the particulate debris that is forming that you observed in the goat study.

In your experiments where I think leaching might give you a sense of chemicals or residuals that might be present in the material, the particulates aren't going to be identified in that kind of testing.

Yet, you do have evidence that particulates do form, and I would anticipate that they would, given the materials that you are using with the short carbon fibers incorporated into the material.

Do you have any sense -- have you done any kind of wear testing to look at particulate formation, have any sense of the kind of load the patient might expect over the lifetime of one of these devices?

MR. CHRISTIANSON: None of the bench testing that we have done has evaluated the potential for the production of wear debris.

DR. BOYAN: Do you anticipate doing that?

MR. CHRISTIANSON: We currently have no plans to do a bench test for the production of wear debris. As you observed, from reading our goat study report, we did indeed see a small amount of particulate matter in the fibrous membrane immediately adjacent to the struts of the cage in the histological sections that were analyzed.

The pathologist who read these studies is also very familiar with the phenomenon of osteolysis that one

sees in total hip and total knee arthroplasty.

He looked specifically for this phenomenon and did not see it. The fact that we had some limited amount of particulate generation, in his opinion, does not lead to a clinical concern that I think you might be asking about.

DR. BOYAN: I am actually not, because the study that you did perform, in my estimation, was relatively short term.

I am not viewing this as a major negative. I am just bringing it to awareness that particulates were definitely observed, and it is likely that they would be there.

No implant in the body is expected to last forever and there needs to be some plan for that, even with cages, that you do have a device that you now have evidence would produce particulate, that could long-term have a consequence, just to be aware of it.

MR. CHRISTIANSON: We certainly are aware of that. As you may recall from my introduction, we have been selling these cages worldwide for a number of years.

To date, we have not had a single report of a human case where anything approaching this type of phenomenon has been reported.

DR. BOYAN: How many years total has the

experience been with the implants?

MR. CHRISTIANSON: Since 1991 in Europe.

DR. YASZEMSKI: The submission to the FDA that I looked through stated that, at the discretion of the surgeon, a posterior lateral fusion could be added to make a 360 fusion.

It wasn't clear to me, as I listened to the studies today -- and this is the question I will ask to start -- of the 221 surgeries, were any of them 360s? Were some of them, were all of them, or were they all anterior fusions with the additional posterior instrumentation?

DR. PERSENAIRE: We don't have the exact numbers available right here. Having gone through the data numerous times over the past couple of months, I can say that it is the majority of cases who received also a post-lateral fusion.

I draw that conclusion by looking at the raw case record forms where I saw the fusion marks checked.

For the record, we evaluate the fusion at the interbody level.

DR. YASZEMSKI: That was my follow on question, so I will just state it. If the fusions were both anterior and posterior lateral, to which or to both of those locations was the radiographic scoring applied?

Was it indeed, as you suggested, only applied to the anterior portion of the fusion?

DR. PERSENAIRE: It was only applied to the anterior portion of the fusion, also partly because some literature suggests that in the presence of a solid anterior column support there is no sufficient compression on the posterior bone grafts to keep it alive, and it slowly resorbs. I refer to an article by O'Brien.

DR. YASZEMSKI: I think it would be nice to know, as we go on, just exactly how many were 360s and how many were anteriors only.

My next comment and question is with respect to the 40 percent removal. Again, I noticed in the instructions to the surgeons in the label, it was mentioned, at least in the copy that I have, that removal of the VSP is recommended after fusion has been demonstrated.

I will ask first, is that truly your position still, to recommend removal? I didn't think I heard that as the presentations were going on.

MR. CHRISTIANSON: The label that you reviewed is a copy of the package insert that we currently put with our VSP plates and screws.

As you may recall from our presentation, those are currently available under a 510(k) in the United States,

labeled for severe spondylolisthesis at L5,S1.

The removal language in that insert, in fact, was directly negotiated with FDA at the time our 510(k) was cleared.

In fact, FDA had some very specific requirements in the removal language.

In the context of this PMA, we have been discussing with FDA a slight modification of that removal language.

I am not certain if you saw the current version. We are a little bit softer in the package insert, as we have currently negotiated with the FDA on our recommendations for removal.

We now have a section called consideration for removal and we list the considerations that the clinician should be aware of in discussing with, and deciding with his or her patient, whether or not a subsequent removal should take place.

DR. YASZEMSKI: I think my last question or comment, I should say, with respect to the removal is, I am looking through my notes. I think it was 87 or so patients in the study who had removal, but the reasons for removal weren't collected on the data forms.

That would be real useful and interesting from my

perspective to be had, and could probably be gotten from a patient chart to understand why so many of them came out.

DR. BRANNIGAN: I would like to comment on that. I think I did probably more surgeries in this series than other investigators, and accounted for a lot of the removals.

In my cases, I followed the FDA's instructions, that removal should be considered a routine event.

I explained that to my patients before the surgery, and they realized that they would have about a 35 percent chance of removal, and that my recommendation followed FDA's recommendation that it should be done routinely.

There were some patients who came back feeling no pain and full function that simply did not want to have the second surgery and I allowed the request.

Other than that, I recommended routine removal of the VSP implants.

DR. YASZEMSKI: Do you still feel that should be the case, or only in response to a specific patient complaint regarding the hardware?

DR. BRANNIGAN: Simply based on what we know about orthopedic surgery in general, the AO group who studied internal fixation of fracture healing for probably longer

than anyone began recommending in the late 1960s that internal fixation devices be removed once bone healing has occurred in weight-bearing bones.

In my practice of general orthopedic surgery, I removed 100 percent of the implants for fracture fixation, and I believe that is still a current orthopedic recommendation.

DR. YASZEMSKI: I have one more comment, if I might. With respect to the dural tears, I think Dr. Hall mentioned in his presentation that the dural tear rate in this study was higher than the upper limit of that reported in the literature in a few historical studies.

What I was wondering, I think this ties in with the talks about surgeon training. Were the surgeons in this study all experienced spinal surgeons, from both the neurosurgical and orthopedic spine surgical fields? If so, was there anything in particular with respect to the types of patients or the types of surgeries that can explain the rather high incidence of dural tears?

DR. HALL: I will defer the experience of the surgeons to Dr. Brantigan and maybe to Mr. Christianson. But in terms of looking at the clinical study and looking at it statistically, there was one thing that was borne our statistically, that the prior surgery was a significant

factor in the dural tears.

I think, as you know, that is a significant factor. When you add doing a posterior interbody fusion on both sides to a prior surgical procedure, you are going to have an increased instance of dural tears with those two factors.

As far as comparing that to the literature, who do you compare that to? What study has that high a rate of prior surgery combined with doing the posterior interbody fusion.

DR. PERSENAIRE: I think I would also like to add, in this study, every tiniest little hole which occurred during, for example, scar removal was noted as a complication, even if it was primary repair and had no further clinical consequences.

I think it is very hard, also, to find any literature which has been as diligent as we have been in this study to note such complications. I think that is part of the explanation.

DR. BRANNIGAN: May I follow up on that? I think Dr. Glass said -- and probably misspoke -- that our control for the DDD group were all non-fusion surgeries. Actually, it was non-pedicle based surgeries. Most of them were fusion surgeries.

There was one that was not a fusion surgery. It was the O'Sullivan article treating current discectomy with repeat discectomy.

If you look at the complications in that study, they did not list the number of dural tears. They did list the number of neurologic complications.

Of them, foot drop, or nerve root damage, occurred in over five percent. That was double the rate in our study.

Their rate of repeat surgery for repair of dural tears was over five percent. That is double what was in our study.

They had one spinal cord injury, which was a complication we did not have in our study.

So, the dural tears that have no neurologic consequence, I feel, are not routinely reported in the literature.

The ones that are important are the ones that have neurologic consequence. Our numbers are one-half of the numbers reported in non-instrumentation, just repeat discectomy studies.

Based on that, I think we are within reasonable grounds for a study of a case group.

MR. BOYAN: Dr. Yasemsky, do you have any further

questions?

DR. YASZEMSKI: No, but I will make just one quick comment before I move on. With respect to the dural tears, you have already told me on the label insert that I read -- it may have not been the current one -- but it did say zero percent in the complications list on the one that I have. I just wanted to point that out, in section two, draft advice labeling. Thank you. I am done.

DR. PERSENAIRE: We have the latest version of the labeling here, and it says at the one month time point, it gives a dural tear rate of 1.7. That represents the patients that actually had subsequent manipulations done because of the dural tear. Those are the dural tears that we considered to be significant.

DR. CHENG: There were 44 patients with dural tears, were there not?

DR. PERSENAIRE: The total, yes. That includes those patients I just described as having a tiny nick during the removal of scar tissue, which was repaired right on the spot and did not have any clinical symptoms following the surgery.

DR. CHENG: So, the two percent is dural tears requiring a second operation.

DR. PERSENAIRE: Yes.

DR. CHENG: Not overall dural tears, which would be 44 out of the 219 or so.

DR. PERSENAIRE: Yes.

DR. WILKINSON: So you are making the assumption that a small dural tear that was sutured intraoperatively does not increase the risk of later arachnoiditis or other complication?

DR. PERSENAIRE: I do not make that assumption. I would refer rather to the clinicians who see the patients on a daily basis.

DR. BRANNIGAN: Of course, there is risk of a dural tear. We would rather not have them. We separately report the incidence of neurologic complications post op.

In those cases in which no neurologic compromise occurred, we feel that it is not a significant complication.

DR. WILKINSON: But arachnoiditis usually shows as pain and not as a deficit.

DR. PERSENAIRE: In our patient group, I agree with you, that it may lead to arachnoiditis. In this patient group, 77 percent had at least one previous surgery, and I think 25 percent had more than one previous surgery, it would be very hard for us to determine what the causative agent of arachnoiditis might be. I don't deny the possibility but I am not sure how we could identify that.

DR. BOYAN: Are you comfortable with the --

DR. WILKINSON: If they admit that it can cause damage, then I don't think you can not mention it as a complication. To just say zero percent dural tears is, I think, inaccurate.

DR. BOYAN: I think we can bring that up again in the questions. There are some specific questions about labeling that we can reinvestigate this at. Dr. Janoski?

DR. JANOSKY: I would like to visit one of the issues that I raised a little earlier. I have in front of me -- this is Clinical Data and Statistics, Volume E, Book 13, which is a series of outcomes from GEE models, looking at the effects of covariates on the outcomes, namely the three outcomes of clinical outcome, pain outcome and fusion.

As I look through here and as I listened to the discussion today, I see that a very profound, differential effect on all three of these outcomes, given potential covariates that might be going on.

Namely, one of the biggest predictors that seems to present itself in terms of outcome of either one of these is Workman's Compensation.

Issues of site, which is telling me that there are differential effects depending on who is doing the surgery, can you address this and try to allay some of my concerns as

to this overall success that we are seeing is for a very, very small population.

MS. QUINN: I am Linda Quinn. I am an independent statistical consultant. I have no financial interest in Acromed or the IF cage system.

I did perform the GEE analyses. I do want to remind the panel, though, that the primary goal of this study was to look at end point analysis. So, the covariate analysis is very much exploratory.

The covariates that were identified are what they are, and the sponsor has agreed to add them to the labeling.

DR. JANOSKY: If we think about going back to the pool-ability of these data, very few variables were examined to see whether they actually were pool-able or not, namely sites.

Where can I find data to show that the sites were comparable? Of concern, as site is popping up when I look at it in these longitudinal, is there a univariate analysis to show me that the sites are comparable, taking a look at the different surgeons and what are the outcomes for the types of patients that they are actually performing the surgery upon.

MS. QUINN: In the original PMA submission we did look in a univariate fashion at a whole range of demographic

type variables, including site, number of previous surgeries, number of levels treated, obesity, Worker's Compensation and legal.

In a univariate fashion, most of them did not come out significant. We did believe they were pool-able. There were some significant differences, and to control for that when we were comparing to the control articles, we adjusted the confidence intervals that were presented in our presentation for those significant covariates.

DR. JANOSKY: So, you can't say blanketly that in these tables that I have seen today in terms of overall success, that that success is indicative of any patient that walks in the door that was in the study?

MS. QUINN: No, I believe that the range of patients that we have seen yes, that those overall success criteria are indicative of any patient.

DR. JANOSKY: If I go again and look at these GEE models, seeing that site, preoperative is coming up numerous times, level of function preoperatively is coming up numerous times, body size as a marker of obesity is coming up numerous times, it is showing me that there are differential effects as the outcomes, based on those presenting characteristics.

MS. QUINN: That is correct. Once again, the

sponsor has agreed to identify those in the labeling. They do appear.

DR. LAURENCIN: I just have a few questions, first about the material itself. What was the compound being used or the material of the implant?

MR. CHRISTIANSON: The material is polyetherketone, ether ketone ketone. That is the generic name of the chemical chain. The fiber is a carbon fiber.

The material composition of the cage is 70 percent polymer, 30 percent fiber.

DR. LAURENCIN: That was switched, though; right?

MR. CHRISTIANSON: That is correct. In the IDE study we, in fact, had the reverse proportions, 30 percent polymer, 70 percent fiber, in the cages that we produced and were studied in the IDE.

DR. LAURENCIN: When I read this, it wasn't clear why the switch was made. There was one line that said it was cheaper and another line said it was better mechanics. What was the reason for the change?

MR. CHRISTIANSON: It is not a matter of cheapness. It is a matter of manufacturing efficiency. In the validation studies that we did, when we went to the material that is injection moldable, again, it is the same polymer and the same fiber, different proportions.

We did mechanical testing to see how it would perform compared with the other material. As it turns out, it was even stronger in bench test performance.

DR. LAURENCIN: Now, the studies in terms of the implants. How many implants were the 70/30 and how many implants were the 30/70?

MR. CHRISTIANSON: In the IDE, 100 percent of the implants were the 30 percent polymer, 70 percent carbon fiber.

DR. LAURENCIN: How many patients is that?

MR. CHRISTIANSON: Two hundred twenty-one patients.

DR. LAURENCIN: So, the implant going on the market would be?

MR. CHRISTIANSON: The reverse.

DR. LAURENCIN: The reverse?

MR. CHRISTIANSON: Right.

DR. LAURENCIN: How many patients have been implanted with that new design?

MR. CHRISTIANSON: That current polymer product we have been manufacturing for sale in the rest of the world since 1996. So, we have been selling that for approximately the last two years.

DR. LAURENCIN: So, the implant that you plan to

implant here in the United States, though, hasn't been implanted in any of these patients?

MR. CHRISTIANSON: In the IDE study under report today, that is correct.

DR. LAURENCIN: So, it is going to be a different implant than the IDE studied.

MR. CHRISTIANSON: We maintain that it is the same implant. It has the same geometry, same modiolus of elasticity. We do acknowledge that it has a different combination of polymer and fiber in it.

DR. LAURENCIN: It is processed differently, though?

MR. CHRISTIANSON: The manufacturing process is different as well, that is correct.

MR. BOYAN: Are you going to ask, Dr. Laurenson, if I can interject, did you do all the chemical studies on the new combination?

It was the IDE implant the preclinical data was performed on; is that not correct?

MR. CHRISTIANSON: The chemical analyses that we did before the IDE study began was on the pure polymer. FDA said they were interested in this polymer because it had not been submitted for implantation in humans before.

They wanted us to maximize the amount of potential

unknown material. So, we, in fact, tested the pure polymer.

DR. LAURENCIN: I think that is one correction to make. I think that in the presentation of the summary by the FDA personnel, it was said over and over again that the implant had undergone in vitro testing.

The implant has never undergone in vitro testing. The polymer has undergone in vitro testing, but not the implant itself.

Now, the carbon fibers that are part of the implant, how are they made? How are they processed? There are a couple of lines in here that say they are processes with aquillonitril? Is that true or how is it processed?

MR. CHRISTIANSON: I have to admit that I am not intimately familiar with the processing steps to make the carbon fiber. I would have to get that information and supply it later. We do not have anybody here who can address that particular issue.

DR. LAURENCIN: Why was the decision made not to perform any in vitro preclinical studies on the carbon fibers?

MR. CHRISTIANSON: At the time we submitted our test plan to the FDA, we presented them with a bibliography of previous use of carbon fiber, and the published information about its biocompatability.

At that time, FDA agreed with us, that the biocompatability was recognized for the carbon fiber itself. FDA never asked us to do that information, nor did we think it was necessary to do, based on the published information.

DR. LAURENCIN: Were these carbon fibers processed in the same way in the literature that you do for your carbon fibers?

MR. CHRISTIANSON: The literature that we submitted to FDA at the start of this IDE study had multiple manufacturing processes.

DR. LAURENCIN: Do you use one of those manufacturing processes?

MR. CHRISTIANSON: Yes, we do.

DR. LAURENCIN: Now, the implant itself, which implant was used in the study for the sheep?

MR. CHRISTIANSON: The goat study --

DR. LAURENCIN: Goat, sheep, it is close.

DR. BOYAN: Four legs.

MR. CHRISTIANSON: The implants used in the goat were the same implants as used in the human IDE study.

DR. LAURENCIN: So, in terms of implants in terms of this study, in terms of the animal study, in terms of the human study, it is going to be a different implant.

If approved, a different implant will be used

clinically, if this is approved, than is in those studies.

MR. CHRISTIANSON: Again, we would maintain from a biological point of view that they are the same. It is merely the proportion of the polymer to the fiber that has changed for manufacturability reasons.

DR. CHENG: Can I ask a question? I don't mean to interrupt you.

DR. LAURENCIN: Go ahead.

MR. BOYAN: Go ahead, Dr. Cheng.

DR. CHENG: You are no doubt aware of the experience with high density polyethylene and the subsequent experience to use carbon fiber in an attempt to improve the longevity and the wear characteristics of that polyethylene.

Clinical experience showed that the addition of carbon fiber to those implants for knee arthroplasties showed catastrophic early failure.

Knowing that, that the addition of carbon fiber to a polymer can result in major changes in the clinical result -- in this case, catastrophic failure -- how can you assume that the result won't be any different in yours, or at least perform tests to look into that issue?

MR. CHRISTIANSON: I used to work for the company that manufactured that particular material. I am not aware that there was a catastrophic failure because of the carbon

fiber reinforced polyethylene.

I am aware that it was found clinically unacceptable because, while the manufacturer did demonstrate that there was less wear, as a matter of fact, there was still some wear.

Because of the carbon fiber materials in that articulating surface, where the VK<y component was articulating against that tibial insert, the production of debris was highly pigmenting because of the nature of the carbon fibers.

I am not aware that there was any report of any adverse immune response to the particulates produced. I am aware that it very deeply stained the tissues and it was clinically unacceptable to see, in fact, black tissue from the pigmenting coming off that carbon fiber reinforced polyethylene surface.

We believe that we have a completely different situation here. We do not have any articulating surfaces against the fiber that is embedded in our polymer.

DR. LAURENCIN: Let me just ask. I think you briefly said that the implant that you propose to use clinically is actually being implanted in Europe right now?

MR. CHRISTIANSON: And the rest of the world; that is correct.

DR. LAURENCIN: How many years has that been?

MR. CHRISTIANSON: As I said, we began to use that material in 1996.

DR. LAURENCIN: How long have you been using the implant with the composition that you propose to use in the United States, how long have you been using that in Europe or elsewhere in the world.

MR. CHRISTIANSON: If I understand your question correctly, you asked, how long have we been using in Europe the implant that we studied in our IDE study?

DR. LAURENCIN: No, the implant that you propose to use here in the United States?

MR. CHRISTIANSON: Two years, since early 1996.

DR. LAURENCIN: Since 1996?

MR. CHRISTIANSON: That is correct.

DR. LAURENCIN: Do you have any clinical data on that implant? That is the implant that you propose to use here in the United States.

MR. CHRISTIANSON: No, we do not.

DR. BOYAN: I would like to make one comment, if I may, since it is on the same subject matter, and then go on to Dr. Wilkinson.

I was actually relieved that you reduced the carbon fiber to 30 percent, rather than them being

biologically equivalent. The combination is biologically very different.

The biology, in terms of an immune response, I think five years ago that is what we thought where debris elicited. It is five years later and we are starting to understanding the where and particles and particulates and carbon fibers have effects on cells in bone, environment, that are not technically immune responses but still clinically could be very important.

Not to take away from this, but those are not equivalent products. They have a potential to be biologically different.

There might be micromotion as the fibrous connective tissue will form at the interface. This is not an implant that has direct connection with the surrounding bone bed. It does have a fibrous capsule as most implants do. It is not unusual in that.

Just to put some biology in its place in this recorded proceedings, I think we do need to keep the biology of how the cells in bone respond to materials clear in our brains and be aware of it. We need to be aware of it.

DR. LAURENCIN: One final question. The necropsy results for the rats and sheep -- goats, whatever.

MR. CHRISTIANSON: The goats.

DR. LAURENCIN: I have sheep on my mind, but necropsy results, what were they like in terms of looking for carbon fibers in lymph nodes, spleen, et cetera?

MR. CHRISTIANSON: In the case of the rat study, again, that was pure polymer implanted in them, after you wanted us to maximize the amount of polymer exposed to the animals.

We harvested multiple sites in the rat -- brain, liver, lung, spleen, lymph nodes. No polymer or no effect of the material could be detected anywhere in the organs.

DR. LAURENCIN: I am most interested in the carbon fibers in terms of the goat study.

MR. CHRISTIANSON: In the goat study, we have the pathologist here who examined all those slides. He is prepared to show you some sections of lymph nodes that were harvested.

He explicitly looked for the particles in the filter organs and did not see it at any single time point in any organ tissues.

DR. LAURENCIN: Is that histology available?

MR. CHRISTIANSON: Yes, it is.

DR. LAURENCIN: Why were there only two goats at the two-year time point? It looks like you had five goats at every time point, but at the two year point you were down

to two.

MR. CHRISTIANSON: Yes, we were. The original protocol called for five animals with cages to be harvested at each time point, six months, one year and two years.

These animals, after their surgery was completed and they were healed, they were released back into their free-range environment. They were in a very large, corralled area.

Unfortunately, some of the goats succumbed to the environmental conditions. One was lost to a coyote. One got his horns stuck in a feeding trough and broke his neck. Unfortunately, we did not have five animals left to study at the five-year time point.

DR. WILKINSON: I have a few questions, but more comments for which I don't really want an answer. Is that appropriate to do now?

DR. BOYAN: Absolutely. This is our opportunity. This is our open discussion.

DR. WILKINSON: All right. What I am hearing is that we are presented with a device which is really a composite of two devices.

The manufacturer, to satisfy me, needs to show that the combination is significantly more effective than any device alone, and at least comparably safe.

That information was not presented to me as a reviewer prior to my coming here. This is relevant also for the biomechanical data and for the animal data.

I have real concerns about the clinical inclusion criteria. For instance, some patients were operated for leg pain only, and most of us realize that there are usually easier ways to cure leg pain.

Surgery was done for failed fusion, but we are not told how that failure was defined and confirmed.

If that fusion was multiple levels and only one level failed, we are not told how many levels were reoperated, and whether it influenced outcome to have one level refused, three levels, two levels, one additional level already fused.

The clinically significant -- all clinically significant motion segment levels were fused, we are told, but we are not told how this was defined, especially in the group of patients with multiple level degenerative spine disease.

All of us know that is a factor of aging more than disease, and if a patient is aging in their spine, they will have multiple level hypertrophic degenerative disease without necessarily having multiple levels of clinical significance.

Then the criteria for degenerative spondylosis, also, I agree with Dr. Grobler, there is a tremendous difference between degenerative spondylosis and ismic, if it is congenital spondylosis.

The data for clinical effectiveness analysis is not broken out by number of levels fused, one, two or three, at least not that I could easily find in this set of documents.

What we are dealing with, you see, is a very heterogeneous population, many different indications, many variables that Dr. Janoski has already referred to.

The fusion analysis did not include dynamic radiographs, and I agree. While the steel construct is stable, the bending film is less significant.

Certainly by two years, when the pedicle screws could have loosened, we are after all looking at two different devices.

We need to know, is there motion around the pedicle screws as well as motion in the interbody, which can conceivably happen even though the steel doesn't break.

I am not impressed, simply, that bone bridging across a fusion area is sufficient for stability. That bone must be stably attached at both ends.

Small areas of attachment may provide only weak

fusion strength, and we are not told how, with plane films only, the area of actual bony fusion at both ends was determined.

We aren't told, for instance, how the increased density of fusion bone was quantitated, and whether there is a difference in density between allograft bone and autograft bone.

We are not told if fusions were rated as stable, even if their VSP steel components were broken loose or migrated.

This is another important factor -- it comes up occasionally in the complication table, but this is not included in the seven point scale for fusion stability.

We are told that disc spaces don't collapse. Well, if you have a stable VSP device, that doesn't prove that the Brantigan cage is stable. The VSP may be holding the disc space open and have total non-union of the Brantigan device. Collapse alone isn't enough.

Pain analysis, individual patient outcomes are certainly much more important to me than composite pain scores.

Only five patients in the entire series preoperatively were rated as having mild or no pain. It is unclear to me why those patients were subjected to such

extensive surgery, another variable in the patient selection.

All patients improving by at least one point between their preoperative and post-operative pain at 24 months were considered successes.

I would like to know how many improved only to a level of two which is severe pain, or even a level three which is still rated as moderate pain.

Patients who remain in severe or moderate pain would certainly not be very convincing successes. Yet, under the criteria used here, they could be rated as successes.

We aren't told about pain outcome in those patients operated for leg pain only. There were relatively few, fortunately, but I think it is important also to know, are we dealing with leg pain or sciatic pain.

The neurologist who made her presentation earlier today is acutely aware that sciatic pain --

DR. BOYAN: Wait. Was that during the last couple of hours?

DR. WILKINSON: Yes, Dr. Glass.

DR. BOYAN: Okay.

DR. WILKINSON: All right. I heard earlier today and it was sounding 3:00-ish, 2:00-ish; earlier this

evening.

DR. WILKINSON: Earlier tonight. She is certainly aware that there is a difference between the pain induced by stimulating a sciatic nerve and the pain induced by stimulating a thecate joint, both of which are felt within the leg.

Functional analysis, also I am more interested in how individual patients improved than composite scores. How many improved by one step, two steps or more.

If we look at the charts presented in the summary of safety and effectiveness, showing preoperative and 24-month functional rates, those results could have been achieved if 40 to 50 percent of the patients had attained no functional improvement.

That is all right if those patients were functioning beautifully prior to surgery, but then why did they need surgery.

For a patient not to improve functionally, I think, is a real deficit.

Twenty-two patients of the 177 evaluated at 24 months, or 13 percent of the entire group, were listed as being still totally incapacitated or unable to participate in activities outside the home; certainly not a success in my book.

We are not told how many others lost functional capacity, at least in the material.

Most similar studies include a return to work or usual occupation category, but we aren't told whether that happened here.

DR. BOYAN: Dr. Wilkinson, are there more pages?

DR. WILKINSON: Only two.

DR. BOYAN: You promised me a few comments. Could you hit some highlights here?

DR. WILKINSON: Let me hit a few highlights and then I will let you read the total list of comments here later on.

The FDA amended facts dated 9-23-97 refers to seven deaths in the study, but none are reported in the table of complications and the summary of safety and effectiveness.

That table, incidentally, is very poorly organized.

Finally, the intraoperative blood loss, averaging over 1500 milliliters -- that is three units of blood per patient, average, for the entire series -- means that one-half of the patients lost more than three units of blood during their operation.

One patient died of interoperative blood loss, we

are told, but that doesn't appear in the complication table.

Is this comparable with other series? I happen to use an interbody pliff technique with, unfortunately, Dr. Brannigan, better results than what you referred to.

In a series of 65 patients, only two of my patients required any blood transfusions. I think this is a serious complication, excessive blood loss, a serious restriction of the use of the device, at least. I will quit.

DR. BOYAN: I would like to give Acromed an opportunity to respond. Maybe they could identify a few things that they would like to respond to.

DR. PERSENAIRE: I have a very quick one on the VSP keeping the disc space open. There is an article, again by O'Brien, Holty, published I think in the European Spine Journal as well as another publication, which shows rings with VSP still losing their disc space height. I don't think the VSP device alone is capable of keeping the disc space open.

DR. WILKINSON: Which further proves that the VSP device does loosen.

DR. PERSENAIRE: I refer to Dr. Hall to answer that question.

DR. WILKINSON: You just told me it did.

DR. BRANNIGAN: I would like to address some of these many points. Of course, they do touch some important areas of concern.

I have been writing very furiously as you talked and I hope I can touch on a number of them.

First off, the interoperative blood loss averaging 1600 ccs, certainly that is an important consideration. I would like to point out, though, that the severity of this mix of patients was much greater than on many other studies.

We had every one of our DDD patients who had at least one prior failed surgery, on average two prior failed surgeries.

There were multiple levels. For a proper perspective for a one-level pliff, I would like to point out that our cases of spondylolisthesis at one level with no prior failed surgery averaged 575 ccs of blood loss.

I think that the blood loss is determined quite a bit according to how many levels a surgeon does and how many prior surgeries.

Certainly I would have a very strong warning to any beginning surgeon not to attempt a three level cage pliff without a great deal of experience, because the blood loss will be a lot more with an uneducated surgeon.

A second point, you mentioned the pedicle screws

can loosen. The loss of disc space height, we have observed that numerous times around a donor bone pliff.

With the collapse of that bone there is loss of the disc space height, in spite of the pedicle screws. On numerous occasions we have been able to reconstruct that anteriorly and stretch that height back out without actually having loosening of the pedicle screws.

Certainly they sometimes loosen, but I would like to quote a Larson study in 1996 in the Journal of Spinal Disorders in which he used flexion extension X-rays with the pedicle screw construct.

His conclusion was that in the presence of pedicle screws, the flexion extension X-rays did not help in elucidating those cases that were successful or failed fusion.

DR. WILKINSON: Even at two years?

DR. BRANNIGAN: I believe it was two years. Larson, 1996.

DR. WILKINSON: I don't recall he had the time in there.

DR. BRANNIGAN: You asked about the need for data for the combination of devices. I would like to refer to Jack Metcalf, slide labeled A-JB-01.

This emphasizes the success rate of fusions by

numbers of levels treated, from one level, two levels and three levels, as reported by every author that we have been able to see over the last 20 years.

As you see, one level fusions, by the literature, averaged a claim of 91 percent fusion. Two levels, by the literature average 76 percent fusion.

DR. WILKINSON: Fusion by what means?

DR. BRANNIGAN: This is what was reported by these authors.

DR. WILKINSON: My question was fusion with a VSP.

DR. BRANNIGAN: If you will let me finish, I will get to the conclusion.

DR. WILKINSON: Good.

DR. BRANNIGAN: The three levels, on average, the literature claimed a 65 percent fusion rate. Now, in comparison to what we are doing with the cage and VSP, at one level we have 99 percent fusion success, at two levels 100 percent fusion success, and at three levels, 11 out of 12, or 92 percent fusion success.

The very definite advantage of combining these devices is to gain fusion success independent of the number of levels treated.

I believe that there has never been a study without pedicle screw instrumentation that has had fusion

success independent of the number of levels from one to three.

Whereas, I am sure that there are certain circumstances where caged fusion without pedicle screws of whatever type would be worthwhile at one level, I think the treatment at two levels or three levels absolutely requires the combination.

As you can see, this is a published report of the previous approved system. Can I say that?

DR. BOYAN: Yes, you can.

DR. BRANNIGAN: The BAK published report, although it had 94 percent fusion at one level, this fell out to 71 percent at two levels.

So, this shows the definite cases in which I believe that there is statistical proof that the combination of devices gains significant advantage.

Another point on this, slide number 21-IDE-PPT-21-J.

DR. WILKINSON: On this slide, you are not implying that you don't intend to market your device for single level use?

DR. BRANNIGAN: I think that single level is an area that we have not fully compared the success rate of cage alone or cage plus VSP.

However, we have data that we believe shows that we had better results in our study than other cage studies without pedicle screws.

Therefore, I think this is relevant to do this at one level, although we certainly believe that surgeons should have the choice of the different systems.

DR. WILKINSON: It would have been nice had the panel had the advantage of that data also. I didn't have it.

DR. BRANNIGAN: To follow up, we have our success rates by one, two, three and four levels treated, for all of these parameters.

Fusion success, as we have said, is consistent across the levels. Pain also was consistent up to four levels had less pain relief. However, that was just two patients.

Function was consistent across the levels, and the overall result at four levels was less, because that was only two patients. One of two had a good result.

So, up to three levels treated we had very consistent quality of fusion and clinical results.

DR. BOYAN: Dr. Wilkinson, would you allow me to move on to Dr. Hollman, and then we can come back if we still have some issues? It is now a quarter to 10:00 and we

still have to go through the questions.

DR. HOLEMAN: I have no questions.

DR. BOYAN: Nothing. Okay, and Dr. Zokitas?

MR. SILKAITIS: In the interest of time, I have no comments.

DR. BOYAN: Okay, are there any other issues that need to be brought forward now that will not be covered under the panel questions, the questions directed to the panel?

DR. GROBLER: Yes, I have one question. We discussed the issue of the cage, the substance has somewhat changed.

I recall that there was also a question of a fifth generation screw implant, the pedicle screw. Is that true?

MR. CHRISTIANSON: Yes.

DR. GROBLER: Can I go through the same questions we went through before? When did that come on the scene and how many of the study was done with that specific screw?

MR. CHRISTIANSON: One hundred percent of the patients treated in the IDE study were treated with what we call the fourth generation screw.

You will recall that I mentioned previously that this product is currently available under a 510(k) in the United States labeled for severe spondylolisthesis.

We presented the mechanical test comparing the fourth generation screw to the fifth generation screw to FDA in our 510(k) application.

As a matter of fact, we only market today, under the 510(k) the fifth generation product.

It is logical for us that we seek approval for the device that we intend to market with the Brantigan cage today. Therefore, we are seeking approval for the fifth generation product.

DR. CHENG: I have a question. You just presented some information as to what the data is stratified by level fused.

The FDA has asked us as a panel whether or not the clinical data for fusion of greater than two levels is supportive of the safety and effectiveness of this device.

Your stratification and breakdown of the data, was that provided to the FDA?

MR. CHRISTIANSON: Yes, it was.

DR. CHENG: Then why are we being asked that question, or has some analysis been done of the data just based on three and four level fusions, because that would answer that question, or have you done it yourself?

MR. CHRISTIANSON: We have presented the data to FDA and we, in our draft labeling, are seeking approval for

treatment of one, two or three panels. It is the FDA that is asking the panel if that is appropriate.

DR. CHENG: There are, as I recall, 14 patients or perhaps 12?

MR. CHRISTIANSON: Jack, can we have slide 21 back, please?

DR. CHENG: Approximately 95 of your patient series involves two-level fusions and there are 14 patients with three or four level fusions.

Do your data hold up for those three or four level fusions? I haven't seen that data presented.

MR. CHRISTIANSON: As Dr. Brannigan just discussed, you see here that there are 12 patients here with three level fusion, two with four level fusion.

The balance was either one or two level fusions. I believe we can review these data with you again if you want to be thinking about that question now.

DR. CHENG: What I meant by holding up, I meant under some statistical analysis such as has been done with the whole study. As you can imagine, 14 patients is rather a small number to make an assessment.

DR. BRANNIGAN: May I follow up? These three levels were written into our IDE program with very good statistical justification.

That was based on the IDE study done for the pedicle screw VSP system approximately 10 years ago. As you know, that system never came to panel.

With the analysis of that, there were many patients that were done over three, four and five levels. When we looked at that data, the cut off of pool-able data was one, two and three levels fused, and four and five levels under that study had a significantly decreased success rate.

That is why we labeled three levels as the cut off in this study, and I believe our data supports that, even though the number of three levels is small.

DR. CHENG: Is the number of dural tears higher with the number of levels fused?

DR. PERSENAIRE: We did not do that particular analysis, but we would be happy to do it.

DR. CHENG: Is the amount of blood loss greater with the number of levels fused?

DR. PERSENAIRE: Yes, it is.

DR. CHENG: So, the safety in terms of a greater number of levels fused may be different than the other two levels of fusion.

I am wondering whether or not the safety and effectiveness of the device for three and four level fusions

remains the same as for one and two level fusions.

DR. BOYAN: I think that is a perfect lead-in to the panel questions. That is absolutely a smooth transition. I would like to thank the representatives of Acromed and thanks the FDA reviewers and the lead reviewers.

Now, I would like to ask you to come back up and lead us through the questions.

In the interests of time, because there are lots of questions, I am not sure that -- we the panel all have the questions in front of us.

The questions are going to go up onto the wall. So, maybe you could read maybe the bottom line of the question.

MS. MORRIS: Actually, I have a short version on the screen.

DR. BOYAN: She was thinking right with me; great. Short version.

MS. MORRIS: I will let you go ahead, or would you like me to read what is on the screen.

DR. BOYAN: Read what is on the screen.

MS. MORRIS: Based on the sponsor's assessment scale, what is the minimum category that should be used in determining success in fusion.

DR. BOYAN: We will start with Dr. Zokitas and

come this direction.

DR. SILKAITIS: In terms of a minimum category, I would recommend the panel to assign the category that is typically used in clinical practice, as opposed to the most ideal.

I know bone takes seven years to totally heal and lines totally dissipate. In clinical practice, there is maybe a more liberal policy in terms of that.

DR. HOLEMAN: I am going to yield to a medical decision on that.

DR. WILKINSON: I would suggest category seven plus 24-month flexion extension films -- flexion extension films at 24 months.

DR. LAURENCIN: I just suggest category seven.

DR. JANOSKY: Pass.

DR. YASZEMSKI: I think category seven would be sufficient.

DR. HALE: My concern is that there is no measure of motion included anywhere in this category.

DR. GROBLER: Category seven.

DR. KERRIGAN: I think some motion assessment should be performed.

DR. CHENG: I think the clinical standard would be lack of motion on flexion extension films in category seven.

MS. MORRIS: Thank you. Did we adequately address that question for you?

DR. BOYAN: I believe so.

MS. MORRIS: Okay, next question. Should the assessment of pain be based on the data collected on overall pain, or should pain be assessed at a specific site of the original complaint of pain.

DR. BOYAN: Let's start with Dr. Yasemsky.

DR. YASZEMSKI: I think the patient came in because of the original pain and I would suggest that the original pain be in the category of assessment.

DR. HALE: Is it unreasonable to think that you could do the assessment based on both scales?

DR. BOYAN: Both scales meaning back and leg?

DR. HALE: Meaning overall as well as specific site. Is there a need to limit it to only one or the other?

DR. BOYAN: I would leave that up to you to comment on.

DR. GROBLER: The preoperative pain area.

DR. KERRIGAN: To go along with Dr. Hale, I believe that both should be evaluated, both overall and the specific.

Again, it just gets back to what are the long term consequences, things we don't know. Is there some new pain

that we need to know about?

DR. CHENG: I would state only the original site of pain.

DR. SILKAITIS: No comment.

DR. HOLEMAN: I would say if the initial assessment was made, your baseline data was based on the complaint of pain prior to surgery, then I think the data collected to determine success should be based on that. I would say both back and leg.

DR. WILKINSON: I think we are dealing both with safety and effectiveness. If a patient has developed pain in the back that they didn't have prior to surgery, that is a comment on safety. I think both criteria need to be included.

DR. LAURENCIN: I think it would be good if you had both criteria included.

DR. JANOSKY: Both criteria, but I would put them in a hierarchy and would say that the specific site would be the primary and the overall would be the secondary.

DR. BOYAN: Could I get some clarification? On the decision for both, are we talking about both leg and back, or are we talking about the site plus overall?

DR. WILKINSON: I referred to overall pain plus site specific.

DR. BOYAN: That is what I heard. Is there anybody who has a different opinion? Okay, thank you. Question number three.

MS. MORRIS: 2-B. What is the best method of assessing patient pain to determine clinical success?

DR. BOYAN: I guess the fastest way is to just keep going around and then we get it all out in a hurry. Let's start with Dr. Wilkinson.

DR. WILKINSON: The beset method is both of the above or the same answer to A, plus patient satisfaction.

MS. MORRIS: So, you mean to use the five-point scale that they have already presented in addition to patient satisfaction.

DR. WILKINSON: A five-point scale, and you have other questions about this later on. The five-point scale should show a pain response in the none to mild category or at least two-level improvement to be satisfactory pain improvement. Then there should be a patient satisfaction score as well.

MS. MORRIS: So, we are talking two scales?

DR. WILKINSON: Yes, two scales per patient. That is, not statistical composites.

DR. HOLEMAN: I really do think that the instrument -- that is what we refer to -- the instrument to

assess pain should be one that would allow the investigator to determine whether the pain is related to function in addition to just pain in general.

Pain can be aggravated by the activity engaged in at the time. I would think that the pain should be assessed on activity, function and site.

MS. MORRIS: Would these be separate scales or a composite?

DR. HOLEMAN: Composite.

MS. MORRIS: Dr. Zokitas, do you have a comment?

DR. SILKAITIS: Are we talking about this specific PMA data or are we talking about in general?

MS. MORRIS: This is specific to this PMA.

DR. SILKAITIS: So, this is the data that the company has?

DR. BOYAN: The data that the company -- the sponsors -- presented was the pain score of a five-point scale for the overall pain.

DR. SILKAITIS: These three questions refer to how the data is going to be combined?

MS. MORRIS: Yes, what would be the best method of assessment for this device.

DR. SILKAITIS: In that regard, are there validated ways to combine these data? It is nice to say we

should do this or do that, but are there validated studies that show that this is going to tell us something?

DR. BOYAN: I think that I can't answer that specifically. I think what they are asking of us would to state what we would feel would be the best method of assessing patient pain to determine clinical success.

You are saying that you want to be sure that we are not inventing a new method at this point in the game. I think that is a fair comment. Did I articulate what you were trying to say?

DR. SILKAITIS: Exactly.

DR. BOYAN: Can we go on around to Dr. Cheng?

DR. CHENG: The appropriate measurement would be an outcomes measurement which is validated for the patient population being studied.

As Dr. Hollman indicated, I do not know whether or not the instrument by the North American Spine Society has been validated for this patient population.

A pain scale like the one used here doesn't take into account medications taken as analgesics and whether or not reductions have been accomplished and other functional status measures.

I am sorry, there are function status measures measured there.

DR. KERRIGAN: I have no more comments to add than what has already been stated.

DR. GROBLER: I just had a question. Did the patient produce the pain score or was this a clinician extracted score? Was this a self-assessment score?

MR. CHRISTIANSON: The case report form had that five-point scale on it.

DR. GROBLER: Who filled the form out?

MR. CHRISTIANSON: Either the clinician or a data manager at the clinician site.

DR. GROBLER: So, that was not an independent person that filled it out?

MR. CHRISTIANSON: In the case of a data manager, that would be.

DR. GROBLER: How many of those were data managers and how many were the clinician. Do we know?

MR. CHRISTIANSON: We will see if we can find it for you.

DR. GROBLER: I think there should be an independent scale. It should be a self assessment and I think a pain score from an established questionnaire such as the NAS battery of questions or the SF-36 would be the preferential way to go as part of a subscale where it is a self assessment from the patient. This was not done and I

would like to see that.

DR. HALE: I have nothing to add to what has already been said.

DR. YASZEMSKI: I think when we are talking to the patient about doing surgery we make a risk/benefit analysis. I think that is the kind of thing to do in retrospect.

I believe I saw Dr. Hall actually present that. Brad, correct me if I am wrong, but there was a scale where the patients said, my result is such that, knowing what I went through, I would do it again. I would take something like that as a good indicator on a pain scale.

DR. JANOSKY: Nothing to add.

DR. LAURENCIN: Nothing to add.

DR. BOYAN: Okay, have we handled that one to your satisfaction?

MS. MORRIS: I Believe so.

DR. WITTEN: I would just like to clarify that the question as you mentioned relates to how the data collected in this PMA could be used to determine clinical success, and not what other scale could be used for measurement.

I don't know whether anybody has any additional comments about how to use the data that we have before us to determine success.

DR. YASZEMSKI: We still have question 2-C and I

think that is where your question comes.

DR. BOYAN: Let's do question 2-C.

MS. MORRIS: What amount of improvement in pain would warrant a success for this outcome parameter.

DR. WILKINSON: As I said just a few minutes ago, I think using the scale that was used -- that is, using the data that is available in this study -- there should be a demonstration of minimal or no pain or at least a two out of five reduction in overall pain.

DR. BOYAN: I guess we will just keep going.

DR. HOLEMAN: Basically a reduction in overall pain. I would not so far as to indicate a number to say statistically significant.

DR. SILKAITIS: I have nothing to add.

DR. CHENG: I don't have anything to add either.

DR. KERRIGAN: I agree with Dr. Wilkinson. That is what I was getting to in my comments. I think that a one point isn't sufficient, but a two-point improvement would demonstrate some improvement.

DR. GROBLER: I have nothing to add.

DR. HALE: Nothing to add.

DR. YASZEMSKI: I am just going to reiterate my last comment, that if the patient feels the level of pain relief they got was worth what they had to go through, that

is success, regardless of what the absolute level is, or the number of levels they jump.

DR. JANOSKY: One point reduction in pain, but calling to mind that there were some patients saying they had no pain going into the study.

That is a consideration for what you are going to do with them. It is impossible for them to show a one point reduction.

DR. LAURENCIN: Nothing else to add.

MS. MORRIS: 3-A. Do you consider the removal of a VSP component as an expected surgical procedure?

DR. BOYAN: Could we combine, I think, A and B?

MS. MORRIS: Sure. Actually, we could combine all three.

DR. BOYAN: Right.

MS. MORRIS: Do you consider a second surgical procedure or do you consider a recommended second surgical procedure for the indicated patient population for the PMA.

Then the third part was, when should the removal of the VSP components be considered a treatment failure.

DR. BOYAN: Let's start with Dr. Grobler.

DR. GROBLER: I think the study set out to remove the instrumentation. I think if you enroll patients in a study, it should go all the way. That would have answered

our fusion question as well.

I think at this time, as part of the study I think it should be done. As part of the long-term, I think it would depend if the implant has any untoward effects.

Do you consider as a recommended surgical procedure? As part of the study, yes. In the long-term it is probably not.

Thirdly, when should the removal be considered a device failure. If it is removed for another reason than electively, having decided beforehand that it should be removed.

DR. WITTEN: I would just like to clarify 3-C, and 3-C relates to how we are going to look at the data after this discussion, which is when should the removal of VSP components be considered as a treatment failure.

I guess this question would have been clearer if it said as a treatment failure in the study.

MS. MORRIS: I had mentioned earlier that I made a mistake. I changed the word from device to treatment. It was an error on my part.

DR. BOYAN: Does everybody understand Dr. Witten's clarification? We are limiting our discussion to this specific study, for 3-C.

DR. GROBLER: Did I understand correctly, then?

This is considered a treatment failure?

MS. MORRIS: If the VSP components were removed, when would you consider that a treatment failure?

DR. GROBLER: My answer would stay the same.

DR. HALE: In response to the first two questions, I would say that if that information was presented to the patient, that this was an expected or potentially expected surgical procedure, then I would say yes.

With regard to the third question, among many other possible reasons for considering it a failure, I would list specifically device failures, if there is any deformation of breakage of the screw or plate system, or loosening of the mechanical interface between those two devices and the biological tissues.

DR. YASZEMSKI: With respect to the first question, removal is part of the study, so I would consider it to be expected.

With respect to the second question, notwithstanding the recommendations to remove fracture fixation devices, I would hesitate to recommend that to my patients routinely who have fusions. I would not let them know that it would be expected I would be removing the spinal instrumentation.

With respect to the third question, I would say it

was a treatment failure if the patient, having continued systems, got another operation for those symptoms and part of that operation included removal, be it because that instrumentation was broken or because it was taken out and another fusion attempted, then I would consider that a treatment failure.

DR. LAURENCIN: Just a question. In terms of the other countries where these studies are going on, what percentage have hardware removed?

DR. PERSENAIRE: In preparation for this meeting, I have, indeed, done a kind of poll throughout Europe and have gotten the information on roughly 2,500, 3,000 patients treated by different physicians. The series varied from 25 to over 600.

I found that there is one surgeon in Amsterdam, in fact, at Free University who routinely, one year out from the surgery, removes the plates and screws, inspects the fusion and then closes the patient, being over and done with.

The reason that he does that is he mainly uses this in relatively young patients with lytic spondylolisthesis.

If I look at some other sites, I have one site in the United Kingdom where out of 145 patients there has been

one reoperation where the VSP has been removed.

So, it varies very widely from a couple of percentage point to 100 percent.

DR. LAURENCIN: Have those 145 patients been followed clinically for outcomes?

DR. PERSENAIRE: They were not part of an official study, but certainly in England and even moreso in the Netherlands, the common medical practice is that a patient continues to see the same physician. There is less shopping around than you see in some southern European countries and I have witnessed, to some extent, here in the United States.

I would say that the data from the Netherlands are reliable as to the actual return of patients.

DR. LAURENCIN: The first question is whether this is something that is necessary or not. If there is other experience that says it can stay in, then it wouldn't be an expected procedure.

It seems that half of these that were removed were removed for some other reason than planned removal.

On the statistics they have elective partial removal 15.8 percent and partial removal 20.8 percent, which sounds as if they are making that distinction, that half of them were removed because of some problem like pain.

MR. CHRISTIANSON: The data to which you are

referring in the complications table from our originally submitted summary of safety and effectiveness, FDA asked us a very similar question.

The case report form listed a number of for-cause reasons for the surgeon to go in to remove hardware, such as broken screws, which some other panel members have currently listed.

We have reported those. There was a blank line called other, that the surgeon would fill in with words of his own choice.

Those words varied widely from elective removal to other reasons. We found it not possible to systematically present those data in a meaningful way because of the open-ended nature of the way the surgeons recorded those reasons for removal.

DR. BOYAN: I think that is the purpose of the question from FDA to us, is the very issue you raised, Dr. Laurenson. Do you have any other comment you want to make about this?

DR. LAURENCIN: No.

DR. WILKINSON: As far as the study up to this point, removal was part of the study and so I think that it should be considered a planned part of the study and not a complication necessarily.

From this point on, however, my own practice is not to recommend removal of the hardware. I would hope that the study could be amended so that it is not an obligation to remove the hardware.

For item C, if the components are removed because of device failure or if, when the device is removed, the patient's pain and function is worse, then I think we are dealing with a treatment failure.

DR. HOLEMAN: Nothing.

DR. SILKAITIS: Nothing.

DR. CHENG: A question to Dr. Brannigan. Would you have this device removed from your spine of you were asymptomatic and it was implanted in your back?

DR. BRANNIGAN: If I was asymptomatic?

DR. CHENG: Right.

DR. BRANNIGAN: No.

DR. CHENG: So, why did you do that to your patients?

DR. BRANNIGAN: Because of the FDA recommendations.

DR. CHENG: So, you let the FDA convince you to do an operation you didn't feel was necessary?

DR. BRANNIGAN: I explained to my patients what the recommendation was, and that in my experience and in

hardware in general orthopedic conditions, that the recommendation was very clear that these things be removed.

I have had many experiences in general orthopedic implants that a patient told me they had no pain whatsoever, and then when I took the implants out they came back and said, I feel so much better, doctor.

There are some subtleties to that, that I think should be decided between the patient and the doctor. I have never insisted that anyone have their hardware removed, but I explained the recommendations and the reasons for it, and about a third of my patients elected to have the patient out.

DR. CHENG: When you tested for motion, what did you do when you found motion present? Did you still take it out or did you tighten up the screws and let it in?

DR. BRANNIGAN: I think that is a very important question. I think the proper answer is, as Dr. Wilkinson said, if you take that hardware out and it is not fused, that patient will have much more severe pain.

So, I think the answer is, if there is motion -- what I always did was to reinforce the fusion and put hardware back in.

DR. CHENG: So, there were times where you planned to take it out but instead, you put it back in.

DR. BRANNIGAN: I had that in one case and that was a posterior lateral fusion above the cage fusion. The study allowed us, at a degenerated disc level, to extend the fusion with the posterior lateral fusion one level above.

In this case, the cage fusion was solid, the posterior lateral fusion was not solid. So, I reinforced the posterior lateral infusion with bone grafts and reinserted the screws and plates.

DR. CHENG: While my personal feeling is that I don't feel it is a necessary thing to do all the time and, if it is part of the protocol, then the results should be measured after the removal of the device.

DR. KERRIGAN: No comment.

MS. MORRIS: Question four. Is the overall success definition used by the sponsor appropriate, or would you recommend any modifications?

DR. BOYAN: Let's start with Dr. Janoski.

DR. JANOSKY: Based upon the discussion that we had just had for the first three, I would suggest a revision of the overall success to the success in fusion as defined by category seven, and improvement in specific site pain and essentially leave the rest as they are.

DR. LAURENCIN: Why don't you skip me for this one.

DR. WILKINSON: I would like to see the overall success criteria changed so that success in fusion is defined by category seven only, improvement in overall pain be matched in improvement in patient satisfaction, that simple maintenance of function is not enough unless the patient had normal function prior to the surgery, and then they shouldn't have had surgery, and that no further surgery related to revision or removal of the cage or problems with the VSP device.

DR. BOYAN: Dr. Wilkinson, I am shocked. You forgot the flexion and tension X-rays.

DR. WILKINSON: Thank you, madame.

DR. HOLEMAN: I think I will go with what they have here, but where it says improvement in overall pain, I will go with that.

Maintenance in function, I would say that a patient who has elected to have surgery had minimum function before going into surgery. I would think there should be an increase in overall function after surgery.

I would have to agree with Dr. Wilkinson that there should be improvement, increased function as a result of surgery as opposed to just maintenance of what was prior to surgery.

Also, patient satisfaction. I think overall what

the patient feels or what he thinks about the overall surgery would be much more indicative of what he perceived as success, what he or she.

DR. CHENG: The only modifications would be level seven only. The X-rays, improvement in function Dr. Hollman mentioned, and the requirement that the questionnaire be self administered and not clinician initiated.

DR. KERRIGAN: I agree with the last three reviewers.

DR. GROBLER: Category seven improvement and back and leg pain, function, muscle strength complication the same, patient initiated outcome, and then include VSP, I think both should be assessed.

DR. HALE: Nothing additional.

DR. YASZEMSKI: I am going to agree with everything that has been said, specifically that a category seven. I also think that to be certain that apples and oranges aren't being mixed here, I would really like to see, of the 221 patients, how many had posterior fusions, so we are not comparing 360s with anteriors, and report it that way.

DR. LAURENCIN: Nothing to add.

DR. CHENG: Was the sponsor not going to respond to Dr. Yesemsky's request about the posterior fusion?

DR. BOYAN: We can have him respond if you would like to. Would someone like to respond to that statement? Dr. Yasemsky, would you repeat it?

DR. YASZEMSKI: Some of the fusions were anterior and posterior. It would be nice to know how many were anterior only and how many were anterior and posterior, and then present the results stratified by those two groups.

DR. BOYAN: Mr. Christianson, do you have the answer?

MR. CHRISTIANSON: We don't have the answer with us here. The monitor who is intimately familiar with this says she believes that it was fewer than five percent who didn't also have a posterior lateral fusion. We can stratify the data that way.

MR. YASZEMSKI: I would just add to that, when that is done, it would also be useful, for instance, for those who had their instrumentation removed and their fusion explored, was the posterior fusion explored manually or, as I have heard, force apply or moment applied to the functional spinal unit and an assessment of motion based on either tactile sensation or a visual inspection made.

It would really help me to understand whether this was perhaps the posterior fusion was solid and the anterior one wasn't.

For instance, did an inspection visually of the posterior fusion and manual palpation of the posterior fusion reveal solid fusion, but perhaps this wasn't a radiographic seven on the radiographic scoring. It would be nice to see how those two mix.

MR. CHENG: Could I clarify? Did I understand you correctly, 95 percent of the patients had both an anterior and a posterior lateral fusion? Is that what you are telling us?

MR. CHRISTIANSON: That is what the study monitor whispered in my ear, yes. That is from being familiar with the data in general. We have the data back at the office. We just can't produce that number for you here today.

DR. CHENG: I don't recall that being mentioned in the methodology or the protocol in terms of doing a posterior lateral fusion.

DR. YASZEMSKI: You are right.

DR. CHENG: This is a completely different operation you did than you are reporting, then, in 95 percent of the cases. Am I right?

MR. CHRISTIANSON: I am not sure I understand your question, Dr. Cheng.

DR. CHENG: You are presenting the safety and effectiveness of an anterior spinal fusion with

instrumentation posteriorly, not a fusion posteriorly. You just told me that 95 percent of the patients had a fusion posteriorly.

MR. CHRISTIANSON: The protocol obviously called for an interbody fusion using either the Brantigan IF cage or the donor bone in the cases of the randomized cases.

All patients received the VSP plates and screws, and the posterior lateral fusion was at the surgeon's discretion.

It said that in the protocol, so that was a standard part of the protocol. All of the fusion results that we have been giving you here today have all been on the interbody space.

DR. CHENG: So, the outcome assessment of the interbody space pertaining to the cage, which is what you are proposing today, can only be assessed by the radiograph.

MR. CHRISTIANSON: Correct.

DR. CHENG: Not by the patient's or clinician's pain measurement or function or any of the other outcome measurements you have listed here.

MR. CHRISTIANSON: Again, I am sorry, I don't understand your question. The fusion results reported in our study are for the interbody fusion level, and all levels had to be evaluated as fused for the patient to be called

fused.

DR. BOYAN: May I take the chairman's prerogative of stating it another way, and then you can tell me if I am getting it right.

The issue at hand here is that many of these other measurements of patient satisfaction, of overall -- many of the other clinical assessments that were done were based on the reduction in pain as far as the patient was concerned.

That may have been because of the fusion -- not the interbody fusion. It may have been a consequence of the fact that there was another fusion that was occurring. They got an extra special treatment.

What I think Dr. Cheng is saying is that the only thing you can really ascribe to the interbody device is the precise radiographic fusion; is that correct?

DR. CHENG: Yes, that is correct. Were the posterior lateral fusions with otogenous grafts or autologous grafts or how was that done?

MR. CHRISTIANSON: The protocol called for the interbody fusion to be autologous graft only. If the physician elected to do a posterior lateral fusion and there was sufficient otogenous graft, the protocol allowed for that.

The protocol also allowed for the posterior

lateral fusion to be augmented with allograft.

DR. WILKINSON: Madam chairman, may I comment on that as well? I, too, was struck, as Dr. Cheng was, that in the surgical technique descriptor that the panel was furnished, there is no description of posterior fusion that I could find.

It may have been in your protocol, but that is not what you gave this panel to evaluate.

I was given a statement that makes no mention whatsoever of fusion. Now I discover that 95 percent of the patients had a significant addition to the operation that I was not aware of.

DR. GROBLER: Could I have a question? I need to clarify this for myself as well. If you look at the surgical technique section -- I haven't looked at it now, but I think it was mentioned there that the section was taken out to the transverse process and bone added.

I can't recall that there were figures given on how many in actual fact had bone graft done or what was asked, and also how many of those were allograft or otogenous bone.

That will also influence the pain scale. How can you have pain over the donor site if no bone was taken?

That also brings -- I am somewhat confused as

well. I think what I need, at least, is how many cases -- are we saying that 95 percent had posterior lateral fusions as well?

Is it possible, theoretically, that the posterior lateral fusion could be solid and the interbody fusion could be a non-union?

In fact, the patient will then have a fused functional spinal unit, but the interbody fusion device was not assessed. That is a question.

DR. HALL: The interbody fusion device, that is the area that was assessed to be a fusion by the radiographic criteria in all cases, was the interbody fusion. That is the criteria that was used in all patients.

In no cases was the posterior lateral fusion used as the criteria of whether or not the patient was fused.

As far as the question about -- I am surprised. When you do a dissection to get the plates, to get the exposure to put in the screws and plates, to not put bone there, to me, from my standpoint, I can't imagine not putting bone there.

I guess I am surprised that you are surprised that bone was put in on those cases.

DR. GROBLER: I guess my concern is that in a study group I tried to give -- I think we need to get the

variables as low as possible.

If we have an interbody fusion, an intertransverse fusion, and an implant device, I am not sure what I am assessing.

DR. HALL: We used as our end point -- we used several end points. On the fusion end point, it was the interbody fusion end point.

We felt that was the one that was best assessed, because we had a radiolucent device, that we could assess that one very well with plane radiographs, two taken in the standard fashion and then plus with the caudal angle to get the better view of L5,S1, where you saw the wonderful X-rays, how well it is visualized. That was felt to be the key thing.

My experience has been that if you get an interbody fusion, those posterior lateral fusions are going to be fused or insignificant.

DR. GROBLER: You do have more blood loss, just to bring another factor in. There is more blood loss when you do that, I guess.

DR. CHENG: I don't think the question is whether or not you should do it, if you do the dissection. It seems logical to me as well, your comments about why wouldn't you just go ahead and do it.

However, I need to comment that I feel a little bit misled about what was done here, as this was brought out only upon questioning, that 95 percent of the people had a posterior lateral fusion as well, and there are not comments about pain at the donor site, hemorrhage, infection.

We all know there is significant discomfort at the otogenous graft donor site.

DR. HALL: Let me assure you, there was no attempt to deceive you in my clinical presentation. Let me assure you of that.

The donor site pain was assessed. That was one of the pain things that was assessed and it is in the data that was given to the FDA. Donor site pain was assessed on every single patient.

DR. BOYAN: We had that information provided to us. I think it is late in the hour and we are starting to have a little bit of memory loss. We are all of us a little bit older than 30, so we are forgiven. The information was provided to us.

DR. WILKINSON: However, it doesn't appear in the table.

DR. BOYAN: They provided us, though, with more than was in the -- do you have the documentation from the big box?

DR. WILKINSON : This is a table that is provided in this summary of safety and effectiveness. There is no category that I see for donor site pain.

MS. MORRIS: On the case report forms, it is listed, as I stated in the full form of the question, are all the sites that were assessed on the case report form.

DR. BOYAN: I think we have one more. Dr. Kerrigan, did you answer this? You did? Okay. I think we are ready for the next question.

MS. MORRIS: Question five is, are the amount of dural tears, amount of blood loss and the length of hospital stay, as assessed at 24 months, considered within the realm of clinical expectation for the device and the procedure at this time.

DR. BOYAN: We have discussed this in the general discussion pretty much at length. As we go around the room, maybe we could limit the discussion of this specific question to any new information or some very specific comments that we would like to make.

DR. CHENG: I think I have already asked you about the dural tears, so I won't comment any more about that.

DR. KERRIGAN: Pass.

DR. GROBLER: Can I ask, the 2.8 percent, there were 44 cases that had dural tears out of 221. Of those,

2.8 had a second procedure. The second procedure was due to a missed initial dural tear or why would a second procedure be done if it was found to have a dural tear at the time of surgery.

MR. CHRISTIANSON: The percentage that you are referring to are those patients who had a reoperation and, upon the second operation, it was determined that the cause was an unrecognized dural tear, that was repaired with the second surgical procedure.

DR. GROBLER: I have some concern regarding the dural tears, but I think we need to see it in the context of the severity of the cases, having cases that had secondary procedures.

I think the amount of blood loss, 1600 is too high, is pretty high. Length of hospital stay, do you have a range? It is seven days average. Is that between two and 20 days or do you have a range?

MR. CHRISTIANSON: It will take us a minute, but I believe we have that data.

DR. GROBLER: I can get that. That is basically it.

DR. HALE: No comment.

DR. YASZEMSKI: I think for a combined anterior and posterior infusion, 1600 milliliters is not too far off

from what I would expect. Seven day hospital stay is probably okay.

I think that this is, again, just a personal comment, since I do this type of surgery, too. I recognize that in a revision a dural tear happens a lot. I think I would report those that did happen and were closed at the index operation, so that surgeons contemplating doing this can have an accurate estimate of what they can expect. I would report all 44 in your table.

DR. JANOSKY: No comment.

DR. LAURENCIN: I definitely agree. If you are trying to introduce a procedure to someone and trying to let him know where the complications are, if one out of every five patients in a study group -- which I guess you are getting people who really know what they are doing and who don't go through a big learning curve -- if one out of every five are having that problem, you have to put that down.

DR. WILKINSON: Within the realm of clinical expectations, 44 tears, no, 1600 milliliters of blood, no, seven days on average, yes.

DR. HOLEMAN: No comment.

DR. SILKAITIS: No comment.

DR. BOYAN: Now we are on to question number six, which is a mega question. You are going to make that short

for us; right?

MS. MORRIS: Does the clinical data support the safety and effectiveness of the subject device for each indication treating one and two levels. If not, what data are necessary.

DR. BOYAN: Okay, we start with a surgeon. Let's start with Dr. Grobler.

DR. GROBLER: Can I pass on that? I will come back.

DR. BOYAN: Dr. Hale.

DR. HALE: No comment.

DR. YASZEMSKI: I want to say yes.

DR. JANOSKY: I pass for now.

DR. LAURENCIN: The clinical data supports the safety and effectiveness of the device that was tested. Again, I think the big question is whether you can accept the device that has a different fabrication method, different composition, and in fact the processing of this involves the use of polysolfone. They talk about small amounts of polysolfone that is in the material.

If I am reading it right, there is a little bit different design, too. They say when making this, they first made it with a polymeric strut from the mechanical testing.

Then they took the polymeric strut out and made the walls a bit thicker, when they were modifying it for the final form to be used.

So, if an implant has a different design, composition and fabrication method, can you approve that design based on a study that has a different implant design. That is my question.

DR. WILKINSON: My answer is no, I do not think safety and effectiveness has been shown for any of these categories, part of the reason being what Dr. Laurenson has just stated and part of the reason being in the six pages I have been asked not to read.

DR. HOLEMAN: No additional comments.

DR. SILKAITIS: No comment.

DR. CHENG: I believe the procedure itself needs to be modified to include the posterior lateral fusion if it was done in 95 percent of the cases. So, I think that definitely needs to be indicated.

My answer to the question is a maybe. The reason I say maybe is I think there are a couple of issues that should be addressed before final approval is given.

The one Dr. Laurenson has mentioned. The other one, my mind slips me at the moment. It has to do with splitting the analysis out into greater than two or less

than two and seeing the statistical review of that.

I don't have any comments with regard to spondylolisthesis in terms of degenerative or lytic.

DR. KERRIGAN: For all the reasons that have been discussed, my answer is to do the clinical data support the safety and effectiveness, the answer is no.

I think the second part of that is, if the data do not support the approval, what data are necessary. I think we have covered it fairly well in question four, with some of the other comments.

DR. GROBLER: I would think for category one, DDD yes, category three yes, category two no.

DR. BOYAN: That kind of leaves me. I was going to take the position of not making a comment. I guess, as a person who feels strongly that Dr. Laurenson's comments are fair, even given all of that, I would agree with Dr. Grobler and Dr. Yasemsky, that despite the fact that they have used a very different device, they have presented data from other countries that is quite convincing with the device that they are asking for approval for.

DR. LAURENCIN: The device they are talking about -- it is maybe a little bit confusing. They have been doing studies from 1991 on, but the device that they are talking about only started being implanted in Europe in 1996, for

two years. Are there any study data on that device?

DR. WILKINSON: Even if they do, we don't have it. We were not allowed to review that data before coming here. The FDA has a requirement of eight weeks in advance for data submission.

DR. LAURENCIN: I remember they said they had six years of data, and that struck me, too. I think it is only two years since they have been using it in Europe.

MR. CHRISTIANSON: Again, we maintain that the shift in the polymer fiber matrix is best addressed by bench testing.

We have extensively tested the device to show, in the laboratory, how this device performs mechanically. We believe that those data are sufficient and the device modifications are minor enough that they can be accepted based on the bench test data.

DR. LAURENCIN: Are you stating that by changing the composition, changing whether it contains some polysulfone and changing some of the design, just histologically, there can't be any change in its response histologically?

MR. CHRISTIANSON: That is correct. We don't believe that those changes would result in different histologic responses.

DR. LAURENCIN: That is interesting.

MS. MORRIS: Could I ask for a clarification?

DR. BOYAN: Sure.

MS. MORRIS: Is your response both biological response as well as mechanical performance?

DR. LAURENCIN: Certainly biological response would be number one. For instance, the fibers that you use in one case, the one in the IDE, what are they like. What are their dimensions.

MR. CHRISTIANSON: The length of the fiber is continuous in the device implanted in the IDE. The diameter is identical to the chopped fiber that we use in the device.

DR. LAURENCIN: So, you are saying it is one long fiber?

MR. CHRISTIANSON: In essence, yes.

DR. WILKINSON: It is not chopped?

MR. CHRISTIANSON: The device implanted in the IDE is not chopped.

DR. LAURENCIN: So, the IDE is one long continuous fiber, whereas in the --

MR. CHRISTIANSON: Multiple long continuous fibers.

DR. LAURENCIN: In the implant proposed, it is a chopped fiber. What are the lengths of the chopped fibers?

MR. CHRISTIANSON: Three millimeters.

DR. LAURENCIN: So, we are talking about a number of small, three-millimeter implants versus long continuous fibers.

MR. CHRISTIANSON: Recall, that in both cases they are bound in the polymer matrix within the device, so that from the biological point of view, what the patient is seeing, the interface between the bone and the implant is virtually identical.

DR. HOLEMAN: No.

DR. LAURENCIN: I cannot agree.

DR. BOYAN: Absolutely no. I guess that is a biologist answering and a chemical engineer on the other side, biologists, it isn't.

To the cell, they can count. They can sit on the surface and they can count those fibers. They know how many are there. It will not necessarily be a smooth surface, and the biology really could be different.

To say that it absolutely isn't without having done the experiment is a really dangerous sort of place to be.

DR. YASZEMSKI: Dr. Boya, may I ask one quick question?

DR. BOYAN: Yes.

DR. YASZEMSKI: Is the polyetherketone ether ketone ketone, to any extent biodegradable and, if so, over what time span?

MR. CHRISTIANSON: We have seen no data that indicates that this material is biodegradable. In fact, to do the incoming quality control on the material, it has to be dissolved in concentrated sulfuric acid to do the chemical analysis. It is highly inert to all normal solvents.

MS. MORRIS: Could I make a clarification or a request? In this question there is a portion that states, if not, what data are necessary.

Is there a recommendation that would address this issue, in terms of what additional data would be necessary?

DR. LAURENCIN: The histological testing of the implant itself, and also implantation in an animal model.

MS. MORRIS: So, we are saying an in vivo test in animal model and bench testing, not new clinical data?

DR. LAURENCIN: And clinical data in terms of implantation in a human would be good.

The polysulfone that is in the process of making the small, chopped fibers, if you implant that into tissues, does that leach out? Does it diffuse out at all? Does it have any effects when it does that?

MR. CHRISTIANSON: To our knowledge, no, it does not. It is bound in the polymer matrix, you will recall. The only potential effects that we thought the polysulfone might have is that polysulfone, when used as the primary polymer in other implants, has been subject to environmental stress cracking in a lipid environment.

So, we tested these cages in a lipid environment to see if there was a degradation and response.

DR. BOYAN: I think I would like us to move as quickly as we can through the questions. We are actually repeating ourselves. If we can find a way not to do that, that would be great.

MS. MORRIS: This is a subpart of the question referring to spondy. If there is sufficient data to support approval for spondy, should the type of spondy be specified to match the types represented in the study cohort?

DR. BOYAN: Let's just address that very specific question. Dr. Cheng, would you like to start this?

DR. CHENG: I don't have a strong feeling one way or the other.

DR. KERRIGAN: No. The first part, no.

DR. BOYAN: Do you have something else you want to add on? This is your chance.

DR. KERRIGAN: No.

DR. GROBLER: I think it should be categorized. There is as vast difference between an ismic and a degenerative spondylolisthesis. I would hate to think that some cases of ismic or degenerative would undergo two and three level fusions.

They is actually one mention of a three level fusion in a spondylolisthesis case. I think for that reason we need to know which cases, what was done on which category. I think it makes a difference and I would like to see that difference.

DR. HALE: I will defer to the clinicians on this one.

DR. YASZEMSKI: Yes. When the demographics statement was given at the beginning of the study the two groups were matched.

I think one would likely to find that a group of ismic spondylolisthesis patients, for example, and a group of degenerative spondylolisthesis patients would likely be very different.

DR. JANOSKY: Nothing to add.

DR. LAURENCIN: Nothing to add.

DR. WILKINSON: My responses really pertain to all three parts of six.

DR. HOLEMAN: No comments.

DR. BOYAN: Okay, we are ready.

MS. MORRIS: Seven. Does the clinical data from one to two levels and the limited data from three to four levels support the safety and effectiveness of the subject device for the use in treating greater than two levels for each indication. If not, what data are necessary.

DR. BOYAN: Let's start off with Dr. Yasemsky.

DR. YASZEMSKI: No, I think the numbers at the three and four levels are too small. Fusions of that magnitude, I think, are done fairly infrequently.

To specifically say it was safe and effective for those specific level indications would take many more patients, as the power analysis for one and two levels has shown.

DR. JANOSKY: I would say no, it dose not.

DR. LAURENCIN: I would agree, no.

DR. WILKINSON: I would say no, and what data is necessary, comparison with other devices, studies with the same device that they intend to market, not the device that they were making two years ago.

They clearly need more patients, but I hope they don't get them, because I don't think a lot of patients need four level fusions.

DR. HOLEMAN: No comments.

DR. SILKAITIS: No comment.

DR. CHENG: No, I agree with Dr. Yasemsky's comments.

DR. GROBLER: No.

DR. HALE: No, I agree with the other panel members.

DR. BOYAN: Next question.

MS. MORRIS: With respect to component removals addressed in question three, can you recommend appropriate device labeling to address this issue?

DR. BOYAN: We had someone lying in wait on the device labeling issue. Do you remember who that was?

Dr. Yasemsky, we are back to you. We will just keep going around.

DR. YASZEMSKI: I would suggest not recommending instrumentation removal in the label.

DR. JANOSKY: I concur.

DR. LAURENCIN: I agree.

DR. WILKINSON: I agree.

DR. HOLEMAN: I will agree.

DR. SILKAITIS: I agree.

DR. CHENG: I agree as well.

DR. KERRIGAN: I agree.

DR. GROBLER: I agree.

DR. HALE: I agree.

MS. MORRIS: Can I ask for clarification? That would mean that there would be a specific label that would say, do not remove the VSP components.

DR. BOYAN: I didn't hear that. I heard that it is not necessary to remove. Is that correct for the surgeons?

DR. LAURENCIN: Correct.

MS. MORRIS: Should specific surgeon training be a requirement for this device and procedure?

DR. BOYAN: Dr. Yasemsky, you are back on.

DR. YASZEMSKI: I think this is a complex issue and I hesitate to say yes, and at the same time I hesitate to say no.

It is a question that is, again, a cautioned maybe, and here are my reasons.

I think, as our sponsor representatives have said, this is a formidable surgical procedure. Surgeons who have, by virtue of their training and their experience the necessary expertise to carry out posterior lumber interbody fusions and intertransverse process fusions, I think will know what additional information they need, which may vary, depending on the surgeon, from reading the materials sent with the device to calling colleagues who have done it, to

calling representatives from the study group who have done it, and do it very safely.

On the other hand, I believe a surgeon who is not experienced in these techniques will not be safe simply by attending a course which then serves as justification for him or her to do the procedure.

DR. HALE: Can I ask for clarification? Are we asking to require the sponsor to provide the training or are we actually asking to require surgeons to attend the training?

MS. MORRIS: Good question. My interpretation is that the sponsor would provide training.

DR. HALE: Then it would be up to the surgeon's discretion whether or not they chose to attend that training.

DR. WITTEN: We are asking you for your recommendation about what type of training there should be, if you are recommending that there be training. I think a description of what type of training you think would be needed would be most useful.

DR. BOYAN: Dr. Yasemsky, as you were -- yes, Dr. Albert?

DR. ALBERT: One of the things we take into consideration as we go to a device approval is a restriction

to trained individuals and describe what that kind of restriction means.

What that restriction does is that the company be assured that their customers are appropriately trained to use their device.

We don't control the practice of medicine, but we do have authority over the manufacturers, and it is a restriction as to who they may distribute their device.

Generally, they provide the type of training but they are not required to supply personally the training. They describe what that training is, because they know the device best and the procedures best, that will support safe and effective use of their product.

The requirement is not that they then go out and provide the training program. That has been a question before.

It is for them to describe what type of training and we include a restriction of their distribution of devices to trained individuals.

DR. BOYAN: What I would like to do is, just ask Dr. Yasemsky, when you addressed that question just a second ago, as you addressed it -- I heard you say it would be good to be trained, but I didn't hear you make the kind of answer that Dr. Albert's answer would have led you to make.

DR. YASZEMSKI: I would suggest not restricting sale to person's who received some specific training just to satisfy the conditions of that sale.

I would, however, like to see the company offer training that surgeons could take advantage of if they wished to.

DR. JANOSKY: I would have a suggestion strongly that specific surgeon training be obtained; not required, but a suggestion.

DR. LAURENCIN: I think because of the complexity of the procedure involving two different technologies combined together, I think there should be some strong wording about training for surgeons.

DR. WILKINSON: I agree completely with Dr. Yasemsky.

DR. HOLEMAN: I will agree that the surgeons should be trained. Even when I say that, I am not so sure, even if the sponsor provides that, how that is going to be monitored.

Who will say that a surgeon who has not been trained cannot perform this surgery?

DR. BOYAN: We could recommend that a surgeon be certified to perform this surgery, however it is that he or she becomes certified. That is something we could do.

MR. CHRISTIANSON: May I make a comment in that regard?

DR. BOYAN: Sure.

DR. CHRISTIANSON: No company can certify that a surgeon is qualified to do the training. That, quite frankly, is a job for the medical societies and local medical boards.

DR. CHENG: Mr. Christianson, that is just what I was about to say. This is big surgery. It is formidable, as Dr. Yasemsky has said. There were big complications. There was an interoperative death. There may have been deaths elsewhere, overseas. I don't know.

I feel it is a very dangerous situation to require a manufacturer to assess a surgeon's capability. That is not their purview.

There are established organizations for doing that in the United States. If you feel that, or if this panel feels that special expertise is required, that should be -- given that decision or how to go about satisfying that condition should be forwarded to the appropriate accrediting body.

DR. KERRIGAN: I defer to the decision of the surgeons.

DR. GROBLER: I think training should be available

as a hands-on course. There should be centers of excellence available for the surgeon. At the end of the day, it is between the surgeon and the patient. It is his prerogative if he is going to take up the opportunity.

There are two reasons. One is, your clinical assessment with your patient and the second is the legal aspect.

I bet that at no time is the sponsor going to stand in the box with you if there is a legal case going. I think it is between you and the patient and it just needs to be available.

DR. HALE: I would agree. I would like to see the company provide the opportunity for the training, but I don't think it is necessary to require the surgeon to have that training.

MS. MORRIS: Question 10 says, is severe osteoporosis a contraindication for this device and procedure.

DR. BOYAN: Can I start? I think it is. Dr. Hale?

DR. HALE: I would say definitely, yes. With this device we are talking about a material that is as strong or stronger than normal bone.

Once we start having weakened bone structures due

to osteoporosis, there is an increased likelihood that the vertebral body is going to collapse and it is going to fail. I would say it is definitely a contraindication.

DR. GROBLER: I agree.

DR. KERRIGAN: I agree that it is a contraindication.

DR. CHENG: I agree.

DR. SILKAITIS: If the manufacturers desire that, there must be a reason, so I agree.

DR. HOLEMAN: That is a medical decision, so I pass.

DR. WILKINSON: I agree.

DR. LAURENCIN: I agree.

DR. JANOSKY: I agree.

DR. YASZEMSKI: I agree.

DR. BOYAN: That was an easy one. Next?

MS. MORRIS: What other labeling recommendations can you provide with respect to the package insert, patient information brochure, and the surgical technique manual.

DR. BOYAN: Let's begin that one with Dr. Laurenson. Are you prepared to answer that one?

DR. LAURENCIN: I think that all the things we talked about in terms of talking about the procedure, in terms of talking about the possibility of implant removal,

although it may not be recommended, if it is a clear possibility I think it should be discussed.

Also, the surgical technique manual, again, resolving the question of whether a posterior lateral fusion is done or not.

DR. WILKINSON: I have a number of things, many of which are relatively minor wording, which I will skip.

I think the device labeling should include a warning or precaution that prior anterior lumbar interbody fusion can result in potentially higher complications, as the manufacturer has told us.

I was puzzled by the warning that says avoid coupling of stainless steel with the IF cage. I assume that means you are not allowed to have the VSP implant come in direct contact with the cage. I would like to know how safe is safe.

The list of complications really needs to be cleaned up.

In the patient information brochure there should be mention of risk of harm to the nerve root, which is a distinguishing risk of pliff as opposed to posterior or posterior lateral fusion.

DR. HOLEMAN: With the patient brochure, I guess what is not included here would hopefully be provided by the

physician.

That has to do with the level of activity and how soon those kinds of things can be increased.

DR. SILKAITIS: No additional comments.

DR. CHENG: If this is approved, I think there should be some changes to the label. One, that posterior lateral fusion is required, or should be done, as the data were reported when that was done in most of the cases.

I think there should be post-market follow up of possibly wear debris. This is a new material and, as such, it does involve carbon fiber.

The experience with carbon fiber in polymers has been unacceptable wear debris characteristics. I think the responsibility for that should lie with the manufacturer.

In addition, I think the percentage of dural tears should be restated to indicate the true number, not just the cases that underwent reoperation.

I think the panel ought to review, or perhaps the FDA ought to review the amount of complications that occurred in overseas patients who had this operation done, as to whether there is a significant number of complications that might occur that might go on the label as well.

DR. KERRIGAN: Nothing to add.

DR. GROBLER: Most aspects I think we have

mentioned. Bracing; is this something that was used on a continuous basis? There was some discrepancy. Bracing was mentioned at one time. What was the bracing protocol?

DR. BRANNIGAN: It was a light brace, either a clam shell or some other light brace that was specified for a period of one month, specifically to provide muscle support and rehabilitation during the initial activity phase. It certainly doesn't support the bone.

DR. GROBLER: I have no further comments.

DR. HALE: Nothing to add.

DR. YASZEMSKI: Nothing to add.

DR. JANOSKY: Nothing to add.

DR. BOYAN: Have we addressed that? We have just decided not to pursue question 12.

MS. MORRIS: Yes.

DR. BOYAN: This concludes our questions and we are going to omit the last question. Okay, thank you very much.

We are now at that moment and, because it is so late, I am going to take the chairman's prerogative of making a few announcements before we get to the very end, because everybody is going to jump up and run.

I would like to remind the panel that we have a closed session here at 8:00 o'clock tomorrow morning. If

for any reason it is not going to be in this room, there will be signs that tell us what room we should go to.

Also, I want to remind the panel to remove all of your stuff from the room. This is not a secure room tonight. If you bring your things back tomorrow, it will be handled for us at that time by FDA staff. We have to clear the room out of all documents that may have been brought in during the day.

Now I would like to turn the floor over to Ms. Nashman, who will read us our instructions.

MS. NASHMAN: Once again, I wish I could just omit reading this, being as we have heard it before, but that is incorrect.

The panel recommendation options for premarket approval applications.

The medical device amendments to the Federal Food, Drug and Cosmetic Act require that the Food and Drug Administration obtain a recommendation from an outside expert advisory panel on designated medical device premarket approval applications that are filed with the agency.

The PMA must stand on its own merits and your recommendation must be supported by safety and effectiveness data in the application or by applicable publicly available information.

Safety is defined in the act as reasonable assurance, based on valid scientific evidence, that the probable benefits to health under conditions of use outweigh any probable risks.

Effectiveness is defined as reasonable assurance that, in a significant portion of the population use of the device for its intended uses and conditions of use when labeled will provide clinically significant results.

Your recommendation options for the vote are as follows:

One, approval, there are no conditions attached;

Two, approvable with conditions. You may recommend that the PMA be found approvable subject to specified conditions such as resolution of the clearly identified deficiencies which have been cited by you or by FDA staff.

Prior to voting, all the conditions are discussed by the panel and listed by the panel chair.

You may specify what type of follow up to the applicant's response to the conditions of the approval recommendation you want, for example, FDA or panel.

Panel follow up is usually done from homework assignments to the primary reviewers of the application or to other specified members of the panel.

A formal discussion of the application at a future panel meeting is not usually held.

If you recommend post-approval requirements to be imposed as a condition of approval, then your recommendation should address the following points:

A, the purpose of the requirement; B, the number of subjects to be evaluated; and C, the reports that should be required to be submitted.

The last recommendation you could make would be that of not approval. Of the five reasons that the act specifies for denial of approval, the following three reasons are applicable to panel deliberations:

A, the data do not provide reasonable assurance that the device is safe under the conditions of use prescribed, recommended or suggested in the proposed labeling;

B, reasonable assurance has not been given that the device is effective under the conditions of use prescribed, described, recommended in the labeling; and

C, based on a fair evaluation of all the material facts in your discussion, you believe the proposed labeling to be false or misleading.

If you recommend that the application is not approvable for any of these stated reasons, then we ask that

you identify the measures that you think are necessary for the application to be placed in approvable form.

Again, following the voting, the chair will ask each panel member to present a brief statement outlining the reasons for their vote.

DR. BOYAN: I was just struck, that is like being on an airplane and listening to the thing when the mask comes down. You think you have heard it every time, but it never hurts to hear it one more time.

Okay, before the beginning of the voting process I would like to mention, both for the panel members and for the record, that the votes taken are votes in favor of or against the motions made by the panel.

Votes are not for or against the product.

I would like to ask Dr. Grobler, would you like to make the motion?

DR. GROBLER: I was wondering, why pick on me.

I think everyone would agree, at this time of the evening and after what we have listened to, this is to me, at least, a fairly complex case.

I think to expedite the discussion I am going to put a motion on the table and let it go from there.

I move a motion to approve with conditions to be decided and suggested by both the panel and the FDA staff.

DR. BOYAN: Say that one more time?

DR. GROBLER: To be approved with the conditions, and there are many of those, to be outlined by the whole committee.

I feel to some extent that to put down every single thing that we discussed, I feel this is something that needs to be done in the committee, but I think that should put on the table as far as what conditions need to be brought into this approval.

This is only a motion, I presume.

DR. BOYAN: Right. The motion is to approve with conditions, conditions to be described by all of us, and maybe we could just go around the table and list the conditions.

I need a second before we can start adding the conditions. Does everybody understand what the motion is? It is sort of an open motion right now.

It is a motion to approve with conditions, pending a listing of the conditions, because there are so many of them that Dr. Grobler would like us to take some time to go down that list.

DR. LAURENCIN: A point of information? A question to Mr. Christianson. The implants in Europe that have been placed, have they been placed with VSP-4s or VSP-

5s?

MR. CHRISTIANSON: Both, but currently VSP fifth generation.

DR. LAURENCIN: How many VSP-5s have been placed with your new modified implant?

MR. CHRISTIANSON: I can't answer that question off the top of my head. I would have to go back to the company and research that information for you.

DR. BOYAN: I sense we are kind of confused here. Maybe you could find another way of phrasing what you said so we could have a motion to work from.

DR. WITTEN: I don't think this will help clarify what you need, but I will tell you what we need in clarification.

The motion has to be either for approval or approvable with conditions, and the conditions stated.

DR. BOYAN: Right, we are going to state them. I am just trying to decide, can we get away with not stating at them at this exact motion in a parliamentary way, and while the motion is sort of a motion in progress, as opposed to being a firmly stated motion.

DR. WILKINSON: May I make a point of order, Madam Chairman?

DR. BOYAN: Yes.

DR. WILKINSON: Since the motion has not been seconded, the floor is still open for a second motion.

DR. BOYAN: It certainly is.

DR. GROBLER: That is partly the reason why I did it this way. I think something needs to be on the table and we decide or get another motion.

DR. BOYAN: Let me state the motion that is currently on the table so that we understand it, and then we can wait for a second or we can have another motion while we are waiting.

We have approvable with conditions, the conditions having been discussed for the last five hours, and these conditions will be outlined by us in just a minute, if that motion is seconded.

Okay, let's try another motion. Can I have another motion?

DR. LAURENCIN: I make a motion for non-approval.

DR. BOYAN: Do I have a second for that motion?

DR. WILKINSON: Second.

DR. BOYAN: So, now we have a motion to not approve and a second. This motion is now open for discussion. Is there any discussion?

MS. NASHMAN: In your discussion if you could describe the issues that you would like to be addressed to

make the application approvable, if you could include that in your discussion.

DR. LAURENCIN: The main issue, as I see it -- clearly there is a lot of good work that has been done on this project, and also, the implant itself may be efficacious in that it is in many cases treating a very bad problem, revision back surgery.

The fact that the implant itself is going to be a different implant from the study is, to me, a major flaw and major question.

The fact that we don't have a good handle even anecdotally in terms of what the experience is in Europe right now in terms of implants, we don't know even the numbers that are being implanted and what are the outcomes on those.

Apparently there is no study going on right now that is assessing this.

I do believe that an implant that has a different composition, that does have a different method for fabrication, that contains the carbon fibers -- you know, in the orthopedic world, carbon fibers always raise a red flag.

In this case, we are going from a carbon fiber design which you would think would be more benign, which is a long-fiber design, to a fiber fragment design. I do

believe that puts us more at risk.

I do believe there are differences. If I take micron sized polyethylene and place it in bone, it has a very different responses to centimeter wedges of polyethylene placed next to bone and it will have a radically different response. I do believe those factors are there.

DR. BOYAN: Let's just keep going. Dr. Wilkinson?

DR. WILKINSON: I am concerned also that we are now asked to review a device which is not the same device. The VSP design is not the same VSP design.

There is a posterior fusion that we weren't expecting that came out of the woodwork somewhere.

We don't have any biological equivalents for this new material. The chief proponent has agreed that this radical type surgery is probably not necessary or superior for a single level fusion, and the panel has advised the FDA that there is insufficient data for a three or four level fusion.

Basically we are left with an indication for only a two-level fusion, but that is not what all of this data is about.

I just think the device as used in this way, these two devices used together in this way, is a very dangerous

and major undertaking.

For me to approve it, I need to be convinced that it is really superior to either device used alone. We were not given that data.

DR. BOYAN: Dr. Hollman, do you have anything?

DR. HOLEMAN: No.

DR. BOYAN: Dr. Zokitas?

DR. SILKAITIS: No.

DR. BOYAN: Dr. Cheng?

DR. CHENG: Are we voting on this?

DR. BOYAN: No, we are not voting. Right now we are trying to give advice to the sponsor or give advice to the FDA as to what we would have like to have seen to make this be an approvable --

DR. CHENG: It sounds like we have agreed on something, though.

DR. WILKINSON: I think we have to vote first.

DR. BOYAN: I think you are absolutely right. As I was listening to you, it almost sounded like we had voted. We have not voted. That is right.

DR. CHENG: Why don't you call the vote.

DR. SILKAITIS: Dr. Boya, I wanted to ask the manufacturer regarding the device design. If the issue is the one in the IDE would be marketed, would that change the

conclusion or opinions.

DR. WILKINSON: It would change the PMA.

DR. SILKAITIS: No, if they were to market the device that was studied in the IDE, would that change the opinion regarding the data that has been presented.

DR. BOYAN: Do you want to ask the panel or do you want to ask the --

DR. SILKAITIS: First of all, the manufacturer and then the panel members.

MR. CHRISTIANSON: Obviously, we presented all the mechanical testing data, all the preclinical testing data, in the context of the device that we studied in the IDE.

We would be perfectly happy to work with FDA to address the issues of differences between the device we studied in the IDE and the device that we are seeking approval for here today, if that would change the nature and the complexion of the discussion we are having right now.

DR. BOYAN: I am kind of at a loss as chairman as to what to do right now.

You called the question, right? Did you want to do that, Dr. Cheng, call the question and have us vote on this motion?

DR. CHENG: I believe the outcome appears to be obvious, unless I am wrong.

DR. BOYAN: I think we should do the exercise, though, and find out if it is.

DR. WILKINSON: I would like to call the question.

DR. BOYAN: We had the question called, and I think when you call the question, that is it. Why don't we start the vote, then, with Dr. Hale, who is over here quietly.

DR. HALE: The motion, again, is to outright disapprove; is that right?

DR. BOYAN: Right. I will restate the motion in just a second. If the motion is carried, then all of us have an opportunity to speak as to why we voted the way we did.

If it doesn't carry, then we get a new motion on the table and we get a second and we vote for that one.

So, at the present time the motion is not approvable, and it has been seconded. We have been asked that, if that motion carries, that we provide information to FDA that would help them help the sponsor in coming up with a proposal that would be approvable.

If we can just begin the voting, and let's begin with Dr. Hale, then we can determine whether or not we need another motion.

DR. HALE: No.

DR. BOYAN: You are voting against the motion and the motion is to not approve.

DR. HALE: That is correct.

DR. BOYAN: Dr. Yasemsky.

DR. YASZEMSKI: Yes.

DR. BOYAN: You are voting to not approve.

DR. YASZEMSKI: Correct.

DR. JANOSKY: I am voting to not approve.

DR. LAURENCIN: Yes.

DR. WILKINSON: Yes, not approve.

DR. CHENG: Yes.

DR. KERRIGAN: Yes.

DR. GROBLER: This motion, yes.

DR. BOYAN: The motion carries to not approve. It was seven to not approve, one to approve, and now I would like to start again with Dr. Hale and explain why you took the position that you took, and remember to try to give as much information as you can give that would be helpful.

DR. HALE: I think I basically agree with the comments the rest of the panel made. I guess my reason in voting the way I did is that I felt that this was something that the sponsor could have potentially addressed as a condition, since they contend that these two different cage designs are essentially equivalent.

It seems like it would be a matter of presenting the data to support that. There were obviously other issues that would require additional data to present as well.

To me, it was a fine line whether to give them that opportunity to present that as a condition or to just flat out say no.

DR. YASZEMSKI: I think that the number of things that were uncertain was too large for me to vote no. As suggestions I would say that it appears that data from Europe is imminent, if it started in 1996.

From a chemical engineering perspective, along with Dr. Laurenson, I am concerned about the change in potential material/biological interaction with both the composition change and the processing change.

I think that should be addressable very soon, with two-year data from Europe. If there is no significant difference from the European data to that which was presented in the IDE, then I would be happy, personally, to say that the processing and composition change is fine.

I think that the stratification of anterior and posterior fusion would help me quite a bit. Both how many were done -- which I think has been alluded to, that is, 95 percent, but also, since there is a posterior fusion now, it would really help me to have a rating system applied to it,

so I could make some comparison for myself between radiographic fusion posteriorly, radiographic fusion anteriorly, and patient outcomes.

With respect to the potential for getting the European data, it would help me to see what surgeons are doing in Europe, and are surgeons using the device in an anterior only fashion.

If that is the case, is there any difference in the complication rate when this device is only used anteriorly.

I think, again, we get to this discussion of what the end point for the device is and what the end point for the operation is.

I think the surgeons are going to make decisions as to whether to do a front fusion only, a back fusion only, or a back and front fusion, and then make the assumption that the patient outcome is in some way related to attainment of a fusion.

I personally would be happy, from an effectiveness perspective on the device only, if the goal of the device was to help a fusion happen and a fusion happened in enough cases, that would be enough for me to say the device is effective.

I would leave the decision as to whether fusion is

the appropriate operation for the patient's problem to the surgeon.

DR. JANOSKY: I voted the way I did because of non-reasonable assurance of particularly effectiveness and safety also.

In addition to the issues I raised today, I think it is very important that subgroup analyses be done to identify whether each one of the patient groups, especially coming with different personal characteristics and different starting points, whether we are seeing the same effect in terms of effectiveness. So, definitely subgroup analyses.

DR. LAURENCIN: Again, I am dismayed that the device that is proposed to be placed in patients is different from the device that was under study.

I think that there are differences that are there. If one looks at the form and function of the material in terms of the way it is proposed, the new material the way it is proposed to be used, there is more cause for alarm and more cause for pause in using that matrix type of design. That is the main reason.

I do believe that there is European data that is there. It is almost near the two-year point. I am sure that if you have used a new device with a new VSP I am sure there must be a study going on in terms of looking at the efficacy

of making such a change. That information and that data could be utilized.

DR. WILKINSON: I think I already stated my reasons earlier.

DR. CHENG: I guess I am sorry I can't in good conscience believe it is in the public's best interest to vote other than how I voted.

The reason why, well, we have gone over it before, but just to reiterate very quickly, the device is different than was presented to us with the data as supportive of a different device.

I think that reanalysis of the data really, with regard to the number of levels of fusion, really would be very helpful.

There is some additional data which you probably have from the other European studies in regard to the complications and so forth that could support the application in a much more strong manner.

I would also suggest that all of the issues that were brought to light by this panel committee and any others which were not discovered by the panel committee be presented in a frank manner.

DR. KERRIGAN: The reason I voted the way I did is I don't feel we have reasonable assurance of safety and

efficacy.

I started off thinking that the criteria -- I didn't agree with the criteria of patient success and pain. I think the other comments that were brought up, particularly Dr. Laurenson's was the most key one, is even stronger why I feel these have not been demonstrated.

DR. GROBLER: I think the study needs very strong work to be done on the implant data. I think that is still achievable.

If it can be shown that the new carbon fiber construct is similar to the previous one, that can be recouped.

The European work definitely needs to be brought in.

Fusion assessment needs to be geared into both the anterior and the posterior lateral. We said about the three and four level fusion. I also have concern about the spondylolisthesis group, in fear that we may be over-treating some of the one level spondylolisthesis pathology. That is all I have to add.

DR. BOYAN: I would like to thank all the panel members. I especially would like to thank the representatives of Acromed. It was a beautiful presentation that you made.

I would thank again the FDA. I did make one mistake. I told the panel members that they had to be here for a closed session at 8:00.

They have to be at panel training at 8:00 and the closed session is at 8:30. Now I would like to turn the meeting back over to Ms. Nashman.

MS. NASHMAN: Do you have a comment?

DR. WITTEN: I just would like to thank the panel especially for helping us out for a long day, and thank the sponsors and the audience and the other people in the FDA who came to hear this presentation and discussion.

MS. NASHMAN: I guess I would just like to third the thanks and especially for the clarity at this late hour. At this time we are adjourned until tomorrow morning.

(Whereupon, at 11:20 p.m., the meeting was adjourned.)