00001

CENTERS FOR MEDICARE AND MEDICAID SERVICES
Medicare Coverage Advisory Committee

November 30, 2006

Centers for Medicare and Medicaid Services
7500 Security Boulevard
Baltimore, Maryland

00002

Panelists

Chairperson
Alexander H. Krist, M.D.

Voting Members
Mark V. Boswell, M.D., Ph.D.
Barbara D. Boyan, Ph.D.
Kim J. Burchiel, M.D.
A. Mark Fendrick, M.D.
David R. Flum, M.D., M.P.H., F.A.C.S.
Jeffrey G. Jarvick, M.D., M.P.H.
Stephen L. Ondra, M.D.
Laxmaiah Manchikanti, M.D.
John S. Kirkpatrick, M.D.

HCFA Liaison
Steve Phurrough, M.D., M.P.A.

Consumer Representative
Charles J. Queenan, III

Industry Representative
Kim K. Kuebler, M.N., R.N.

00003

Panelists (Continued)

Guests
Tom Faciszewski, M.D.
Jon D. Lurie, M.D., M.S.

Executive Secretary
Michelle Atkinson

00004

TABLE OF CONTENTS

00005

00006

PANEL PROCEEDINGS

(The meeting was called to order at 8:05 a.m., Thursday, November 30, 2006.)

DR. PHURROUGH: Good morning, folks. We're getting ourselves ready here in just a minute, so if folks will take their seat, we'll get started here. I yell pretty well, so I'll start. I am Steve Phurrough, I am the director of the coverage and analysis group and the representative, CMS representative to this particular panel. Welcome today. I'm going to repeat this over and over and over again. I have been repeating this for the last three months and no one seems to believe me. We are not doing a national coverage determination on spinal fusion. We are not doing a national coverage determination on spinal fusion.

(Laughter.)

The purpose of the Medicare Coverage Advisory Committee is to provide recommendations to CMS as to the strength of evidence around a particular technology that we ask them to review. The purpose of the advisory committee is not to tell us whether we should or should not cover

00007

something. It's not their role, it's not a role that we allow them to assume, because even if we are having the, this particular committee meet during a national coverage determination where we may in fact be asking them what's the evidence around the particular issue that we are addressing in that NCD, we're still not asking them to tell us whether we should cover it or not. We never

ask them that. In fact, they are prohibited by their charter from doing that. I just wanted to clarify that. I have had this interview now for 12 to 15 times, and not always accurately -- well, I should say that it's usually accurately quoted, it's just not believed. We're not doing a national coverage determination.

Now the reason we have these kinds of meetings is to stimulate many of you who are involved in the treatment of patients, in the development of evidence or in the development of technologies to recognize what this particular group and what CMS believe about the state of evidence around a particular technology and to stimulate you to do something about those gaps, if those gaps are present. We think we have done that successfully in some other MCAC's, or at

00008

least that has begun, and that is their role here today, is the evidence sufficient for us to be comfortable as a community, health care community in the use of spinal fusion in degenerative disc disease, or are there some gaps that need to be filled, and if there are some gaps that need to be filled, we will be encouraging you to help fill those gaps over the ensuing months to years.

So that's the goal here today. We expect a lively discussion. It would be a different MCAC if the discussion was not lively. There are time lines to be met. Those time lines will be met. If you are a speaker, you have a specific of time to speak and when that time is up, your speaking is up. So you may be cut off. There are a lot of people who want to talk, you have your time, and you need to finish within that particular period of time.

There are some specific formal things that we need to get done to begin with, the microphones are now on, so I can quit yelling, and I will turn this over to Michelle Atkinson, the executive secretary, and let her start with the official business.

MS. ATKINSON: Thank you, Steve. Good

00009

morning and welcome, committee chairperson, members and guests. I'm Michelle Atkinson, the executive secretary for the Medicare Coverage Advisory Committee. The committee is here today to discuss spinal fusion for the treatment of low back pain secondary to lumbar degenerative disc disease.

The following announcement addresses conflict of interest issues associated with this meeting and is made part of the record. The conflict of interest statutes prohibit special

government employees from participating in matters that could affect their or their employers' financial interest. Each member will be asked to disclose any financial conflicts of interest during their introduction. We ask in the interest of fairness that all persons making statements or presentations also disclose any current or previous financial involvement in any company that manufactures tools used for the diagnosis or treatment of spinal problems. This includes direct financial investments, consulting fees, and significant institutional support. If you haven't already received a disclosure statement, they are available on the table outside of this room.

00010

We ask that all presenters please adhere to their time limits. We have numerous presenters to hear from today and a very tight agenda, and therefore cannot allow for extra time. There is a timer at the podium that you should follow. The light will begin flashing when there is two minutes remaining and then turn red when your time is up. Please note that there is a chair for the next speaker, and please proceed to the chair when it is your turn.
For the record, the voting members present for today's meeting are Mark Boswell, Barbara Boyan, Kim Burchiel, Mark Fendrick, David Flum, Jeffrey Jarvick, Stephen Ondra, Laxmaiah Manchikanti, John Kirkpatrick. A quorum is present and no one has been recused because of conflicts of interest.
The entire panel, including non-voting members, will participate in the voting. The voting scores will be available on our web site following the meeting. Two averages will be calculated, one for the voting members and one for the entire panel.
I ask that all panel members please speak directly into the mike. You may need to

00011

move the mikes and share them. And lastly, please remember to discard your trash in the trash cans located outside. Thank you very much, and I will turn it back over to Steve.
DR. PHURROUGH: One last introduction before the rest of the panel is introduced. Unfortunately, I will need to be stepping out of the meeting for brief periods of time during the day. Dr. Marcel Salive, who is the division director for issues to include spinal surgery, will be filling in on those occasions to provide CMS input. I have no other issues, so we'll turn it over to Dr. Krist.
DR. KRIST: Thank you all for coming

here today. I want to start by thanking you for the work you have done to get here today, and CMS for putting this together, for the technology assessment, for all the presenters today, as well as the panel members.
As Michelle was saying, we have a tight schedule, and one of my jobs is going to be to keep us on schedule, so I apologize if I end up cutting you off. If you could pay attention to those lights and try to stick to the time lines, that would be good, because I'd like to make sure

00012
that the panel has enough time to ask questions, to clarify the evidence issues that they have, as well as have time to discuss all the topics that you lay down for us.
For speakers, when you get up to the podium, if you would introduce yourself, where you're from, and also the conflicts of interest to begin with, and we'll start actually with doing the same thing for our panel here, so starting at the end with Tom, and introduce yourself and say if you have any conflicts of interest.
DR. FACISZEWSKI: Tom Faciszewski, with the Marshfield Clinic, and I have nothing to disclose, no conflicts.
DR. LURIE: Jon Lurie (inaudible).
MR. QUEENAN: I'm Charlie Queenan, a management consultant, and I have no conflicts of interest.
MS. KUEBLER: My name is Kim Kuebler, I'm associate director of the medical publication Respiratory, and am the industry representative. I have spoken on (inaudible) and received grants from PhRMA, Abbott Laboratories and the Michigan Department of Community Health. I have also been a contact point for the Advancement of Medical

00013
Technology Association, which included representatives of Avinet, Abilene Health, Des Plaines Spine, Innovative Spinal Technologies, Medtronic and Zimmer.
DR. KIRKPATRICK: I'm John Kirkpatrick, I'm from the University of Florida, an orthopedic spine surgeon, and I have to beg ignorance. My conflicts were reviewed as far as stock holdings and they were close to threshold, but I have not been informed whether they have to be disclosed. Michelle?
MS. ATKINSON: No.
DR. KIRKPATRICK: So I'm fine, thanks.
DR. MANCHIKANTI: Laxmaiah Manchikanti, Pain Management Center of Paducah, Kentucky. I'm an interventional pain physician. I'm also CEO of a group of pain physicians. I do not have any

conflicts of interest.
DR. ONDRA: Steve Ondra, from Northwestern University in Chicago, I am a spine surgeon. I have grant and support, as well as consulting.
DR. JARVICK: I'm Jeffrey Jarvick, a radiologist at the University of Washington, and I'm a co-founder and stockholder of a company

00014
called (inaudible).
DR. FLUM: I'm Dave Flum from the University of Washington, and am a general surgeon and researcher. I have NIH support for research but no conflicts.
DR. FENDRICK: Mark Fendrick, general internal medicine at the University of Michigan. No conflicts.
DR. BURCHIEL: Kim Burchiel, Oregon Health and Science University neurosurgery, and I have nothing to disclose, no conflicts.
DR. BOYAN: I'm Barbara Boyan, from Georgia Tech and Emory, I'm a tissue engineering faculty member, and I am totally conflicted as far as being a consultant for MTF, BioMed, Medtronic, Zimmer, and I'm on the board of directors of two companies, both of which have spine products.
DR. BOSWELL: Mark Boswell, Texas State University Health Sciences pain management. No conflicts.
DR. KRIST: I'm Alex Krist. I'm at Virginia Commonwealth University in the department of family medicine, and I have no conflicts to disclose.
And I think at this point we'll turn it

00015
offer to Dr. Feinglass to do the CMS presentation of the background topics and the voting questions for discussion.
DR. FEINGLASS: Good morning. I want to thank you all again for coming out very early to Baltimore. To reiterate what Dr. Phurrough said, we are not doing an NCD on spinal fusion, just so you hear it again. As many of you know, surgical intervention for degenerative disc disease is really a very controversial issue on many levels. Because of this controversy, we have brought you all here today to hopefully evaluate the impact of this treatment, to target this benefit to the Medicare population and shed some light on the current state of the evidence in the field.
I also would just at the outset of this meeting like to acknowledge the long awaited SPORT trial which was published last week in JAMA. You will notice that SPORT will not be debated in much

of the evidence presented today because the data published thus far deals with information in an area not addressed by the MCAC today. And also, the data that is available publicly was just published last week. We commend the SPORT

00016

researchers for their work and encourage others in the field to look at doing large scale RCTs, and look forward to the rest of the data that will come from the SPORT trials.
So to begin with the official reading of the questions. Today's hearing is to consider spinal fusion for the treatment of low back pain secondary to degenerative disc disease.
The first question: What level of confidence does the evidence provide in addressing the outcomes needed to determine the effectiveness of lumbar spinal fusion for low back pain due to lumbar degenerative disc disease? Rate it from low to high. Discussion questions: Is the relief of pain the appropriate primary outcome, or should it be restoration of function, return to work, or something else?
Question 2: What level of confidence does the evidence provide for characterizing the complications, adverse events and other harms from lumbar spinal fusion for degenerative disc disease, both short term and long term?
Discussion questions: What does the variability in surgical risk depend on? As this procedure is permanent, are there other potential long-term

00017

harms that have not been discussed?
Third question: Based on the evidence presented, how likely is it that lumbar spinal fusion for lumbar degenerative disc disease improves clinical outcomes as compared to conservative treatment, both in the short term and long term? Discussion questions: What are the causes of low back pain? Is patient selection important, and if so, what are the clinical and/or patient characteristics that are reliable predictors of satisfactory outcomes? If there is an absence of evidence of long-term benefit, would evidence of short-term benefit be sufficient to justify a fusion procedure? If one clinical trial were to be done, what should it be?
Question 4: Based on the evidence presented, how likely is it that the various fusion procedures improve health outcomes for lumbar degenerative disc disease? Consider these procedures both with and without instrumentation, short term, long term, and then with instrumentation and without instrumentation, and then we ask about specific procedures.

Discussion: How important is patient selection relative to the type of procedure? What criteria

00018

are used to select the type of fusion procedure? Question 5: What level of confidence does the evidence provide that radiographic interpretations are correlated with clinical outcomes for lumbar spinal fusion due to lumbar degenerative disc disease? Is there uniform agreement regarding terminology for radiographic interpretations? Question 6: Based on the evidence presented, how likely is it that the results generalize to the Medicare population for, A, relief of pain, and B, complications, adverse events and other harms? Discussion questions: Do studies need to be done in the Medicare population to strengthen the conclusions? Discuss the impact of age and comorbidities. Those are the official questions.

DR. KRIST: Okay. Next we'll hear the technology presentation from Dr. McCrory.

DR. MCCRORY: Good morning. I'm Doug McCrory, assistant professor of medicine at Duke University. Oh, I need to speak into the mike. There we go. Well, while we are waiting, I'd like to acknowledge my co-authors, two of whom are here today, Dr. Turner, a neurosurgeon who does a lot

00019

of work with spine surgery, who is on faculty at Duke, and Dr. William Richardson, who (inaudible) today, who provided an invaluable amount of evidence to this project. Here we go. I'm on the medical staff at Duke University and my co-authors' conflicts are listed there.

I want to start by, first of all, I think my remarks today are going to try to summarize the report. I'm not going to be able to provide all the details of the technology assessment, as it's rather lengthy and detailed. I want to go through a few slides, I hope it's not too repetitious for many members of the audience here today, and then get into the questions that we addressed.

So you know, first of all, we were concerned with lumbar spine disc disease. I reviewed this with members of our own staff who were not very well versed in this and go into it in some detail. I wanted to draw a (inaudible) facet joint, and we were interested not only in the (inaudible) disc, but also the disease that occurs from early (inaudible) spondylosis. The rationale for lumbar spine fusion

00020

really relates to relief of back pain or radiculopathy. It has been typically considered after conservative measures have not provided symptomatic relief. The fusion risks are fairly well described in the literature. They include risk of anesthesia, perioperative risks, as well as short and long-term risks associated with the spinal nerves.

Just to review the range of procedures that we're considering as part of the spinal fusion, anterior fusion or ALIF is a procedure often used in middle-aged patients who have symptoms, usually disc degeneration at a single level. They are often posterolateral fusion, which can be single or multiple levels, often utilizing bone grafts or metal instrumentation. And sometimes it's combined with lumbar decompression or laminectomy.

Circumferential fusion combines anterior and posterior, including interbody fusion done posteriorly, PLIF, and the particular choice among all these procedures varies as the degree of variety that is used, depending on comorbidities, symptoms, and the optimal approach to the lumbar spine is made on a case-by-case basis.

00021

The final slide, just part of the motivation for the technology assessment I believe was the increase in spinal fusion surgery that had been seen, and this slide from Dr. Cowan's presentation compares the thoracolumbar fusion rates, the cervical fusion rates and the lumbar rates. Now this is data from all ages, but the subgroups have shown that the increase in the rate of fusion has occurred not only in middle-aged patients but also in those over 65 years of age, which is the target population here.

So more specifically, the background for this report was that there was no systematic evaluation on the efficacy or safety of lumbar spinal fusion in the elderly population, and elderly patients may be clinically different because age-related changes in the spine may mean that they have a different disease, it may include diffuse disc and facet degeneration. And as people age, they generally get more comorbid conditions that could affect the procedure's efficacy and safety.

So, that brings us to the key question that we were charged to address with the technology assessment. In patients at least 65

00022

years of age with degenerative disc disease and/or degenerative joint disease of the lumbar spine,

what is the evidence regarding indications and outcomes, including the adverse events and overall net health benefit, of lumbar spinal fusion as compared to non-surgical conservative treatment or other surgical strategies? So you know, we wanted to pay attention to the age range of the patients which we were particularly interested in, primarily we were interested in outcomes, and the fact that we're interested specifically in how well they did and whether there were adverse events, compared with conservative management or other surgical procedures.

So, a few slides on methods. The way we decided to approach this question was to look at a variety of sources. We did look to web sites for professional society guidelines or preexisting systematic reviews. We wanted to look at the most recent data, so we also looked at news articles, data from the FDA, and solicited advice from several experts. We did a comprehensive search relying on MEDLINE, but we limited that to a primary search of literature since 2002; we relied on systematic reviews and other data to fill in

00023

for previous data previous to 2002.

So the purpose of the MEDLINE search was to identify the recent trials or recent series of cohorts. We did limit it to groups with more than 50 patients, and based on conservative treatments.

Our inclusion criteria was guided by the question. We were considering any lumbar spine fusion surgery, so regardless of the approach, anterior, posterior or both, and we looked at data on the use of instrumentation as well as non-instrumented fusions. Comparisons were nonsurgical management, which would include pain medication, treatment, injections, and rehabilitation strategy. We specifically did not include chiropractic treatment. We were also interested in comparisons with lumbar arthroplasty.

The set of outcomes that we were considering is also enumerated right here. We looked at two time frames, early and late, as defined in the questions. We were particularly interested in patient outcomes such as quality of life, disability measures, pain and pain medication use, in particular narcotic use. We

00024

looked also at adverse events including perioperative complications, later reoperation, and longer term outcomes such as adjacent segment disease. We considered radiographic fusion as an outcome, but only if it was ancillary to one of

the other two sets of outcomes, either the efficacy outcome or one of the adverse event outcomes in bullet three.
The Oswestry Disability Index, we looked at the disability as it is currently. We named this specifically in our methods as an item we were going to look for, specifically looking at the 10-item instrument that looks at activities of daily living, and each of the 10 items is scored with six response choices. Those scores are then standardized on a 0 to 100 scale. There has been a bit of background psychometric work on the scale with standard error measurements, that's about four points, but the minimal point and important difference, that is the difference in score that would correlate with the change in how the patient perceives this disability level on a global scale is in the range of 10 to 15. And indeed, the level of 15 was selected by the FDA for use in the, I believe the arthroplasty versus (inaudible)

00025

as being a clinical significant difference, in contrast to the change on visual and pain scales (inaudible) on a 0 to 100 scale, and for the RDI, somewhere in the range of 2.5 to 5.5.
In general, these sort of clinically meaningful global measures, where patients describe being much improved or very satisfied, can be seen from these satisfaction scales, and just incidentally, the minimal clinically meaningful changes in conservatively treated back pain is generally less smaller than those who are referred for fusion to treat their back pain.
So, we also had an organizing principle just in terms of our approach to the literature. We felt that the clinical rationale for selecting a procedure depended very much on the clinical presenting symptoms as well as the subjective degree of pain, as well as the patient's interest in surgery. We organized it according to the syndrome of axial back pain and spondylolisthesis. We weren't really able to be as fine in our gradations of the clinical syndrome as we would like, pretty much because of the limitations both ways, the information we had at hand, and it's not well described in the literature.

00026

I apologize that this slide is not terribly readable, but let me go through the results of our literature search. So just considering the Ovid MEDLINE searches, we had several independent subsearches, we did some supplemental searching that resulted in 125. And then we had about 273 review articles. Also, we looked at references from Cochrane's reviews and

other sources.
The total number of articles that we ended up looking at was 1,391. We excluded about a thousand of those because they weren't focusing on one of the areas of our key questions, but detailed (inaudible) many were excluded. The final evidence report includes citations of about 82, and granted, this is not an exhaustive list of all the literature, but the articles that (inaudible) these topics in order.
The first is dealing with the axial back pain, and what we did is look at comparisons of lumbar spinal fusion without surgery, and then data on anterior lumbar interbody fusion, posterior procedures, and arthroplasty. What we will do is look first at spondylolisthesis and then deal with some of the other areas, such as

00027

incidence of adjacent segment disease, look at other complications, and then we'll look at the difference between instrumented versus non-instrumented fusion from previous reviews. We'll look specifically at a group of studies that were performed in patients over 65 years of age and deal with a little bit of complications in that context. And then finally, we'll look at techniques to augment fusion.
So here are the results. Studies of lumbar fusion for axial back pain, I want to emphasize that there is no randomized clinical trial that has a direct comparison of lumbar spinal fusion with non-surgical treatments in the population of those greater than 65 years of age. We did identify four random studies of fusion that were not done in the U.S., for axial pain versus rehab, and all of these were done with a posterior or mixed fusion. Most of this data for this procedure are coming from other series of cohort studies.
I'll show you this graphic on the four randomized trials. For each trial, I'm showing both arms of the study. The open circle on the right represents the Oswestry score

00028

preoperatively, the closed circle on the left shows the Oswestry score postoperatively, so the arrow indicates a change in the Oswestry pre to post surgery. And then looking at both of the arms, you can get an idea what the difference is between those arms.
Looking at the Oswestry, you can see the reduction in the Oswestry score was between 10 and 15 for both the rehab arms and the fusion arms, and this difference between the length of that arrow and the length of that arrow would be

the difference in efficacy between those two procedures. As you can see for the four studies, the differences between the effects were somewhat different. The Fritzell study was statistically significant but the others did not show statistically significant differences, with the exception of Fairbank. Fairbank, I do want to note, was powered to detect a difference of about four points on the scale, and in fact the difference reported was 4.1, which is less than that which is usually recognized as a clinically significant difference.
And I just wanted to discuss for a moment, there are different ways of approaching

00029

the clinically significant, this was greater than 10, it's not quite 15, so you might argue that it doesn't matter clinically, but there was variation in the way a lot of the studies were validated. But if you're looking at a clinically important difference as to whether surgery significantly exceeded the effect of rehabilitation, then clearly a four-point difference doesn't get near that level required to be a clinically meaningful difference. So although the Fairbank study does show a statistically significant difference, we're not clear that it is a clinically significant difference between the pre and post surgery scores, compared with rehabilitation.
So with that framework in mind, the Fritzell study also did not achieve that 15-point difference between the fusion and the rehab arm.
Now, these studies (inaudible) and why is that? There were some differences in the patient population in terms of how long the patients had undergone conservative management and there were also differences in the degree of intensity of the rehabilitation therapy. For example, the Brox trial had one of the more intensive rehabilitation regimens, (inaudible)

00030

involving those components, and like Fritzell, seemed to be the most intense rehabilitation intervention. Now, that said as a hypothesis, the details were not terribly well described, and I'm not trying to pretend that's a conclusion that you can reach from these studies.
So in summary, you know, we believe that these do provide evidence of before and after clinically meaningful change in the stability index and shows a difference between fusion surgery and conservative treatment, but I don't know that we would necessarily call that important.
So moving on to the ALIF studies, as I

said before, we used an isolated single-level disc degeneration in adults with a positive discogram. The one theoretical advantage of (inaudible), show you data from a number of uncontrolled series as well as trials that compare ALIF with arthroplasty. These were primarily designed to show that ALIF are, that arthroplasty was better or worse than other possibilities such as posterior.
One of the first things I wanted to notice about these studies is that the change from

00031

before and after surgery on the ALIF arms was of a greater magnitude than we saw with lumbar fusion. So in every case here, it exceeded a change of 15 on the disability index, so from before and after, by frequency and disability. However, you know, there were no trials here that compared ALIF to conservative non-surgical approach. Some of these did compare to the arthroplasty arm but are not shown on this slide. So despite the fact that the change in the disability rating is of greater magnitude, the lack of a comparison may make it difficult to draw a conclusion about how effective ALIF is, as a particular procedure is considered to be conservative therapy.
So again, looking at posterior lumbar fusion, you'll see direct comparisons between different types of posterior fusions. Those include the (inaudible) separately to posterior fusions, (inaudible) or a circumferential fusion which was done as either an ALIF or -- I'm sorry, an anterior -- sorry -- a circumferential fusion which was done either anteriorly or entirely posterior.
You can see from this slide, the -- there was a reduction in scores, and what I showed

00032

you before was 24-month outcomes, and that was a reduction of about 13 points. And the point of this slide was to illustrate that the changes in visual and pain scale between these procedures, the experience was actually fairly similar, so these are not significantly different from one another. These groups aren't huge, but the power (inaudible) clinically important.
Last but not least, the RCT comparison, this was not a random controlled trial, these were patients who had initial observational studies looking at outcomes for different populations, but it did allow one fairly large sample sizes, a comparison of the pre-op to one year post-op, and could be randomized by the type of procedure. And as you can see visually, I apologize for the reproduction of the slide, the slides are in

order, so this is ALIF, PLIF and/or TLIF, posterior fusion, and 360-degree fusion, and they're on the same order over here. As you can see, the before and after here were all somewhat similar, there were some slight differences in that the PLIF and TLIF were, reductions were a little less than the others, and the posterior fusion was a little more than the others, but the

00033

differences were not very large. They didn't exceed that clinically significant threshold. To look at some additional data from the circumferential fusion, going with the anterior-posterior or entirely posterior two-level approach, we saw that these, the changes in the Oswestry were somewhat more variable, in a few cases the differences were large and exceeded 15, in many of the cases they were smaller and didn't. Some of these data came from randomized controlled trials but many had variations in technique, for example, the (inaudible), but many of them were just not controlled studies. So we didn't have any other cases where there was a control sufficient to reach conclusions about the efficacy relative to the conservative therapy arm. To show some data on the arthroplasty trials. The arthroplasty trials were done in specific indications in the study populations, so in all of those studies I will show you that the patients were single-level degenerative disc disease with only axial pain, or not only axial, but also a positive MRI or discogram measured around that level. These studies were all done in patients with an average age of around 40. But as

00034

I mentioned, they were inferior trials compared with ALIF, and all of them give data of no more than two years. On this study, I want to point out that the before and after Oswestry differences were generally greater than 15, and that's in terms of the mean change. I did want to illustrate, only one of these studies showed the proportion of patients who individually had a clinically significant change from before and after that exceeded 15 points. So here in this, you'll see that 68 percent of patients had a change in disability. Now prior to the study showing the changes in disability, and that's obviously when you're dealing with means, some patients have a better outcome, some patients have a worse outcome. We were looking for these sorts of results (inaudible) little bit for clinically important (inaudible) actually experience a large disability. Again, the drawback in interpreting

those arthroplasty trials is that none of them were compared to conservative treatment. We're going to switch gears now and leave the axial back pain studies and go on to the spondylolisthesis studies. As compared to the

00035

axial back pain studies, these spondylolisthesis studies, the goal is usually treatment of patient complaints, radiculopathy or spinal stenosis. And the, I wanted to primarily highlight the one randomized controlled trial that not only compares fusion to conservative treatment, and that is the Moller and Hedlund in 2000, and followed up then by Ekman in 2005 with long-term follow-up data. I'm showing you data here on these two panels on pain, and I must apologize, the original Moller and Hedlund study describes this as being measured with a disability index and then Ekman describes them in further studies, so I'm not sure whether we're (inaudible) identical correction to try to present everything in numbers that relate to, but these are, should be interpretable along with the Oswestry.
So the Moller-Hedlund study showed results in favor of surgery and compared to exercise at one and two years in both pain and in the disability rating index. And these differences here at one to two years look like they exceed 20 in terms of the pain scores and were just about 15 in terms of the disability rating index. One of the remarkable things about

00036

the Ekman study was it was one of the few that had really long-term follow-up data about the effect of, both the surgery, but also those patients who were maintained on conservative therapy. And remarkably, there didn't seem to be a great deal of difference between the conservative patients and those who had surgery. So the thing I wanted to point out was the trajectory of the non-surgery patients who didn't really worsen, they didn't get clinically significantly better, but these two groups' experience gets closer together as time goes on as a consequence of the mild decrease there and a mild increase in pain, and the same thing with the disability there, and this result is nonstatistically different for them.
I want to turn now to the instrumented fusion. There has been a lot of data, a lot of synthesis of this sort of data presented by (inaudible) summary instruction, clinically shown to increase the rate of radiographic fusion. The relationship between a radiographic fusion and outcome is only indirectly available and not terribly strong.

I want to show you data from Gibson and all, which compares randomized controlled trials

00037

looking at instrumented versus graft-only fusion. So we didn't want to reinvent the wheel, this seemed to be a pretty reliable analysis, and so this is the fusion data in the seven studies. Most of them had at least a trend for favoring instrumentation in terms of creating a fusion. The next slide shows poor clinical outcome as a measure where patients are asked whether they are better or worse than they were after surgery. And when you combine the data from all these randomized controlled trials, they favor instrumented fusions. Bono and Lee took a different approach, in that they not only looked at random controlled trials as the Cochrane review did, but looked at uncontrolled series. The studies were combined, but there were quite a few, and this shows the total number of patients involved in all these controlled studies and resulted in a large number of patients, and indeed, they did show statistically significant improvements here for any fusion versus no fusion, or, I'm sorry, instrument fusion versus no instrumentation, semirigid or rigid versus none. The fusion rates were not terribly impressive, but again, this is

00038

data from the 1980s and 1990s. When we look at the clinical outcomes as the Cochrane review did, there was a trend, though not statistically significant, for what's better, good or excellent results among noninstrumented fusion. Again, which is (inaudible) that we would not expect from the (inaudible) associated with the instrumented fusion. The only direct data we had from randomized controlled trials that I want to highlight was the Fritzell study. This was using bone graft, the square group was using VSP, and as you can see, the diagonals and the square is (inaudible) but also very similar for disability outcomes. The good fusion rates can be achieved without instrumentation, there may be an effect over time, and I think I'll hold it for the last slide, that in terms of open fusion, that the instrumentation, but the good fusion rates may be obtained without instrumentation. Okay. What I would like to highlight for a moment, the adjacent segment disease data, this required looking at long-term studies that correlated disease at levels either above or below fusion, and we looked at a bunch of uncontrolled

00039

studies that provided long-term outcome data at specific levels of the spine, and we found that when we were looking at this data relating to adjacent segments, the definitions vary a great deal in terms of how (inaudible) might be defined, and how variability was (inaudible) and we didn't analyze that data.
What we did feel was reliable was that data on reoperation for adjacent level disease which was more precisely measured or more reliably measured. So we actually looked at ten studies that had four to 14-year follow-up and calculated an overall pooled rate that was around three percent per year. We were not entirely comfortable attributing that to the effect of fusion, the original fusion causing increased motion at the adjacent segments, so we tried to figure out what would be an appropriate control for that. We did look at several long-term studies without fusion, which also gave us data on the reoperation rate for symptomatic recurring stenosis, and when we applied a similar methodology, the only difference being that we included reoperation at the same or different levels, we found a reoperation rate of about 2.5

00040

percent per year.
Again, this is a crude way of sort of trying to subtract out the background of progression of spinal disease, but we feel like it provided a little bit of additional information helping us to interpret that three percent per year figure.
We had a hard time addressing the issue of the elderly population in the focus question. One way we attempted to do that is we identified studies that identified older populations, and the populations in this group of studies weren't precisely what we were interested in but they seemed to be the best we could do, with over 55 as being a representative population that might be more generalizable to the Medicare population.
I want to highlight two studies that provided within the same study a comparison of older and younger patients. And the conclusion in each one of these studies was that perioperative complications may be increased in older patients, with 12.5 percent complication rate in older patients versus a 5 percent complication rate in younger patients. While the sample sizes are not very big and the result is not statistically

00041

significant, they do suggest that the perioperative complications do increase. The

Kilincer and Carreon studies suggested that there was a difference in the type of surgical procedure that the older and younger patients were undergoing, such that the younger patients were more often receiving instrumentation, with the older patients less frequently, so I think in addition to the complication rate trend, there also seems to be a trend to use procedures that were expected to be slightly safer in the older populations.

And, I have just a couple more slides, I know I'm running out of time. We looked at a variety of ancillary procedures used with fusion to, as an enhancement. The only data that came out of that as being clinically important was the role of BMP as an alternative to autograft bone or PLF. These data were summarized in the Journal of Neurosurgery guideline. There was a little bit of new data that reinforced that, and we basically didn't have any additional findings beyond that. (Inaudible) has to be contained as it can cause problematic tissue spread out of where it's supposed to be.

00042

I want to turn now to a couple of slides just to highlight some of the methodological issues we encountered in our interpretation of this literature. The first set here is taken from a systematic review from Bono and Lee, 84 reports that they looked at. They noted that the studies were nonspecified and, you know, largely not random studies, and there were other failures in the documentation of details of the procedures. So there was some confusion on what fusion criteria were used to determine whether fusion had been observed as an outcome or not, what the program source was, and what the fusion rate was or at what rates fusion occurred. Now we found that these deficiencies were present in the studies we looked at and the more recent studies that they looked at as well. I wanted to highlight a few things that they didn't comment on specifically. One of the problems is that the lumbar spinal fusion studies tend to be identified by a group of patients for whom the inclusion criteria really is being driven by the procedure, rather than being driven by the patient's symptomatic multiple presentation. This creates problems in interpretation of the

00043

literature because we're not quite sure under what circumstances that procedure might be applied to patients, so we favor more stringent criteria-based approaches for defining patient populations and selection of surgical procedures.

An obvious drawback for the purposes of this panel is that the studies we found were almost uniformly middle-aged people and not an elderly population. One of the main goals of our technology assessment was to try to focus on patients who are (inaudible) such as the suit index or the pain scales and in contrast to a lot of the older (inaudible) global assessments for patients of having to use their own judgment about whether they're either better or worse or, you know, measures that are utilized to determine whether fusion has in fact occurred. We do believe that this focus on more well measured outcomes is important in assessing the current literature and in evaluating going forward. And finally, you know, also for this technology assessment, we were interested in the comparison between surgical therapy and nonsurgical therapy. The nonsurgical controls were really not terribly well standardized and

00044

described and could not be easily reproduced, at least in the papers that we looked at. We noted in more detail, and some of the other speakers will comment on this later, that a greater characterization of the disease needs to be done in terms of the spinal anatomy in greater detail, and also use of (inaudible), and finally after looking at the previous treatments that the patient had undergone, and that includes both surgical treatments as well as other kinds, and it should describe what the process was, and for what type of herniated disc report or what type and what duration of frequency.
My final slide is just kind of on the SPORT study, it is ongoing with patients with degenerative spondylolisthesis and who were randomized for fusion or nonoperative treatment. That data has not yet been published, but should be forthcoming next year. And when we made the slide, there were no other planned U.S. studies that would compare fusion for axial back pain to rehab. Actually, since then I have learned that there was some material in a meeting, in fact there may be a slide, and maybe it's getting underway in December.

00045

Thank you. I believe that used up my time and, if I'm not mistaken, the questions are going to be saved until the question and answer period.
DR. KRIST: Yeah, I think we will be probably getting to that around 11 o'clock on our schedule. Thank you, Dr. McCrory. I appreciate your presentation. Next, Dr. Steven Garfin is

going to be presenting for us.
DR. GARFIN: Thank you. I'm Steve Garfin, from the University of California San Diego. My co-author was referenced extensively in the tech report and earlier, Chris Bono, who is here in the audience somewhere, from Harvard. I in particular have a lot to disclose, it would probably save time to tell you what I'm not involved in, but as the chairman of the department I also receive money from government and private industry. Chris Bono is not a chair and doesn't have to take any under the guise of the whole department, any funding, nor a salary, I understand.
(Laughter.)
The most important thing, number one, we are both spine surgeons. Back pain is a

00046

significant problem. 75 to 80 lifetime prevalence, annual incidence of 15 to 20 percent per year, and it affects all of us at some time in our life. It's the primary cause of disability in patients less than 50 years old. It involves 500,000 workers' comp and personal injury cases. I add that group, because many of those patients become Medicare disability patients less than 65. Chronic low back pain is defined as a failure of nonoperative management. The time period is not known, whether it's three months, six months, one year, it varies. 10 to 25 of the cases result in greater than 75 percent of the cost. Again, high in workers' comp, this will fall into the Medicare population if they don't get back to work.
The natural history course is that they do improve, 90 percent improve within two to six weeks and get back to work with normal function. There are 25 percent recurrences. There are only 11 percent of these that are chronic patients, and few consider surgery. Unfortunately, though, the percentages are small because the amount of people who have back pain, the numbers are very large. So, why is there confusion about

00047

fusion? There's a lot of reasons. The pathophysiology of low back pain is unclear. There's limited support diagnostic tools. There is conflicting evidence and confusing evidence, and I will try to go through these. Axial low back pain symptoms are very vague. Normally there is a normal neurologic, they may or may not have a nonradicular sensory pattern, pain may extend from the back or not, it may be localized, it may have functional limitations or not.
The differential, there are a number of

significant medical problems, tumor, infection, fracture, those are relatively easy to pick up on x-ray or clinical exam. What we're talking about today is axial low back pain, whether it's due to the disc, stenosis which can cause back pain, scoliosis which often does not cause pain, it's there but can cause back pain. All of these taken together above the word unknown can cause about 10 percent of what we know. 90 percent of the patients come in with low back pain associated with various sophisticated terms such as lumbago, and we don't know what caused it. It may be structural, it may be microinstability, maybe some chemical irritation

00048

from the disc or nerves, abnormal loading, sagittal balance. It may be hypotic in that the thoracic spine is hyperextended with low back pain, or the low back pain may be functional, so there's many reasons that we don't know. Potential sources could relate to the discs, the nerves, facets, muscles, ligaments, bone and psyche, and we have very limited ability to differentiate. The disc, as you all know, has an annulus on the outside and then the nucleus on the inside. The big nerves we also know about, the cauda equina and the spinal nerve that exits. However, there are smaller nerves like the sinuvertebral nerve, which goes across the posterior annulus, and that annulus is often left in place during kyphoplasty in the lateral side of the disc. There's also the nerve (inaudible) which runs a couple layers within the disc across the longitudinal ligament and into the intravertebral space. So there's a lot of intrinsic sources that can sense pain, none of which do we know or can pinpoint. There's bones, ligaments, joints, muscles all around the spine. It's a very complex

00049

anatomy, any one of these can cause pain, and we don't have the ability, unfortunately, to localize them to one area or another, which adds to the confusion. The diagnostic tools also are fairly nonspecific. History tells us nothing except my back hurts, maybe except for tumors or infections. The physical exam is often normal. X-rays show standard things in everybody over 25 or 35, narrow disc space, some osteophytes. CT is good for looking at bone but usually not the area we're interested in. MRI is good for disc disease, stenosis, herniated disc, so 30 percent of the population who are asymptomatic has the same

findings. And discogram is what many of us use to pinpoint the location of pain, but it is only 70 percent reliable. It's just not specific as a diagnostic tool.
Outcome of assessment depends on the diagnosis. It could be pain, it often is pain. It could be the outcomes of function, are they back to work, are they getting out of bed, do they hurt much, do they take less narcotics, what's their walking tolerance, what's their activity level with the back involved, what do the x-rays

00050
look like. Seek and ye shall find. There's multiple outcome measurements, and to make this list I just took some selected references, they focus on function, they focus on quality of life, they focus on pain, walking tolerance, timed functional tests, VAS has an outcome, NASS has an outcome, every specialty society has an outcome, and there is no consistent use of any or all of them.
There are challenges to low back fusion with the diagnosis of degenerative disc disease. As stated, there's a lack of RCTs, limited Level I evidence. There is increasingly more fusions done despite this. There's a lack of clinical diagnosis and indication in some if not most cases, and as stated, the adverse events are significant, fortunately not frequent, but this is one area, low back pain surgery or low back pain, if patients don't get better, then they often do get worse. FDA is approving more devices based on either tacit or an older work regarding scientific validity and safety. The costs are now becoming important and there are questions of who gets the treatment and who does the treatment.
And all fusions are not the same. As

00051
you heard, posterolateral interbody fusions, posterolateral fusions, PLIFs, ALIFs, XLIFs, corpectomies, taking out the bone, anterior, posterior and combinations. We also use instrumented and noninstrumented, allograft, autograft, cages, and newer devices that keep coming out, and this is just a list of the various interbody approaches on the top four, and posterior fusions on the bottom three.
Just as an example, interbody fusion anteriorly, using cages, bone, autografts or bone substitutes, usually supplemented with posterior instrumentation, enabling us to take out the disc, which may be to immediate advantage, but I like biologics to supplement a posterior cage. The TLIFs and PLIFs, posterior lumbar interbody fusion and transforaminal lumbar interbody fusion look

alike on x-ray. The PLIFs, however, move the body, incline the body a little higher and there is a higher incidence of damage reported. The TLIF is more lateral and is used to create instant stability and tends to cause less neurologic problems. The standard extreme lateral fusion is used less and less in the literature for low back pain surgery, it is usually done after a

00052

laminectomy or for instability as appropriate for scoliosis and spondylolisthesis, and may have a role for facet pain if we knew what that was or could diagnose it.
Instrumentation does increase fusion rate, we saw that in previous slides and I won't go over it.
Why are there increased surgical rates for low back pain? There's many patients greater than 65 years old today, they have better medical care, more are living longer, their expectations are to be healthy. Gerontologists and cardiologists are some of my biggest referrals for spine surgery, to get patients out of bed and back to where they want to be.
We have better imaging studies, we see more, we don't know what it means, but we see it very well. And then there's better education, not just for spine surgeons, because we have more awareness of spinal stenosis and the differential for low back pain. We know the value of fusion in certain select patients, some of which lead to back pain.
We've also had some paradigm shifts in treatment in the last 20 years. We've had better

00053

instrumentation with shifts in the treatment of lumbar fracture, more aggressive treatment of tumors, surgical treatments for spine infections and a greater understanding of the negative effect of lumbar deformities. This all leads to more surgery in the lumbar spine.
From 1990 to 2000, pedicle screws were approved, cages were approved by the FDA, again, tacit or overt suggestions that these are scientifically valid. There's been literature demonstrating increased clinical outcomes with solid fusion, particularly with stenosis and low back pain, increased better fusions with instrumentation which should lead to better outcomes, long-term follow-up or pseudoarthrosis with some worsening outcomes, and we are more aware of clinical outcomes or clinical measures.
In the last six to seven years we have had improved fusion success with devices, instrumentation, DBM and BMP and other tools, with

decreased complications from fusion as we become more familiar with it, and better low back pain outcomes in newer devices continually compared to the old.
Some DDD/low back pain disorders that

00054
you could relate to are better with fusion.
Isthmic spondylolisthesis, compared to laminectomy, the fusion success rate is much higher, and this started us on the fusion track.
Degenerative spondylolisthesis, a number of articles, particularly those by Herkowitz, showed that the progressive deformity of simple laminectomy, 85 percent success with the addition of a fusion. In Herkowitz's series, first he looked at patients who had fusion not instrumented, for laminectomies, and initially if they had fusion at two years, they had good results even though he had 30 percent pseudoarthrosis. When he followed the same patients for two to five years, those pseudoarthrosis patients deteriorated. When he looked at his patients who had instrumented fusions, the long-term success remained at 85 percent.
Degenerative scoliosis followed the same track. Laminectomy for leg pain tends to progress, and if you add a fusion, the overall results are better.
Low back pain is a different beast.
Fusion is not a perfect solution. The Cochrane

00055
report, as we heard today, no conclusions are possible about the relative effectiveness of anterior, posterior or circumferential fusion, but remember, they looked primarily at randomized controlled trials.
There are some studies to look at that point our attention in the direction of fusion.
Turner in 1992 did a mega-analysis with no RCTs but four non-random studies comparing herniated disc surgery, that is laminectomy, with fusion.
The conclusion, for several low back disorders, no advantage has been demonstrated for fusion over surgery without fusion, and complications with fusion are common. However, we don't do fusions for herniated discs and primary leg pain today.
Malter et al. in 1998, complications in the current study occurred more frequently in patients who underwent lumbar spine fusions than on those who underwent laminectomy or discectomy alone. Again, that's not the operation we do or are talking about. And if you adjusted them demographically, there was in fact a significant difference.

Parker et al. in 1996 looked mainly at back pain with a posterolateral fusion, something

00056

we don't do quite so much of today, and it was like flipping a coin, you either get better or you get worse. However, if you take out the workers' comp patients, it went from a bad result to a 92 percent excellent or good result, and his conclusion then was fusion is good.
There are three RCTs that we heard about. Fritzell from Sweden, fusion is better than a nonoperative routine using physical therapy. Brox in Norway, fusion is the same as non-op, even to the extent of cognitive intervention and exercise. And Fairbank gave some criteria for ODIs, he met that criteria and then said oh, by the way, it doesn't work, and this was a cooperative study that's listed as an RCT. So the best thing, the RCTs don't agree, or agree.
Comparing Brox and Fritzell, non-op people in the Brox study mainly did a much better job of controlling the non-op, whereas Fritzell did a much better job of controlling who entered the study, who got the surgery, and fusions came up better. It's unlikely in the United States to have entered patients who have not had any preoperative care, any nonoperative care before they went to surgery. So if you look at the two

00057

studies, Brox is better for the results of non-op care, because they were given uniform non-op care.
Fritzell, however, did a much better job selecting patients, they weren't just loose, low back pain, you get an x-ray and a fusion, they did more screening, they had a better surgical outcome.
There's a number of non-RCT studies showing fusion works for low back pain, but these are the second best available data, and they showed much consistency, 80-plus percent improvement from 2004 to 2005. Moore's is a retrospective study looking primarily (inaudible).
When you compare Moore to Brox and Fritzell, the indications as you go up are much more rigid, much more inclusive in Moore's study than Fritzell's study and Brox's study.
Interestingly, their surgical outcomes are the same. The more, the stricter indications you put on, the better the surgical outcome.
Fritzell's results are on your left, Moore is on your right, and that's because, is it worse because Fritzell used an RCT and looked at them that way, or is it better in Moore's study because he had stringent selection criteria and did only one, as opposed to three or four different

00058

operations for the pain? We know fusion methods differ, we know non-instrumented fusions don't heal as well as instrumented fusions. And increasingly, we are going to interbody fusions if not circumferential fusions, because the fusion results are better and, therefore, hopefully the clinical results are better.

So what Chris and I take from this is fusion is not a good first-line treatment for discogenic low back pain. Fusion can be effective in select patients who have failed non-op treatment, whatever that is, and I would say that the non-op literature is even softer than our literature. And, interbody fusion seems better than posterolateral fusion alone, and maybe that's because we take out the disc if in fact the disc is the culprit.

There are a number of complications both coming from the graft site and the other devices, but these aren't specific for low back pain from degenerative disc disease. Fusion consequences, there are some long-term consequences, loss of motion, but that's the goal, and it's clinically often not apparent. Metal in

00059

the body doesn't appear to be a major problem. Adjacent segment degeneration occurs due to natural history or we've done something to it; most, fortunately, are not clinically significant. The tech report says second operation required; I would caution that required, it's elective and chosen. And fusion disease has no literature-based foundation.

So our assessment is fusion has a role in treatment of discogenic back pain. Better outcomes are with strictest selection criteria including failed non-op care and more preoperative instability. Our criteria for low back pain surgery is exhausted non-op care and exhausted pain in the patient, x-rays, MRIs showing one or two levels maximum degenerative disc disease. Discogram which reproduces the pain, not how it looks, pain reproduction, at the same one or two levels. And the patient is willing to undergo a rigid part in the back.

Our surgery is anterior discectomy to remove the disc, open the disc space and we put in fusion biologics, and posterior instrumentation fusion. With this approach, the literature-based results are overall fusion rate 90-plus percent,

00060

clinical success of 60 to 80 percent, and some do get worse. Is this good enough? Probably not.

Is this the best possible? If we could diagnose better, it could be better. Is it bad in selective patients? What is really needed, with exclamation points, are better diagnostic tools and assessments to target the pathophysiologic cause and assess the pain.
The ideal study for low back pain, to definitively does or does not help discogenic low back pain, but to do this, we need to clearly define the cause of low back pain. It's not possible at this time.
How do we assure all patients have the same cause or treatment before we enter them into an RCT? What are the clear unequivocal and reproducible criteria? We don't have them now.
Non-op treatment, we don't know that, and surgical treatment, everybody differs.
What is or are the ideal measures?
Pain measure, function, performance, quality of life, fusion x-ray, what else? The surgical success is difficult to assess. X-rays don't tell the story. Flexion and extension laterals alone don't tell the story. CT scans are difficult to

00061

assess unless you have 1.0 to 1.5 millimeter cut reconstructions, which are seldom done, and are able to compare this rigorously with the surgery, which is rarely done, and they often don't compare to clinical outcomes.
Perhaps rather than an ideal study, we need a realistic study. It may not be an RCT. Non-op is an issue, but nothing is proven. I think less rigorous data that is available for surgery. Surgery can't be blinded, shams don't work in a fusion, and patients serve as their own control and you can't follow what occurs once they leave the office. This has been shown currently in the SPORT study by a long large crossover. So a realistic study, everybody enters with chronic low back pain, everybody's outcome is measured the same. We do evidence-based nonoperative care, which may not be possible. Some get better, we study them. Some don't, this is not an RCT. If they fail non-op, they get surgery, we do the same outcome measures on all the patients. The patient is their own control and has to be accepted, because RCTs are almost impossible even in a clear herniated disc, as the SPORT trial has shown.
As an aside, Chris and I do 400 to 500

00062

spine surgeries a year, which is a fair amount. Most are with fusion, but only a few for primary low back pain. We feel we know, but we probably don't, when to fuse for cancer, when to fuse for fracture, when to fuse for infection, non-union,

obvious instability. What we don't really know is who is going to benefit from low back surgery and low back pain. We feel we probably don't offer enough fusion to most of these patients, but we just can't quite get a handle on it ourselves. But we do know in the right patients, it can drastically improve their quality of life. And yes, this is a Medicare population. Discogenic low back pain goes to these other diagnoses, degenerative scoliosis, degenerative spondylolisthesis, and stenosis. And then there's the Medicare disability population which starts out just like I said, in the 30 to 55-year range, and then moves on into these other diagnoses of people, I don't want to say elderly because I'm getting there, older than 65. So, thank you for this opportunity and responsibility, and challenge that it has been an honor to present. (Applause.)

00063

DR. KRIST: Thank you, Dr. Garfin. Next we have Dr. Mirza.
DR. MIRZA: Good morning. It's a privilege to be invited to address the panel and the audience this morning. I'm Sohail Mirza. I don't normally address people, I want you informed, but I appreciate the opportunity. And before I begin, let me just tell you a bit about myself. I'm a spine surgeon who has been now in practice for 11 years at the University of Washington. I have a very busy spine practice. I do perform fusion for back pain, also do a lot of tumor surgery, and I think even though the surgery can often be more complicated than the decision-making involved, you have to coordinate with their chemotherapy and staging and radiation and other treatments. I find it easier, I think, and I feel more comfortable in getting an informed decision on a tumor patient that is suffering back pain and is considering surgery. It's a longer preoperative visit, it's more involved, and I think it's mostly because of this variation. It's hard for me to convey to a patient things that I feel they should know and consider before they make a choice, so a lot of my talk will be focused

00064

towards that end, what I feel, the information I feel a patient should know if they're facing a decision about back surgery for degenerative disease.
To begin, my disclosures, this is the North American Spine Society form for disclosure. I get royalties from Synthes Spine that are licensed by our university technology transfer

office. Our department has endowments from DePuy, Synthes Spine, Surgical Dynamics, also from Medtronic, and I hold one of the lab chairs for spinal research.
One nice thing about going after the other speakers is I don't have to review a lot of their stuff. There is a lot of new technology in spine surgery, and I don't think I do procedures today the same way as I trained 10 to 15 years ago. The implants I use, the technology I use is completely different, not just implants that are left in the patient, but also how we get there, the radiographic imaging preoperatively and intraoperatively, and the biological devices that result. There is a lot of new clinical knowledge, and I think Dr. McCrory and Dr. Garfin have done a very nice job covering a lot of the new

00065

information, particularly the more recent randomized trials.
And I think we can all agree that in general for patients with spondylosis, treatment with fusion is better. If you do it, the procedure with fixation, with instrumentation that holds the vertebrae together as they heal, you get a better healing rate, you get a better fusion rate. It's not clear whether that really translates to better pain relief or better physical function. It does add to the surgery, it takes longer to do an instrumented fusion than a non-instrument fusion, it involves more dissection, more exposure of the spine, and it does have a higher complication rate. And most recently, artificial disc replacement may get around some of the problems for the patient, particularly the late consequences such as adjacent segment disease, but that I think remains to be shown.
So first to talk about variation in decision-making. I think all spine surgeons feel very confident about what they practice and how, the complications, but the reality is we disagree tremendously among ourselves in terms of when we

00066

offer fusion. Fact conditions are fairly homogenous across different populations, it's not that we have more degenerative disc disease in the U.S. than other countries, but if you compare western fusions across each other for how often a fusion is done for a degenerative condition, the U.S. is the highest, five to ten times higher than some of the other European developed nations.
Even within the U.S., there is tremendous variation.
There was a very nice article by

Dr. Lurie, who's on the panel, and Dr. Weinstein just this month in Spine, and the rate of spinal fusion for Medicare enrollees ranges from .21 to 4.48 just within various hospital regions of the United States. So depending on where you live, who you see, you can get a very different recommendation on the type of treatment you should get, particularly on whether a fusion is a reasonable option or maybe even a necessary option for what you have. If you break it down into specific rates, there is tremendous variation across states from 1.8 per 1,000 enrollees, to 9.2, and across individual cities, within cities, and it's not the same across all surgical

00067

decisions, and it's not the same across all decisions within the northeast.
If you have a hip fracture, this is a log scale, so this is 10-fold higher and one-tenth the rate. It's a pretty uniform recommendation, fixation for hip fractures is not something that we disagree on among orthopedic surgeons. For back surgery, for fusion there is a tremendous variation. A little bit for total hip replacement, a little more for laminectomy and discectomy, but a lot for lumbar fusion.
Depending on where you live, the regional variation in terms of whether you get a recommendation for fusion is 20-fold for back pain, 8-fold for laminectomy, and I think that is very, very high.
What are the causes of the variation?
Dr. Weinstein pointed out that it could be lack of scientific evidence, although I think you heard a very thorough explanation of the evidence that is available. Financial incentives and disincentives may both play a role. A lot of it has to do with how you train and what you learn. Surgery is still very much an apprenticeship, and for spine surgery, most spine surgeons go through a

00068

fellowship, and I think a lot of how we approach patients and the kinds of things we offer patients depends on our mentors. And I think new technology has a role, how things are developed, how they are presented to both surgeons and patients, and how they are marketed has a role.
I'm going to skip through some of these things. Just compared to other established orthopedic procedures, the increase just in the five-year interval for lumbar fusion for degenerative disease was 100 percent, compared to 13 to 15 percent for hip and knee replacement.
This shows the same slide. Over an eight-year period the rate of spinal fusion doubled, whereas

the rates for hip and knee replacement went up just about 10 or 15 percent. These are population-adjusted rates.
And this is for the Medicare population, I think it's very relevant for this panel. It's hard to sort out the shades, but the graphs are in order here, and lumbar fusion is this bar. So if you look here, the other colors don't vary much, but this little bar has doubled or tripled in height compared to these other procedures, such as discectomy here, non-lumbar

00069

fusion, and other procedures such as laminectomy and discectomy, so lumbar fusion has increased tremendously even within the Medicare population, and it's a big budget issue for this population.
In 1992, spending for lumbar fusion was $75 million; half of the spinal budget is now spent on fusion.
The increase is partly geared or correlated with technology, and not just technology, I would say a lot of other things happened in the mid 1990s, but the rate of fusion went up at a fair pace early in the late '80s and early '90s and then it plateaued, possibly due to litigation involving spinal instrumentation, pedicle screws in particular, and also partly because of newer technology, particularly the disc excision and interbody fusion cases, and I think also some credentialing issues in terms of neurosurgery, orthopedic resident credentialing.
But it has dramatically gone up since then and this increase is not uniform across all populations.
Also, the percentage of cases, lumbar surgery cases that involve a fusion has gone up.
So not only is more surgery being done, but more

00070

of a fusion type of procedure is being done. Now it's over half of the spine operations on the lumbar spine.
The increase is also not uniform across patient groups. The increase is most dramatic in the older patients, particularly those who are 60 or older. Those are the patients that have had the highest increase for fusion, and I think this primarily relates to fusion in addition to decompression for spinal stenosis in older patients. Previously, many of these patients would have received laminectomy; now often they get a combination of laminectomy plus fusion.
Also, the increase is different across diagnoses. Even though we don't offer fusion as a treatment for herniated discs as a primary treatment, it is often done for that indication.

The increase is greatest for the lumbar degenerative disc conditions and some of these conditions, even though mostly we're talking about herniation disc disease, they are referring to small disc protrusions or bulging discs, which are coded as herniated discs. The increase in spondylolisthesis and spinal stenosis is less dramatic.

00071

So what do we know about efficacy? I think this section of my talk is going to be brief. This has already been fairly covered by Dr. McCrory and Dr. Garfin. There are five trials which have asked the lead question, which is, does surgery work better than nonsurgical treatment? And I think those studies are very, very hard to do. And as the SPORT publication showed just this last week, it is very hard for us in the United States to conduct that kind of randomized trial than in other countries. And I think it's not irrelevant that all of these studies were conducted in Europe where each of the nations have a nationalized health care system, and in fact it was more palatable to patients to enroll in the study and possibly get a fusion type procedure earlier than if they were to just wait out their turn on the waiting list. So during the one year or even longer waiting list in Norway they could conduct these kind of comparison treatments. I doubt that we would be able to do that kind of a study in the U.S.
We tried to do a fusion versus nonsurgical treatment for back pain study back in 1988 at the University of Washington, and we had a

00072

very thorough IRB process to make sure that the patients got an unbiased consent. And out of 28 patients who were given the choices by a physical therapist, none chose randomization. Patients either said they waited too long, they were going to go ahead and have the surgery, or they said they had no idea this is what surgery involved and they were not willing to do it, so we couldn't get patients to accept randomization.
These are European trials that have all been published in the last five years. The Fritzell study had a very long enrollment interval. Most of them, all of them deal with chronic back pain. The difference between the 2003 and 2006 Brox study is that one study had no prior, the first publication, patients had no previous surgery, and the second publication, each patient had had a prior discectomy at least a year before the study. The Fairbank study allowed some patients with spondylolisthesis, about 10 percent.

They were very well designed studies, very thorough assessments, a wide array of outcome instruments administered both preoperatively and postoperatively, and I think done very well across the board. These are just various quality ratings

00073

that we tried, and they all scored pretty well. They scored low on this scale because a lot of emphasis is placed on the blinding of randomization on this scale, and like Dr. Garfin said, you can't blind when you're comparing surgeries to nonsurgical treatment. The largest studies, Fritzell is 300 patients, Fairbank 350 patients, these trials I think were underpowered. The Brox studies really did not have enough patients to show the differences they were aiming to compare.

And I'm just going to quickly mention the change in the surgery group from the baseline to the final Oswestry Index was from 9 to 16 points in these trials. The big difference between them is the change in the nonsurgical group. In the Fritzell study it was essentially none, a 3-point difference with nonsurgical treatment and about an equal change in the three European trials.

The Fritzell study allowed natural nonsurgical healing, patients could continue what had already been done prior to enrollment. The physicians prescribed physical therapy, injections, rehab, and various other things as

00074

they saw fit. These other three studies had a very rigorous inpatient rehab program, the kind that's not available in the U.S. These patients were, in the Brox study they called it a back hotel; they signed in for three weeks, they had five to eight hours of PT and lectures and education each day, and then they had follow-up sessions at six months and one year. So with that kind of intensive rehab, they got the equivalent improvement to surgery.

And the change didn't really meet the threshold for what would be defined as benefit, that is, the difference in the changes in the surgical group subtracted, minus the changes in the nonsurgical group were all less than the 15-point minimum threshold for the benefit of surgery.

A fairly high complication rate. These were prospective studies where they actually defined what they were going to look for and measured it as they enrolled patients and followed them. Higher complication rates in the circumferential fusion group than in the

non-instrumentation group, but still fairly high by most standards.

00075

I already mentioned these differences. I think the Fritzell study didn't really specify the nonoperative treatment. The Brox study had small sample sizes. The Fairbank study allowed some patients with spondylolisthesis. They also had dynamic fixation, so they allowed instrumentation in some patients which did not involve bone grafting or fusion, they called this dynamic stabilization, so patients had rods and screws put in to limit the mobility but not eliminate movement, so I think it's less pure of a study in that sense.

So to look at outcomes from another perspective, we have been looking at population-based outcomes in the state of Washington. We had access to the statewide data for all spine cases done as an inpatient procedure. And we looked at the amount of outpatient surgery and a very small number of patients were having outpatient procedures. About a quarter of the discectomies and laminectomies were done as outpatients, but most fusions are inpatient procedures.

So if you look at a longer horizon, a 10-year outcome in terms of the reoperation rate,

00076

those patients that had surgery in 1990 or 1991, what their outcome was at 10 years in terms of needing another operation on the lumbar spine, and we have no way of saying whether this was at the same vertebral levels or adjacent levels, but that risk is not trivial, about 20 percent at 10 years for fusion, but also very high for laminectomy or discectomy, about 18 percent. The fusion reoperation rate is a little bit lower for the first three to six months and then it's higher in the subsequent years, and I think that just correlates to how we take care of fusion patients, you allow some healing time in the first three to six months before you start talking about a repeat surgery.

Breaking that down by diagnosis, spondylolisthesis was the only diagnostic category from the lumbar degenerative diagnostic groups where fusion had a lower reoperation rate. All the other diagnoses, herniated disc, and again, if you look at the total number of patients, 16,000, the vast majority did not have fusion, a small fraction had fusion, but the reoperation rate was higher for herniated discs, about the same for degenerative disease, disc degeneration, a little

00077
bit higher for spinal stenosis.
If you look for differences over time, earlier treated groups, patients who were treated in 1990 to 1993 compared to a more recent group, and certainly in this interval a lot has happened, and hopefully the new knowledge from clinical studies has improved decision-making, new technology has improved the healing rates from fusion, but even when you compare the earlier to the later cohorts, this is a three-year reoperation rate, the more recent cohort has a higher reoperation rate, and that's all lumbar surgery, fusion and nonfusion. And we've also broken this down to diagnosis and by fusion, and even the more recent fusions compared to the prior fusions have a higher reoperation rate.
This is a study we did on injured workers in the state of Washington. Washington has a very special workers' compensation system where they have population-based data and it's very detailed in terms of the clinical information it contains. And we looked at 2,000 patients who had lumbar fusion in the 1990s for a back problem, for a back injury, excluding those who had falls from heights and other work-related incidents like

00078
fractures and dislocations, so these are primarily degenerative disc disease patients. Overall, the point, the outcome that's most relevant to the workers' compensation board is the return to work ability. Looking at surgical technique for fusion, the disability rate at two years was not different than if somebody had a fusion without instrumentation, if they had cages only, or they had instrumentation only without cages, or if they had both, fusion, circumferential fusion.
Complication rates, and these are not multivariable comparisons, 6.2 percent -- these are complications within three months of surgery. Higher if you have cages alone, 12 percent, or with instrumentation. If you have both, higher complication rates for the first three months.
Reoperation rates at two years, and these are again, unadjusted for the covariants, high, 25 percent two-year reoperation rate for uninstrumented fusions, a little bit lower for the cages, and somewhat lower for the combined cage plus instrumentation. This is just a summary of those things.
If you adjusted for all the covariants which we think would be relevant, things like age,

00079
gender, diagnosis, numbers of levels of fusion, medical comorbidity, psychosocial comorbidity, if

you adjusted for those factors, the disability rate does not differ depending on fixation technique, the reoperation rate is not different, but the complication rate is almost twice as high when you have both as compared to just an instrumented fusion.
And I think the safety side, particularly when the efficacy is uncertain, is very crucial, because there are some risks that these patients may be unwilling to take, and I think it's really important for patients to understand that once they're going into surgery, because this is not something you can undo and go back and start over again, because once a fusion is done, it can set off a cascade which is hard to come out of.
I put this slide in just to show how hard it is to interpret safety data. This is from an epidemiology textbook and it just shows, you know, if you have a percentage difference, this is a difference between percentage of complications or adverse effects between two treatments, that is if one treatment has a 10 percent complication

00080
rate and the other treatment has a five percent complication, when you look at 10 percent versus five percent, that's a big difference. To show that difference as a significant meaningful difference, you'd need close to 500 patients in each treatment arm. So again, it's very hard to interpret safety data, you need very large studies, and randomized trials are not good at providing safety data. They don't have enough numbers, they're too expensive, they don't follow patients long enough to even give us reliable safety data, and I think this really has to come from observational studies.
We looked at stand-alone anterior lumbar cages, specifically that was really popularized in the early part of the cage technology in 1996 and 1997, where surgeons felt that just putting in threaded cages in the intervertebral spaces with some bone graft would be sufficient to prevent movement at that site and allow fusion, and hopefully treat any back pain. But that procedure fairly rapidly fell out of favor, without any real kind of single study or particular event that pointed it out. I think mostly it fell out of favor because surgeons

00081
themselves found out that often they had to go back in and do supplemental fusions to get things to heal.
So we were interested in looking at that literature and seeing why were the earlier

studies so much more commonplace than the later studies. And we did a systematic review, we looked at studies that looked at patients with stand-alone cages for lumbar disc disease, studies that had more than 10 patients, and we ended up with 30 trials. And in particular, we were looking at the safety side of things, and I think I want to just point out that the safety data in spine studies is really not well addressed. Very few studies actually pointed out if they looked at complications. Even fewer studies described how they looked at complications, what kind of surveillance they did and how they interpreted, and often the safety data wasn't even reported. So this is just the number of studies. So if you're doing a fusion for pain, wouldn't you think that would be important? For posterior fusion, all 10 studies did it, but overall, 15 percent of the studies didn't even comment on fusion rate, reoperation

00082

rate was addressed less often, and then other important things like blood vessel injury, organ injury, rarely addressed in some of these studies. Complication rates, we looked at the amount of variation in the complication rates that could be explained by the study design, either patient age, gender, type of study done, and what -- this should be I-squared, which is a meta-analysis on spinal fusion that identified the percentage of variation that can be attributed to non-random variation. So these are differences that are somehow contained in the studies that we had identified, but they are not random differences for the rates of variation. There are very high rates of heterogeneity for things like nonunion, for reoperation, less so for sexual dysfunction, infection. These are more clear things when people report them, those are real. The things like nonunion are subject to variation that is not explained by the study characteristics. We also looked at whether financial sponsorship of the study made a difference, and the only thing that mattered was judging fusions. Studies that had sponsorship generally had lower

00083

fusion rates than studies that did not have financial sponsorship, or lower nonunion rates, much higher fusion rates in these sponsored studies than the nonsponsored studies. The other issue wasn't so apparent in other complications such as reoperation or neurological injury. So I think potential financial conflicts do have a bearing, particularly on these

new technology studies which are often done by people having invested in the technology. It is hard to get public funding to do studies that compare fixation and surgical technologies. And I think just to review some literature, in general, if a study is sponsored, it's been shown in various disciplines that sponsorship leads to more favorable interpretation of the data, and that's true for spine literature also.
In orthopedics, where a lot of what we do is use implants such as knee and hip devices and spine devices, it's also been shown that if a study is funded, it is more likely, funded and I should say published, because it could be that the unfunded studies that don't have meaningful results don't get published, but the published literature, a higher rate of favorable results in

00084

the sponsored studies.
And I think this is important. There are a lot of articles in the press about fusions and about back pain, but I think this one is particularly relevant. This is one of the first ones by Jerome Groopman in Boston. And I think the key point I want to make is that the author and the patients, as described in the study, wished that they had known uncertainties about this diagnosis of back pain, the uncertainties of interpreting diagnostic tests such as discography, MRI scans, and the uncertainties in outcomes of fusion, and that this is a critical part of the informed consent for patients considering these procedures.
So to end, I think lumbar fusion rates have gone up despite any real compelling evidence that fusion is a much better procedure than some of the other alternatives. Fusion for chronic back pain compared to rigorous nonoperative treatment like that in the Brox and Fairbank study is probably equivalent. Compared to routine care such as available in the U.S., fusion is probably better. Safety data are limited, advances in technology at least has improved the reoperation

00085

rates with the fusion. And financial conflicts have a bearing. Thank you.
(Applause.)
DR. KRIST: Thank you, Dr. Mirza. Now we're going to turn to our scheduled public comments here, and we have several speakers scheduled. First will be Dr. Matthews, and just, for, you have 15 minutes scheduled, so around 10 after 11.
DR. MATTHEWS: Good morning, members of the panel, ladies and gentlemen of the audience.

I would like to thank CMS and the MCAC for the opportunity to be here. My name is Hallett Matthews, I'm a spinal surgeon from Richmond, Virginia. I am associate clinical professor of orthopedic surgery at Virginia Commonwealth University in Richmond. I do not receive royalties from Medtronic. I am a research consultant. I concurrently sit on three spine society boards over the last several years, and I will continue my clinical advocacy upon the occasion of wearing a new hat in 2007. From our perspective, I would like to discuss several perspectives on lumbar degenerative disc disease as a continuum of care.

00086

I think it's important to understand the difficulty of the diagnosis, I think it's important also to discuss the complexity of the disease, and also the prevalence in the Medicare population. I will also go over a quick evidence review and some new and exciting data in the Medicare population for treatment of degenerative disc disease.
As you know, Medtronic has similar missions to this group and this panel today, to alleviate pain, to restore health, and to extend life, and we believe it is important that we consider how we can improve it.
Lumbar degenerative disc disease is a continuum of care. Not every patient presents at the same disease state as a clinician. I'm going to put my clinical hat on now, and now I'm in the office trying to take care of and analyze these patients, and their presentations differ. The patients that present earlier in the cascade are often earlier, younger presentations and not in the Medicare population. The patients that present later often have multiple disease processes going on and confounding diagnoses, which makes it very difficult to classify these

00087

patients specifically. So it's very important.
It's also important to understand the progression of this disease which varies with regard to genetic factors, aging factors, that we all age differently, injuries both chronic and acute that happen throughout our lives, environmental factors such as using vibration instruments or having a job that involves vibrations, smoking, and different workplace occupational issues, and also associated with other diseases.
It's also important that we look at the specific pathologies that create this pain or dysfunction. This is a patient that entered my

office about three months ago, and you see multiple disease patterns that are significant. She presented to me with leg pain, she presented to her referral primary care physician with back pain, and she presented to the physical therapist with back and leg pain. Okay? So I'm a clinician in the office. How do I code this patient? I've got five different codes that I can use, and oh, by the way, when you stand the patient up and do flexion and extension, L5 is backward on S1, so she has translational instability in addition. So

00088

that in essence describes our clinical dilemma with regard to coding and entering these patients into diagnostic pools.
It's important also that we understand how the patient presents with a degenerative disc pathway. The medical issue is extremely important. To me, that's the most important clinical thing that we can do, to establish where the patient is coming from and their morbidity of the process. Not every patient presents the same to the physician, he may present with different diseases but similar symptoms, so clinically we've got to decipher those. The history is important to talk about, walking, standing, their abilities to perform activities of daily living, their abilities to do certain exercises as simple as walking to the office, and these are important clinical parameters to discuss as to whether or not disease is progressing, and it is the trend of clinical history and presentation that alerts the clinician as to what has to happen.
Whether they present early or late to design their clinical management, we then often begin a patient education self-management program so the patient can learn to live with their

00089

disease, and after nonoperative care has been exhausted do they become surgical candidates, but only if we define a surgical need. We don't go in and explore patients for back pain. And secondly, we use physical exam imaging to confirm our clinical diagnosis, and this is a pattern that we use in the office to help us establish is this patient at risk.
Part of the challenge with degenerative disc disease is that it's difficult to diagnose, it has many comorbidities, and it's rare to define pure degenerative disc disease in the population, and often my patients have four or five degenerative disc codes as they present.
When we look at demographics, in 2005 there were approximately 54,000 lumbar fusion cases performed in Medicare patients of 224,000

patients that had lumbar fusion surgery, which represents about 23 percent of the entire population. When we look at the prevalence of DDD, 722.52 code, it's only about one percent, which means 99 percent of disc pathologies are not strictly from degenerative disc disease that we talk about, they come from other etiologies. It's a very complex diagnosis.

00090

Also too, when we look at evidence summaries, we discussed earlier this morning one RCT and several observational studies, but it's important to realize what's going on with regard to industry and what's happening with getting better intensification of research efforts. In 2002, Medtronic received approval for an IDE study to look at prospective randomized clinical evaluation of posterolateral lumbar fusion, and these were patients with degenerative disc disease that are being treated with fusion, which is considered the standard of care treatment for those that had failed conservative care management. These patients had already had conservative management and now they're going to get a surgical intervention. The control arm was autograft, the investigational arm was BMP-2. We took the cohort of the autograft control group and subsidized that into prospectively selected patients into Medicare and non-Medicare populations.
What was interesting is that there were no statistical differences between numbers of previous surgeries, leg pain, operative time, blood loss, hospital stay, or complications

00091

related to the surgery. However, for many clinical endpoints, there was statistically better outcomes for the Medicare patients. When we look at back pain summaries, those with zero to 20 intensity and duration scores, preoperative evaluations of 14 had greater than 50 percent improvement in the over 65 patient population at 36 months after surgery for failure of conservative care in degenerative disc disease. Also, you see here that the Oswestry Disability Index was greater than 30 points improved in the Medicare population, for again, patients with significant morbidity and disability had values of 50 preoperatively down to 18 postoperatively. That's a clinically significant improvement, which fits the MDA criteria of a 15-point improvement. We also looked at the patients that were significantly satisfied with their surgery and 92 percent were satisfied, compared with 75 percent that were not satisfied.

If we look also at nonoperative care, we agree that nonoperative care is a good control, it's ethical, it's the standard of care, and to do sham surgery is difficult with regard to the risks and complications in the surgery population.

00092

Nonoperative care lets us also look at the true value, as Dr. Garfin phrased it, of the soft data of nonoperative care. We need to evaluate its true value, and using it as a control group for research will help us with that. Medtronic will enter a nonoperative control with the lumbar fusion surgery next month, so we are going to collect that data which is going to be a very powerful study. Research initiatives are important. The Lumbar Spine Study Group is a group of 30 surgeons in 29 centers, which has a 2,000-patient database that's longitudinally managed by PhDx. Over 73 abstracts and publications have been submitted from within this group. This group will also initiate a ProSTOS study, which is the Prospective Spine Treatment Outcome Study Group, which will look at the community-based results from patient outcomes. Now this is a group that is supported by Medtronic, but it is totally independent and retains full discretion as to what it studies, what devices and techniques it studies, and how it reports the data. Some research limitations are certainly there. The study design has to look at the

00093

complex pathology and how difficult it is to enroll pure cohorts of studies. Remember, this is like a changing coastline, this is a changing disease and a continuous evolution over time in particular technology and techniques. I've been in practice for over 20 years and those techniques have evolved dramatically, and we have purified our surgical technique and our indications, and also lessened the trauma and morbidity. There is community practice variability with regard to who can get the care, who has access to the care, whether there is a regionalization of spine care centers that draw from hundreds of miles that centralize the point of service of these fine surgery procedures. It's important that we understand the differences between rural medicine, academic medicine, and private practice medicine, because they're all handled differently. And also too, nonoperative care needs to be better defined. By using that as a control in the future, we will be able to define this both cost-wise and with its efficacy. And evidence-based medicine obviously will continue to

be refined as we get better control for our entry criteria, our control and accountability after the

00094

studies are completed, and also too, with regard to better outcome parameters.
So we would like to recommend that there would be a multidisciplinary work group with CMS, the societies, and the spine device companies, and the overall objectives of this group would be to determine the appropriate research methods for the Medicare population. Let's get the studies right for this specific group of patients. Also, let's incorporate Medicare patients in our FDA IDE studies so we can get that data set going. Also, let's use that information to help develop specific age, specific clinical guidelines that help us and guide us in what is indicated for the Medicare patients. And also too, look at our outcomes. We saw a lot of outcomes presented today with regard to different outcome measurement tools, but none are specific to the Medicare population. Is it time now to have a senior-specific outcome tool, where we could pick up the sensitivities of the improvement of procedures in the Medicare population, that would help design better treatment options for this population.
I would like to thank the panel and the

00095

audience for the opportunity to present.
(Applause.)
DR. KRIST: Thank you, Dr. Matthews. I just want to remind our scheduled speakers to make sure you introduce yourself and state your conflicts at the beginning. Our next speaker is Dr. Gelb, and you requested 10 minutes.
DR. GELB: My name is Daniel Gelb. I am associate professor of orthopedics at the University of Maryland. I have a consulting relationship with Synthes Spine as well as DePuy Spine. This morning I speak on my own behalf, I'm not the official representative of any particular organization at this time. I have been practicing spine surgery for 12 years, practice locally in Baltimore, and have an active elective practice with a lot of Medicare enrollees that form a large part of that practice, and I wanted to give the panel my clinical perspective so when you come to vote later this afternoon, you understand my experience in dealing with this type of problem.
Spinal motion segments occur in the discs and facet joints, as well as multiple ligamentous attachments. I think rarely does degeneration occur in just a single portion of the

00096

motion segment. Spinal degeneration occurs as a natural phenomenon, and at times that spinal degeneration can become extremely painful. Thankfully this is rare, but the problem is complex, as we heard. Pain can occur from arthrosis, from instability, from deformity as well as neurologic compression.

The first question that was posed is, is there an appropriate measure to measure surgical outcome. Well, I think if you sit with a patient in the office, you could come to understand that they have their issues. Their primary concern is pain, of course. A patient who has debilitating back pain or leg pain that makes it impossible for them to perform simple daily tasks, they can't stand long enough to wash the dishes in the sink, they can't walk in a supermarket long enough to do shopping. When the spine's loaded, the pain comes. The patient may be comfortable when they sit, but they're severely restricted in their activities.

I think the customary pain scales are very helpful, functional scales do give us some insight into the degree of disability related to their pain. The questions I ask patients are how

00097

long can you stand, how far can you walk, how much pain medication do you need to get through the day. I think these are the issues, especially for the Medicare population.

The committee asks if there is evidence that surgery improves outcome as compared to conservative care. I think there's a large volume of evidence out there that is difficult to interpret, but in my opinion, there is no question that there is evidence that surgery, especially fusion, is efficacious in treating spinal degeneration. Not all these studies are randomized, but the evidence is there. The studies of Fritzell, (inaudible), are some of the best studies that we have and best evidence.

Clearly patient selection is a critical factor in determining the outcome of surgery and the benefit. Patients with spinal instability or deformity such as spondylolistheses or scoliosis, especially when it's associated with neurologic compression, clearly benefit from fusion surgery if they fail nonoperative treatment. The scope of surgical complications has been well characterized. This is why surgery is generally reserved for those patients who fail conservative

00098

care.

We give patients nonsteroidal

inflammatories, sometimes if their pain warrants, we give them narcotics. We try physical therapy, we send them for injection. But when all these things fail, surgery may be the best alternative for some of these patients. Patients become progressively debilitated when they are relegated to a life of minimal activity. Surgery can restore the ability to maintain a more active life-style and the loss of function that occurs with the loss of mobility.
Different types of spinal fusion techniques are utilized, and this can be confusing to understand which is used for what and which is better. Anterior fusion may be necessary for standard kyphosis. Pedicle screws are the most efficacious way to stabilize a spine with osteoporosis. These techniques are our tools and they need to be utilized in an equalized basis.
In addition, I would add that internal fixation has negated the need for postoperative bed rest and casting. Patients are more comfortable, they can be mobilized more immediately following surgery. It minimizes the

00099

complications related to prolonged recumbency and prevents the deconditioning that occurs following surgery.
Finally, the committee asks that the evidence be extrapolated to the Medicare population, and I see no reason a priori why this should not be the case. I do not think that there is any question that the evidence is applicable to a patient with spinal instability or degenerative disformity, or spondylolisthesis with spinal stenosis. These are common conditions for patients in their 60s, 70s and 80s. Even the rare patient who comes in this age group who has only axial pain and limited degeneration, although that's a rare case, I don't see that patient as someone different from someone who is in their early 60s versus their late 60s.
As long as a patient can undergo the rigors of surgery from a medical standpoint, having already failed nonoperative care, as long as that patient goes through a clear informed consent process and understands the risks and benefits of surgery, there is no reason why they should not be given the opportunity to undergo that type of treatment. To me the available

00100

evidence is clear that this type of surgery is useful and beneficial, and I hope the committee will take that into consideration when you deliberate later. Thank you.
(Applause.)

DR. KRIST: Thank you, Dr. Gelb. Our next speaker is Dr. Guyer, and maybe at the beginning you can let me know how you all plan on doing this. I understand this is a joint presentation.

DR. GUYER: Yes, and I'll explain. First, I'd like to thank the committee for allowing our society coalition to make a presentation today. My name is Dr. Richard Guyer, I'm the president of the North American Spine Society, I'm an associate clinical professor at the University of Texas Southwestern, in Dallas, and I'm a spine surgeon at the Texas Back Institute.

I would like to explain how we're going to do this presentation. I will do the first part, Dr. David Polly will then give the middle part, and Dr. Charlie Branch will give the last part, for the sake of time constraints. But I'd also like to recommend the other members of our

00101

team that helped put this presentation together, Dr. Dan Resnick, who represents the CNS as well as AANS, Dr. David Wong representing AAOS, Dr. Hansen Yuan representing SAS, and Dr. Steven Glassman, who represents the SRS, and Dr. Charles Mick, who represents the North American Spine Society. These six societies represent 25,000 practicing physicians, and it was through their help that this whole process came together, and I believe that this is a landmark cooperative effort that to date has not been seen.

Pain relief is a primary reason that our patients seek treatment for degenerative disc disease. Improved function can occur with pain relief, but return to work is very complex in the elderly populations. Degenerative disc disease is an evolving process with numerous pathologies and with significant variability in diagnostic coding. It rarely exists by itself in the greater-than-65-year-old population. Nonoperative care, as we've heard before, is always the first line treatment. Clinical experience and patient preference are extremely crucial in determining the proper treatment for each patient individually.

00102

When nonoperative treatment has failed, there are clinically significant benefits to the appropriately selected patient from lumbar spinal fusion for degenerative disc disease. Lumbar spinal fusion does not stop the aging process in the remainder of the spine or the patient. It only addresses that particular painful degenerative segment.

In my talk I would like to focus on the clinical perspectives, existing nomenclature problems, and then Dr. Polly and Dr. Branch will discuss review of evidence, response to panel questions, and make our recommendations. As we are well aware, the degenerative cascade is a process that occurs in a normal aged lumbar spine. It can be affected by genetics in terms of age of onset and the diffuseness of it. Most commonly it does affect the lower lumbar spine, and it is a process, not the result of an injury. We also know from more recent literature that smoking can speed up the process. When it comes to the term degenerative disc disease, unfortunately this refers to a number of pathologies, and I've only listed five here, but it includes spondylolisthesis,

00103

spondylosis, herniated disc disease, degenerative scoliosis. Now in clinical practice, we've heard a little bit about how these patients present, but who is the patient that suffers from degenerative disc disease in this population? Well, they are our parent, they are our aunts and uncles, they are our grandparents. They are people that have had, progressively, increase in low back and/or leg pain that progressively debilitates them. They no longer can play with grandchildren, they can't play golf, they can't even walk from one side of Wal-Mart to the other without leaning on a shopping cart. Our population is healthier, living longer, and we would like it to be more active. But once they fail the nonoperative care that we've heard about so often today, we then will carry out further diagnostic studies, and if indeed they are found to be a surgical candidate, then a frank discussion will be had between the physician and the patient. They will weigh the benefits and the risks, and the patient will then make the decision whether or not the deterioration of his or her quality of life is bad enough to

00104

warrant considering alternative interventions. We operate in this population for many different diagnoses. As you can see here, they range from spondylolisthesis, spondylostenosis, to scoliosis, to acquired spondylolisthesis, and a very small segment is degenerative disc disease. When we further break that down, however, that small 14 percent shows that there are many secondary diagnoses as well. Degenerative disc disease is a very, very broad diagnostic category and encompasses

many pathologies. The nomenclature doesn't adequately define all the pathologies that are present when we use that term as a primary diagnosis. So we must look at both a primary and secondary diagnosis to get a better idea of what the patient's true pathology is.
As you can see, less than one percent of Medicare beneficiaries are fused for pure degenerative disc disease, and we've heard this over and over again in the previous discussions. The scientific evidence shows that the studies for the elderly actually are lacking, but the few that do exist show that these patients do improve. The complexity of the diagnosis

00105

compounds study design and current measures lack the sensitivity to account for all nuances of patient pathologies. Even in the best randomized controlled studies such as those from our Europe, and even in our SPORT study, there are methodology problems. There's also problems in terms of time versus the technology and technique, and we've heard each speaker discuss how the technologies have continued to improve with time. There's variations in community-based practice and there's variations in conservative care, and as Steve Garfin said, we really don't know what good conservative care is either. The evidence-based guidelines are slowly evolving and certainly we will get there, but it is a slow process. Nonoperative care is always the first line of treatment and once we embark on surgery, we have to be very careful in how we evaluate these other studies. Dr. David Polly in the next couple slides will discuss the entry criteria for the various randomized controlled studies that we have seen from Europe. There's variability in treatment regimens and also variability of outcomes. I would like now to turn the podium over to Dr. David Polly, who will continue.

00106

DR. POLLY: My name is David Polly, I am the professor in chief of spine surgery in the department of orthopedic surgery at the University of Minnesota. I'm the secretary-elect of the Scoliosis Research Society and currently receive grant support from the Department of Defense. I have received no royalties, have no institutional support. I consult for Medtronic, but I own no stock. I previously have served as the co-chair of a panel for the Department of Defense and Veterans Affairs for the development of clinical practice guidelines for low back pain, and I receive financial benefit from the Army as a retiree.

In addressing the questions that have been posed to the panel today, we've heard a lot of review of the information about the randomized controlled trials. I think there are a couple of key points to hone back in on.
Number one is that in the European RCTs, the average ODI scores were about 40, and they had not necessarily failed a trial on conservative management to date. In U.S. surgical trials, which I will detail in a little more detail in a minute, the average ODIs averaged in

00107

the 50s, so the patients were worse off. They had failed six months of nonoperative treatment to date, so one might even consider their entry ODI of 50 as being a result or an outcome of the nonoperative management to date. So I think it's important to point out that these aren't identical patient populations.
And the second issue, as detailed nicely by Dr. Mirza, is the challenge for us in trying to do a similar RCT in the U.S. in our current health care system. However, in spite of that, we have a good wealth of well done surgery to surgery RCTs that have been done in the United States, and I had the opportunity to review the aggregate data on 1,800 patients enrolled in these trials and then compare these to the published data on other surgical interventions.
So in this cohort of 1,800 patients from FDA IDE randomized surgery to surgery trials, we looked at the pool of SF-36 data, and this would have been compared to published literature data for other surgical interventions. So here is the key point of the presentation. This represents an SF-36 score, a well recognized metric, with this being at baseline disability, a

00108

normalized patient for this age group would be in a different category, and so this amount of improvement across all studies represents a four-times increase of a clinically important difference of benefit.
John Ware, the developer of the SF-36, has defined a clinically important, not a statistically important, but a clinically important difference as being a 5.4 change. All of the studies showed at least that and typically three to four times that. So we're seeing a different effect size in the U.S. trials than we saw in the European RCTs.
When we compare this to other interventions, looking at total hip and total knee replacements, we're seeing commensurate benefit.
Why does this happen? Total joint replacement is

currently considered to be one of the most well accepted highest value interventions for musculoskeletal disease in the United States today. The amount of benefit derived from these patients is commensurate with total hip and total knee replacement.
In here I have clarified a little bit on the issue of mature versus immature

00109

technologies. The increase in spinal fusion rates, I think represents the adoption of maturing technology. The change in total joint replacement with the increases that are continuing represents application of an already mature technology. If we were to go back and look at grafts in the 1970s and potentially the 1980s, as total joint replacement was being improved, we would see a commensurate increase in its utilization compared to where it is today.
But what does this kind of benefit look like? Do more patients get better? On the right you see the healthy population scores, on the left you see the disabled patient that Dr. Gelb described earlier. And the intermediate column shows that we are able to improve. Do we make them normal? No, we don't, but we make them significantly better in their activities of daily living, which is in general what the patients are looking for.
So, does this data generalize to the Medicare population? Well, because of the way the question was framed, in the past we had not generally broken our patient cohorts into under 65 or over 65. As an impetus from the MCAC request

00110

for information, we did this in our study group, looking at patients who had degenerative spondylolisthesis. Why degenerative spondylolisthesis? This is our most consistent diagnosis from which we had good data to do the comparison. So when we look at the ODI scores in the over 65 versus under 65 patients with degenerative spondylolisthesis, we're seeing that they're starting in the category of a 50-point ODI, showing a 20-point improvement, and that there's an exact parallel in the improvement in the over 65 versus under 65 population. When we compare the SF-36 data, and here higher scores are better, again, we see a commensurate improvement exceeding a clinically important difference for both the over and the under 65 population.
What about other publications? Well, Glassman has a paper on this as well which is now in press, which has also shown an equivalent benefit in these people compared to a younger

patient population, so let's look at their data. Again, the patient population is more disabled than the European RCTs, with the aggregate intake ODIs in the 50s. Obtaining 20 points of improvement, doubling the European RCT

00111

improvements. And ending up with a clinically important difference of improvement, and there is parallel benefit for the over 65 versus the under 65.
In terms of the SF-36 benefit, again, we see a similar trend with the over 65 and under 65, both achieving clinically important differences.
So, relief of pain has been talked about as a measure. The current tool, as mentioned by Dr. McCrory earlier, the visual analog scale, and he cited a 20-millimeter improvement on intensity, and a general consensus also is that a 30 percent overall reduction has been accepted across cancer trials and other trials as being a meaningful difference in pain. And all of the U.S. studies achieved this aggregate result.
Function has also been proposed as an appropriate outcome measure, and here the SF-36 reflects this, and I'm showing you the data showing the improvement in the SF-36, with Ware stating that 5.4 makes a clinically important difference.
We think return to work in an over 65

00112

patient population is problematic so we do not feel this is the best outcome measure.
In terms of complication rates, we have heard a lot of discussion about this today, and there is great variability both in the way the reports are conducted and what the information shows, but we do not feel that we have any data on which to draw conclusions on this today, and we feel that it does merit further study. But this issue is not just complications, but complications that affect outcome. Many of the complications as reported are mere transient events and do not threaten long-term outcome, and may not even prolong the hospital stay or treatment. So we think that it's important to look not just at complications, but do complications affect outcomes, specifically in this cohort, and this is an important need for further research.
In terms of long-term sequelae, it's very clear that fusion is a biologic process, so a well done fusion that is solid, is stable and robust, can adapt to the life of the patient. There are no studies at all that talk about, that

have ever demonstrated a solid fusion that has gone on to arthrosis, resulting in a problem at

00113

that fused setting.
And you heard the debate about the issue of adjacent segment degeneration. That is a challenge for us, to sort out the differences between the intervention and the natural history of the patient. And here I think it's important to talk about revision rates that have some challenges for us. Specifically, if we look at the total joint population, they have a high revision rate too. So a person who has had their right hip replaced and develops degeneration in her left hip would be considered a reoperation by the statistics that we've heard presented earlier. I don't think that's a failure of the primary operation, but rather a representation of disease progression in the host, and I think it's important to keep that in mind as well, that the rest of the patient's spine may experience the continued effects of the aging process which lead to future or further additional issues that may need additional treatment.
In terms of trying to understand the pathophysiologic basis, we heard that discussed, and I just included this as a detailed citation about basic science evidence, that there does

00114

appear to be a physical structural difference between pathologic deterioration versus age-related changes. And the key point here is that we hear a lot of information about finding asymptomatic patients. Well, that's the challenge that we face. We're not talking about asymptomatic patients, we're talking about symptomatic patients who have significant changes that are attributed as being the cause of the pain from degeneration, and we realize that's a difficult challenge we still face.
So, the RCTs have been well addressed today, I don't have anything further to add about the information, other than to reiterate that we think the U.S. patient population cohort is different. I want to reemphasize that the U.S. patients have failed nonoperative treatment prior to entering into these trials, and they seem to be more disabled.
So, we don't have key compelling data to date to identify patient characteristics as predictors of satisfactory outcomes, and that clearly needs some further work.
We do feel that there are separate and distinct patient populations, but again, I would

00115

point out that our surgical patients have failed nonoperative treatment, and the challenges of RCTs in the U.S. are that we would have to have patients that continue treatment in the modality that has currently failed them and agree to be randomized to that.

In terms of long-term follow-up, there are challenges in this patient population. I think we have to recognize the issue of frailty and continued aging in the over 65 population. The expectation is that they will over time experience gradually decreasing function. The durability of the intervention to them, we know that fusion is biologically stable. The issue has become one of pragmatic end points and what is an adequate duration of follow-up, and we think two years to date has been reasonable, and there's been no evidence of further deterioration of the treated segment after that two-year period.

And at this point I would like to turn it over to my colleague, Dr. Charlie Branch.

DR. BRANCH: My name is Charles Branch. I am the chair of the department of neurosurgery at Wake Forest University, and today I represent the American Association of Neurological Surgeons

00116

and the Congress of Neurological Surgeons, as chair of the joint section on spinal disorders. I receive compensation as a consultant from Medtronic and I also receive royalties from Wake Forest University relating to licensing in the field of spinal fusion devices. The AANS and CNS have paid for my accommodations and transportation to be here.

I'm going to continue our society presentation and try to summarize our positions in the next few minutes. Obviously this is a complex question, as complex as the field it addresses. It not only involves diagnosis, surgeon skill, body habitus, comorbidities, but the evolution of the techniques themselves. The draft technology assessment has occurred. When looking at the reported experience on anterior and posterior fusion, it did not identify the superiority of one approach over another. This review did in fact confirm the benefit of instrumented fusion with regard to the improvement in the ODI or SF-36 scores, and I would like to sort of suggest that we consolidate this paragraph of questions into two, does instrumentation help, does interbody fusion help. Even these questions might seem

00117

difficult to answer with absolute clarity.

The post hoc analysis of Fritzell, 2003

randomized controlled trial, fusion had a beneficial effect. The authors did not detect a statistically significant benefit from instrumentation, but frankly the study was not designed to detect such a difference.
The lumbar fusion guideline published in the Journal of Neurosurgery by Resnick, et al., this was a compilation of evidence-based medicine reviews in contemporary literature. And here we find that instrumentation is beneficial when there is presence of radiographic instability, kyphosis, or aggregate instability following the decompressive procedure. And all three of these are associated with the condition degenerative disc disease at one or multiple levels, depending on the severity or the multiplicity of the condition.
Glassman in the publication of their multicenter analysis on fusion outcomes with a variety of techniques suggested in the index question, that is the one that they found that an anterior lumbar interbody fusion showed a greater improvement in SF-36 and ODI results when compared

00118

to posterior approaches, but admits, and the reference is noted in the tech assessment, true benefit that you get from one technique over the other is just impossible to determine. When reviewing outcomes of patients with degenerative spondylolisthesis, spinal instability associated with degenerative disc disease, PSF showed a benefit having true fixation for improving fusion rate.
Kornblum noted that there was an increase in patient symptomatology associated with pseudoarthrosis, and that patients with a solid fusion had a better clinical outcome. In another cohort prospectively studied by Zdeblick published in '93, again, shows improvement in fusion rate in patients with instrumentation.
So in summary, I think there is really good evidence that internal fixation improves fusion rates and improves outcomes when carefully applied in a disabling spinal condition with instability, and current techniques do not appreciably increase the rates of complications.
The evidence does not clearly identify a specific fusion technique or approach as superior, but the more current reports are positive for anterior

00119

interbody fusion with recombinant CMP.
And for some of the minimally invasive techniques, the evidence from contemporary carefully designed and heavily controlled FDA IDE studies is that instrumented surgical fusion

treatments deliver improvements of 20 to 30 points on the ODI scale, and we believe that this is strong evidence that instrumented fusion improves health outcomes in appropriately selected patients.
Question 5 asks, what's the level of confidence about the radiographic interpretations? The draft technology assessment that we reviewed prior to the meeting did not really address this topic in any specific review. Our review of 31 studies did not identify a clear correlation between clinical outcomes and fusion rates and frankly, from the patient perspective, it's the clinical outcome that really matters. There's growing evidence that the accuracy of plain x-rays in the identification of solid fusion is in fact weak, and in recognition of that fact we believe that the current fusion studies assessing fusion status for technology in the lumbar spine are going to be optimized with the use of computerized

00120

tomographic imaging. We believe that the correlation of a solid fusion with a beneficial outcome is strong in appropriately selected patients, but the historical literature reported is in fact weak. This is a discordance that we believe should be resolved and will be resolved with further study.
Question 6. Well, we agree with the draft technology assessment that there are no studies pertinent to this question that focus exclusively on the Medicare population or the impact on health outcomes. Yamashita in a 2006 study stratified results by population, but it is difficult to determine significance.
Glassman in his study found that the older population benefitted similar to the younger, and this observation I think is borne out in many of our practices where the 65 or 75 or 80-year-old patient who's physically active, traveling, playing sports, living life fully, while then having had their previous life-threatening conditions treated successfully, CABG, knee replacement, hip replacement, whatever, now presents with a compound disabling degenerative spinal condition that has failed to

00121

resolve with a constellation of nonoperative therapies. And this condition appears to be caused by a clearly identifiable degenerative disability and neural compression, where the patient benefits from a decompressive stabilizing surgery.
We believe there is a cohort of the Medicare population that is truly comparable with

the younger population, and that there is evidence that the benefit of surgical fusion is generalizable to this cohort. The questions that have been posed for the panel are the same questions we ask ourselves routinely in the office. The ambiguity or lack of clarity in the tech assessment prepared for the panel reflects the collective state of the literature on the subject. When comparing outcomes, we all use the analogy of comparing apples to apples or oranges to oranges, and our presentations, our assessments, our literature unfortunately are not just apples and oranges but are truly an ambrosia, a true fruit salad of RCT, FDA IDE device studies, prospective, retrospective, single, multicenter, cohort analysis, historical, and emerging techniques.

00122

But I think there are several distinct flavors that come out of this mix. First, in a painful degenerative condition of the lumbar spine, nonoperative therapy is always the first choice of therapy, but is not always successful in relief of symptoms even after extended effort. Second, the randomized controlled trials available clearly demonstrate that in patients with disabling back pain with an identifiable degenerative condition that fails treatment for weeks or months, there is a clear benefit to surgical treatment. Third, there is evidence that the benefit from operative fusion is comparable to that of total joint replacement or other accepted surgical intervention. And fourth, there is evidence or at least a strong inference that the Medicare beneficiary is as likely to benefit from lumbar fusion as the younger cohort. Do further studies need to be done? Absolutely. The recently completed SPORT trial, unfortunately, was not designed to answer the questions that we're considering today. From a myriad of publications, though, it appears that this study does affirm the cumulative medical experience in spine care that recognizes that

00123

appropriately selected patients benefit from nonoperative and/or operative care, depending on their unique condition. We also realize that methodologic purity is challenging in a large randomized controlled trial studying these conditions. What's the model of the study we might propose? Well, the design of the Fritzell study probably comes close. Its strict diagnostic, ODI, SF-36 inclusion criteria would have to be determined and included. We can't have another Brox or Fairbank

challenge where the entry criteria in these oft-quoted RCTs are way outside what we consider to be the best medical practice in this country. The cohort of patients, especially in the Medicare benefit group, are going to be difficult to randomize, especially the patient with disabling pain who has already failed nonoperative therapy. This is going to require a tremendous recruitment effort, potentially comparable to the scale of the SPORT trial, making the feasibility of this study challenging.

What are some of the practical studies that we could implement expeditiously to enhance our knowledge? Well, there's cohorts of data that

00124

we've already seen, IDE studies that are certainly relevant. We can initiate observational studies, similar design to SPORT, to the observational cohort of SPORT. Clarification or tidying up of coding nomenclature for processes will provide more meaningful analysis of the large population data set which is in MedBar in the future.

In summary, lumbar degenerative disc disease is a real and maybe painful and maybe disabling entity that presents in a constellation of conditions reflected by the ambiguity of the diagnostic coding nomenclature. It's rarely found in isolation in the older-than- 65 Medicare beneficiary population. In most stable degenerative conditions of the lumbar spine, nonoperative care is the first line of treatment. When nonoperative treatment has failed, there is clinically significant benefit to appropriately selected patients for lumbar fusion. It appears that there is a consistency of magnitude of that benefit to patients older than 65 when compared to those younger than 65, but further study is needed to understand the optimal roles for both operative and nonoperative treatment strategies for these patients whose ongoing aging and degeneration are

00125

going to confuse the analysis of benefit of any isolated treatment. Thank you.

(Applause.)

DR. KRIST: Thank you very much. Now we're going to move to the section of our discussion today with the open public comments, and five individuals have signed up to speak. I would ask that if you signed up to be one of the open public speakers, if you would come to the front of the auditorium here, and we will have you speak into this microphone over here. And I remind you to state who you are, your affiliations, and disclose any conflicts of interest that you might have.

For the panel members, you'll notice there are no breaks for bathroom or anything, it's been pretty obvious. Our break is around noon for lunch, so if you need to excuse yourself for a couple minutes, feel free to do so. I'm going to read these the best I can based on what you've written here on the sheet, so I'll apologize if I get things wrong. The first person who signed up is R. Pocelli, and since we have five folks here, I'm going to ask that each of you try to keep your comments to two minutes.

00126

DR. POCELLI: Okay, we'll go quickly. My name is Richard Pocelli, I presently work for DePuy Spine as the vice president of clinical affairs. I was a former academic spine surgeon at the University of North Carolina. I just wanted to talk to you a little bit about our feelings at DePuy as to what we've heard here today and the future as we move forward.

We think obviously DDD is a complex disease, and as we've seen today, there are limited treatment options. I think what we're trying to portray here, and I think some of the docs did that pretty well, is this seems to be a pretty significant disease for patients. In its most severe form it can destroy a patient's quality of life. These patients seek all standards of there, some of them have been listed here, but we know of others as well that do occur, and no present cure exists for these patients. The social and economic costs of low back pain is well documented, and this disease remains a large public health concern, and I think we all agree with that. Spinal fusion surgery has been the mainstay surgical treatment for the last 40 years. For thousands of patients it has

00127

remained and should remain an important option for those patients, and those who have been carefully selected and have failed nonsurgical treatment. But clearly more research is necessary. And we at DePuy are committed to supporting these efforts, working with all stakeholders so that if a surgical therapy is indicated, the right procedure is performed for the right patients at the right time, and by a highly trained professional.

In the meantime, doctors and patients must carefully weigh the risks and benefits of surgical versus nonsurgical treatment and determine the appropriate course of action based on their individual circumstances and the data we have today.

In summary, DDD is a terrible disease

for which options are limited. More research needs to be done and we need to find newer and better solutions. We along with you have reviewed the evidence. We acknowledge that stronger data are needed to permit these patients to make the most informed choices. We are committing to work with CMS to construct clinical studies to clearly identify the superior benefits of both

00128

conservative management and surgery, so that more patients achieve positive outcomes. We want to thank you for the opportunity to comment, and look forward to more discussion and collaboration in the future. Thanks.

DR. KRIST: Thank you. Next on our list is Todd Albert, and then after Todd Albert I have Steven Glassman, so if you could come forward in a queue.

DR. ALBERT: Thank you. I'm Todd Albert, I'm a spine surgeon from Philadelphia. I'm the incoming chairman of orthopedics at Thomas Jefferson University and the president of the Rothman Institute there. My conflicts, I'm on the medical advisory board of three spinal companies, Case Medical, Genesis and Axial Med. I'm a consultant for DePuy Spine. I was reimbursed for my transportation here today by DePuy Spine, but I'm not being compensated to come here, and I felt strongly and wanted to make a couple comments. I was pleased to be able to hear some of the comments made by the societies, which I felt were excellent, and I will just try to expand on a few of those. I think the most important point to understand, for us to all realize is that

00129

when we use the term degenerative disc disease, we're talking about something with multiple diagnoses which importantly do cause low back pain but also cause leg pain. We think the very positive effects of surgery are in degenerative spondylolisthesis, I think the evidence today is excellent, the quality of evidence is excellent. They include degenerative disease in axial back pain with a flattened disc, and I think few people would argue that we have good evidence that we help patients with surgery for back pain, although it's a rare diagnosis in the over 65 population. I guess I would just remind you that many people covered by Medicare are disabled and much lower than 65, and fall into that group of diagnostic categories studied by Fritzell in the Swedish spine study showing the positive effects of surgery.

So we don't have good evidence for the over 65 population for that solitary diagnosis,

not inclusive of degenerative spondylolisthesis, not inclusive of degenerative scoliosis with spinal stenosis, but patients with significant back pain and leg pain. And I think the evidence that we

00130

pointed out relative to degenerative spondylolisthesis, best done by the Beaumont group in the Fishburn and Kirkwood article, following patients, looking at them, at two years showed no difference between instrumentation and no instrumentation in outcomes, but a significant difference in fusion rate. Fast forward when they reviewed those patients at five years, they showed a significant difference in outcomes for patients who had pseudoarthrosis and those who did not. Go backwards, instrumentation led to higher fusion, less pseudoarthrosis and, therefore, better outcomes. So I think they proved that fusion was better than no fusion in those studies, and they again showed the benefit of instrumentation in a randomized controlled trial.
Finally, I would say Dr. Glassman, who I know is going to speak, has done a lot of work in terms of looking at meaningful differences, and has shown that the degenerative scoliosis population again in the lumbar fusion population, very positive effects of surgery.
I was also an author and a site investigator on the SPORT study. That is, as you know, a study that's looking at three diagnoses.

00131

Only the one first one has been published, the largest prospective trial funded by NIH, and there are difficulties. I can tell you that enrolling in that study and having lived with it, there are incredible difficulties in signing up patients to be inside that trial and then have the patient either have to be randomized to nonoperative treatment end up with searing leg pain where they cannot work or live, and demanding to be switched over to the surgery arm and --
DR. KRIST: Dr. Albert, I'm going to have to ask you to wrap up your comments.
DR. ALBERT: Okay. But there are more trials coming for fusion. I appreciate you listening to me and allowing me the time. Thank you.
DR. KRIST: Thank you. After Dr. Glassman, I think the last person is Sean Aclasia, I apologize if I got that wrong.
DR. GLASSMAN: I'm Steve Glassman, from the Leatherman Spine Center in Louisville. I'm here today on behalf of the SRS. I have conflicts including being a consultant, I receive royalties

from Medtronic, I have researched for both Medtronic and Network Healthcare.

00132

What I would like to do is just briefly talk to you a little bit about the study that was alluded to by some of the other presenters when we looked at patients in the over age group. There were 85 patients, it's in press for Spine journal, but not available to you yet. Maybe there's a copy of it that you have received.
We looked at both generic and disease-specific outcome measures, all prospectively at two years, and the ODI improvements of 20 points in that group for the primary surgeries wasn't a surprise, because those are the patients that we see do well, those who can't get to the grocery store, who can't sit on the bench and watch their kids play tee ball, and we know that those patients get better. But in the older patient population, I think the concern of all of us is, are these patients taking a hit in their general medical state in order to get that benefit in their disease-specific improvement in disability.
And in all honesty, I sort of anticipated that that was what we would find, but we didn't. What we found was an improvement not only in ODI, but a substantial improvement in

00133

SF-36 too. And the reason we compared it to a younger group, which was a 50 to 65-year-old group, was that that's the group that people don't look at and say they might be too sick, we shouldn't operate on them. And yet, there is no difference in the general health measures between those two groups. So I think with our added study techniques and our general health, this is a group at this point that we can give the benefit of these operations without overriding risk.
And so I would just like to talk a little bit about complications. One of the studies that we talked about before, my data as well, we reported on the complications in these older patients, but one of the things we looked at in this study was the two-year outcomes as compared to complications. And there was no deterioration in outcomes in two years in the patients who did or did not have complications. Which is not to say that there aren't any isolated patients who had complications whose outcomes certainly will deteriorate. I think what it reflects is that the majority of these patients, you're talking about minor events, urinary tract infections, ileus, which you know is an important

00134

thing to be on top of, but isn't the kind of thing that's really affecting your outcome, and I think as you look at complication rates, you need to differentiate those issues, and I'd ask you to be cognizant of that. Thank you very much.

DR. KRIST: Thank you, Dr. Glassman. And then our last individual signed up for open speakers, Sean Aclasia. I might be getting this wrong so if you signed up and I haven't called your name, come on up. Okay. Well, at this point, we've finished with our scheduled presentations and scheduled public comments and open public comments, and we're going to turn here and give the panel some time before lunch to ask questions of our presenters. So, this is your opportunity to clarify any of the information that you've heard, and try to get details from any of the presenters. David.

DR. FLUM: Dr. McCrory, I would like to start with the evaluation of randomized trials that you formally meta-analyzed in your report, the Fairbank and Brox studies. Can you give us a point estimate for the sense of the impact of the interventions on the patients?

00135

DR. MCCRORY: I'm sorry, the question is what?

DR. FLUM: Was it a formal meta-analysis with point estimates?

DR. MCCRORY: No, we didn't combine them. You know, I guess we entertained the idea, but we felt like with a small pool of four studies, that wasn't a reasonable way of approaching it. You know, the differences between trials and differences in results are apparently what they are, and I'm not sure that I would trust a single synthetic estimate of effect size.

DR. KRIST: Mark.

DR. FENDRICK: Doug, while you're up there, I found it interesting that many of the comments had different interpretations of randomized trials and stated clear evidence from the trials that there is a benefit to surgery over the variety of therapies. I would presume that there were no data that concluded that in the studies you looked at. Is there any reason you can see, or explain the clear difference between them on the overall assessment of the RCT data?

DR. MCCRORY: I think one of the -- I can think of at least two issues. One issue is

00136

that's talking about, you know, the idea of benefit, is the before to after change in the

index of 15 points is clear enough evidence of benefit, or whether it requires a difference between a change that occurs before and after surgery from a change that may occur after rehab. I think fundamentally, that's the thing that led us to our conclusion that was, you know, sort of, we lacked data enough to say that, to convince us of that.
I think the other issue, I have been reflecting more on this since we submitted the report, and that is that there is just more uncertainty, perhaps more uncertainty when you compare conservative management to surgery, both in terms of what therapy and what, you know, what the aggregate of those therapies might be. So I think when, I think that I could almost argue, or someone could argue that the benefits of surgery are clear, and what's unclear is the benefits of the conservative treatment, I don't think that there is a great deal of value to that, or that they are reaching conclusions that aren't supported by any evidence. I think it's just the degree of evidence that we required.

00137

DR. KRIST: Kim.
DR. BURCHIEL: I was interested in one of the slides you showed that, we discussed the short and long-term outcomes of fusion (inaudible, off microphone). And it strikes me, though, that the missing element there is sort of quality of life -- (inaudible, off microphone) -- increased quality of life issues, but that's not mentioned in your discussion. In other words, there is substantial improvement over time. Was that a conclusion?
DR. MCCRORY: I think the purpose of putting up that slide was more to deal with the issue of post two-year outcomes. I felt like, I think our interpretation of the data was that there is some support there, but there was under two years improvement that was clinically significant from that data, and what happened afterwards wasn't as certain, so that's what happened in that area of the curve.
DR. BURCHIEL: So just basically flipping it around, the outcome collapsed towards the same point, roughly at eight years. There may be no significant difference. But what happens in that intermediate time, the time between short and

00138

long-term outcomes, may actually be substantially beneficial to the patient in terms of quality of life in your calculation, but that's never mentioned in your discussion.
DR. MCCRORY: Or perhaps an omission.

There is a lot of data in the intervening years.
DR. KRIST: While you're there, I would like to hear your comments on the differences in the patients in the European studies versus the U.S. studies for the surgical candidates, and what your perception is of that.
DR. MCCRORY: Well, there was one additional comment that I think I agree with very much, and that's that the baseline ODI disability did tend to be lower in those trials, and I didn't actually comment on it in my slides. There was a fair amount of difference in the starting point and the disability level in the studies that we looked at, and those studies tended to be the lowest.
And I heard Dr. Mirza talk about, they were able to produce some data that was far more detailed than what we were able to regarding the details of the conservative therapies in the Brox study, for example, where, I'm not sure where he

00139

got his data, but he must have some source he was able to tap. So I'm not sure I feel like I had a good enough handle on the varieties of the conservative management or in the baseline patient characteristics to be able to comment any more. I didn't have that additional data.
DR. FLUM: I have a follow-up. So one of the themes that emerged in the other presentations was that in the United States, people who failed nonoperative therapy would go to surgery, implying that in the European studies, they hadn't already gone through a period of nonoperative therapy. My reading of it, I didn't pick that up in the European studies, but I wonder if you could comment on that.
DR. MCCRORY: My reading of it is it's a little hard to determine. They did have lower Oswestry scores. It wasn't as clear to what extent they received nonoperative therapy before. It was our impression, I can't quote you line and verse what the specific studies say, I think that's a generally true statement, but I might want to defer to my colleagues, and Dr. Turner or Dr. Richardson might be able to comment on that.
SPEAKER: I don't think there was as

00140

much specified about, particularly in Brox, about three randomization therapies offered, with the presumption that there was some, but they weren't specified very well in the articles.
DR. KRIST: Well, in one of the Brox studies they had surgery a year prior, right? One of the criteria was they had to have surgery a year before, and I assume something happened over

that one-year period between when they had the surgery and when they were randomized for the second study.
DR. FROM: This seems like a critical issue because I think we have to discuss whether there were significant improvements with nonoperative therapy. Three of the four randomized studies showed significant ODI improvements with nonoperative management. The question is whether or not all the benefits were even achieved in the United States, or it would render our U.S. randomized trial not helpful to this point. So I think it's critical that we do understand what data we're getting in these European studies.
DR. ONDRA: This is another qualification about the difference in the studies.

00141

Do you feel that the selection criteria, you commented on that, that many of the studies were procedurally driven rather than patient symptom driven, or do you think the patient selection criteria led to the differences?
DR. MCCRORY: Well, let me distinguish. When you talk about the procedure driven, we were referring to the series of uncontrolled studies, which many of them were. I think in the randomized controlled trials that were done, the European ones, the issue is that the patient population was not precisely described in terms of the nonsurgical therapy that they had, and indeed, some of them were in the context of the post-randomization or a new trial. So we were aware of that data. I mean, example, one of the studies had citations for the nonoperative therapies that were used in the trial, an when we go back and look at those, they're really not very helpful in providing any detail on, you know, a protocol, or any very detailed collaboration of what was expected to be done.
DR. KRIST: Kim.
DR. BURCHIEL: Maybe you shouldn't be the one in the hot seat here, but my question has

00142

to do with the outset, I mean the whole, I think as Steve said, this is such a vital issue, are these populations comparable or are they not, at the beginning of these studies? David, the other David implied that they weren't. My question is, looking at validated measures, has it ever been validated across populations? I mean, some of these results would imply that the fusion rate in the U.S. being five times or ten times more than other countries. Is it possible that we're just a more disabled population in general, and that's

not a mixture or not a result of more aggressive therapy, but in fact of just how the population norms are different. That's one question I have concerning outcomes, has it ever been validated internationally.
DR. MCCRORY: We did do a fairly extensive follow-on evaluation (inaudible) and did a lot of digging and read a bunch of articles, and it has been described as an acute back pain with people who are starting with us on treatment not involving surgery, some around 40, some around 50, some around 60, and we found a range of severity of level of disability. And one of the unique things I think about the Oswestry as well as some

00143

of the other measures was that there were no floor or ceiling factor, it was like a rolling index. And there were, you know, these different populations had either floor effects or ceiling effects, it was an international problem, and one of the (inaudible) on function was that they seemed to be responsive over a wide range of starting points in terms of level of disability. It also seemed to have validity in terms of correlating with patient's improvement. A lot of people have used it with a 10 percent improvement or something, and the score was determined to be a clinically important difference, and I don't think the data was found to support an absolute point change, so I am actually pretty pleased with the properties of that measure.
DR. KRIST: You had mentioned that the meaningful improvement was about 15 for acute pain but it was a lower number for chronic pain. Is there a number that's established on that, or did I hear wrong?
DR. MCCRORY: Well, the data (inaudible) between 10 and 15 with focus on 15 as the level because the FDA selected that one as

00144

their criteria. You know, one might argue that it could be a little bit lower, but not lower than 10. I think some argue that the level for acute pain sure should be higher, I think the 10 to 15 was specific for chronic pain.
DR. JARVICK: I actually have a question for Dr. Branch, which is a follow-up question to a question Mark had regarding interpretation of the RCTs, and it is that the RCTs essentially showed a benefit of surgery, and I was just wondering if you could comment a little bit more about that in light of the other presentations that have been made.
DR. BRANCH: If you take the surgical

cohorts of each of these studies, in the surgical cohorts of all of these studies, there is a benefit. Whether it was 10, 15, 20, 25, 30, 40, depends on what study you look at. Even in the RCTs, when you look at them and compared them to the nonoperative therapy, depending on how you look at it, you may not have a clinically significant difference between the two in improvement, but every single study had an improvement associated with surgical treatment.
DR. JARVICK: So this was just looking

00145

at before and after.
DR. BRANCH: If you look at the impact of surgery, there is, of all the studies that all of us have looked at, there's one study, Brox 2006, that showed a modest, a less than minimally clinically important difference in the benefit of surgery, but that was the one of probably literally thousands of patients. So there's a benefit of surgery. The question, you know, depends on which patient cohort you assign to which, I think that's what we're finding in the European studies.
You know, if you take patients who, the Fairbank study, look at their paper, what are their criteria for entering the study? The uncertainty principle, and this is quoting their language, the uncertainty principle about it, this is as good as you get. Then a physical therapist concluded that patient was studied, and then the patient is randomized.
The Brox study, degenerative disc disease on a plain x-ray, then they're recruited into a randomized study. So that the folks that had surgery, they got better, but the folks who had nonoperative therapy, they got better. But

00146

these are patients that we would not consider, certainly not consider for surgery surgically in the arms that we randomized, and that was the issue we addressed.
DR. KRIST: Do you want to follow up on that?
DR. FLUM: I would like to extend on that. I think there are few surgical procedures out there that have four, I think if we include the Muller and SPORT trial, there are few procedures that have four randomized trials that have looked at comparative evidence. I think in many ways we have a richness of randomized trials if we change this data a little bit. We may not like the results of them, but there are four of them that have been done, and we may not like the way they have been done, and they may have varying

criteria, and we could improve upon them. But I think one of the questions is, how can we plan fusion studies that will be better, and it seems like Dr. Polly himself has taken the idea of another randomized trial off the table a little bit. And I wonder if it's really true that we can't do a randomized trial given the significant number of tax dollars that are spent

00147

on these procedures.
DR. BRANCH: I think most of us are going to say that if we're going to do a randomized trial, then we have to sort of back up. Right now people who at least surgeons are seeing in the office and considering for the randomized trials that we're doing to compare one technique or one technology or one approach to another, okay? For us to do a randomized controlled trial, we've got to back all the way up to the SPORT entry criteria, and we'll see what happens with the degenerative spondylolisthesis cohort in the SPORT trial. I mean, that's out there, it's coming in, so we don't want to get too far down the trail of designing a new randomized study before we see that.
But we know, one, there is a lot of methodology challenge that's going to have to be overcome there. Number two, we have to sort of back up three or four steps and say which patients are, or when are we going to see patients in their diseased state, and then begin the analysis, randomization, treatment process. So it's either, one, we might commit to doing surgery on these people earlier in the process, that's a

00148

challenging thought, or -- I mean, you even basically commented on the thought of sham surgery. That's an interesting thought, but that's a real tough one. So conceptually, back up, treat people earlier, offer, expose, depending on your favorite, surgery earlier in the treatment process, or do like we're doing now and take an observational perspective and see where that goes. I think there are a lot of options and certainly over the next few months and years, I think we as collective health care professionals and government payers are going to try to sort this out.
DR. FLUM: But your interpretation of those RCTs, I agree with Dr. Jarvick, was emphasizing the surgical arm, and I think SPORT along with three of the four RCTs in this area have also shown that in the nonoperative area, there is also improvement, whether or not it's better or worse.

DR. BRANCH: There are many RCTs that are European that there is improvement in both groups. All of those studies, the patient was entered into the process long before, or much earlier in their disease state than we see in our

00149

environment. Most of us who practice spine surgery, number one, see the patient; number two, wouldn't consider entering him into a randomized trial that would include surgery as a treatment arm. So what does that mean? We don't know, okay? But the studies are the studies, no question. But the folks that got entered into the surgery arm earlier in the process got better.
DR. KRIST: Barbara, and then Steve.
DR. BOYAN: I actually have a couple of questions, and I think they're to Dr. Mirza and the gang of three. The questions really stem from the fact that, and this is coming from a non-surgeon who actually was present in most of the FDA panels for which the clinical studies that we looked at today were presented. So I've seen a lot of these clinical studies in a different context.
And what struck me about them was the fact that surgeons don't just do the protocol that's defined, they do the protocol that's defined plus their own little special tweak on it. And there is a tremendous variation indicated. I'm wondering how much of that data is attributable to a variation, to the secret sauce,

00150

and by secret sauce I mean autograft, allograft, DBM, a little bit of each, whatever it happens to be, that makes that surgeon feels like he or she has treated the patient adequately during the surgical operation. That's question one.
And question two comes to deal with male and female. Most of the studies that were presented here were not adequately powered to say anything about males and females, but there are studies in Spine where they were adequately powered, they were prospectively randomized clinical trials where there was at least a hundred males and a hundred females. And looking at nonsurgical therapies, where they found that there are actually statistically significant differences in how males and females respond in therapy. So how are we accommodating that in the conversations that we're having today about the fact that the demography of this group is not worried about smokers, it's worrying about who the main patient is over the age of 70, because most of the guys have fallen by the wayside, that we have taken into consideration what these outcomes are doing

to my side of the population. Over to you.
DR. MIRZA: I'm not sure, but let me --

00151

I'm not sure I can address either of your questions.
DR. BOYAN: I thought you were the one.
DR. MIRZA: I think each surgery is individualized, and that's one of the challenges of doing surgical trials over various medication trials. There is no uniform standardized surgical procedure. It does depend on the surgeon's experience, on their skill, on their particular preferences, and it goes all the way from the initial encounter with the patient to how you frame the issues and how you present the information, to how you carry out the surgical procedure and what is the postoperative course. It is very hard to standardize that.
DR. BOYAN: Is there some way in the analysis data at the end that a surgeon kept a record of what was actually used in the additives, that they then could go back and see how that impacted the outcome?
DR. MIRZA: I think you already answered that. These studies are very unpowered to study the impact. It would be very difficult for them to try to look at the ancillary effects. And the same applies to gender differences. I

00152

believe they are hard studies to conduct, and to look for subtle difference when the main effect is debatable would be hard to interpret. So I'm not sure we have large enough or rich enough data to really look at the specifics of surgical details and postoperative variations, or the specifics of patient characteristics like gender or duration of symptoms.
You know, I think those four trials that have kind of been discussed a lot, but none of them had MRI as an entry criterion. The second Brox study did mention MRI scan, but only as a condition to indicate there was no residual or adjacent segment disease. But they did not specify how many levels of disc disease they were looking for changes, and all these factors are probably important depending on what they see, that is the surgeons, and depending on the specific procedure to that particular patient.
DR. ONDRA: I have kind of an observation that, from Dr. Garfin and Dr. Polly and beyond, everyone talked about the fact that in the U.S. practice, nonsurgical treatment often is what constitutes conservative treatment, so nonsurgical treatment is always done before we

00153
consider surgery. I think that's the part of the behavioral validation which, as Dr. Garfin pointed out, we don't really know a lot of where we are with this.
And so this gets down to how do you study? And the answer is, we've got four studies randomized, which is one of the most important to researchers, but if we have 20 badly designed studies, 30 badly designed studies, it doesn't really increase our knowledge. So my question is, given the difficulty in SPORT with the RCT, if we've learned nothing over the last couple thousand years of science, is that there is more than one way to the truth. Is an RCT, given the difficulty of getting that in the SPORT study, and it would be even bigger doing that in the degenerative disease groups, is that the only way to the truth? Is there an alternative path where you could look at population studies that would obviate the need for an RCT? So before we focus our thought on RCT alone, it looks like Dave is ready to come out of his chair there, is there any other way to get at this answer? David.
DR. POLLY: As a member of the gang of three, I think there are a couple of comments on

00154
that. One is the issue of practical applied clinical trials as published in JAMA in 2003, suggesting looking at effect size in the analysis, and I think that's one of the points that we need to make in the difference between the U.S. data versus the European data. The effect size in U.S. data is roughly twice that in terms of benefit of ODI, versus the Europeans. Now you can argue, are we seeing a different patient population, a worse off patient population, or is our intervention substantially different? However, the effect size is clearly different in well done surgery versus surgery RCTs, so we think it's reasonable quality level one data.
Now, it's not surgery versus nonoperative treatment, but Dan Resnick just reviewed for us and confirmed what we thought, in three of the four RCTs there was no run-in period of physical therapy, so they had zero treatment to date, which is not really a technical paradigm for us in the U.S. in terms of ethical care, at least as most of us think of it, you walk in with your first episode of back pain, receive no treatment, and are randomized for the surgical versus nonsurgical arm. Most of those patients are not

00155
enrolled for surgery because we feel their natural history is so good.

So the question that you're asking, I think, has several different points. The run-in part, which I think in the U.S. has altered our patient population as it sees the treating clinicians. The second one is the effect size analysis, and I think the effect sizes that we're seeing are substantially different, but that's why you're seeing divergence of information. Patient differences, effect sizes, but I think when you have multiple RCT data relating surgery versus surgery, that effect size is no longer an aberration, that now becomes real.
And that gets back to what some of you were saying, that the question is how much benefit, and I think those anterior interbody fusion studies contained adequate male-female differentiation to suggest that the effect size is commensurate in perhaps the best controlled surgical technique studies that we have. So I think the answer is, we can find that information from large, well done cohort information. We can look at effect size analysis and see in the run-in period, is there a difference in our patient

00156

population, and from that I think there is valid information to be gained outside of RCTs that may not represent the patient population that we are treating.
DR. KRIST: For the panel, when we get to our open discussion, I want to make sure that we discuss RCTs versus this cohort data, and I think that will be an important thing to think about. But right now here before lunch, what I want you to do is think about what information do you need to have to inform you for that conversation, so that's going to be very important. Did you have a question?
DR. KIRKPATRICK: Thanks. First of all, I would like to acknowledge my appreciation for all of those that went to great time and personal sacrifice to prepare your presentations. There's also a number in the audience that I am aware of that are missing out on the Cervical Spine Research Society meeting which started actually yesterday, and I don't know that our panel members are fully aware of the professional sacrifice that some people are making. In fact, we have the past president of that organization in the audience.

00157

The issue that I would like to refocus all the presenters on, if they could for me, is the fact that every one of us has talked about this waste basket term of degenerative disc disease; it has included scoliosis, degenerative

spondylolisthesis, has talked about back and leg pain. My interpretation of the questions posed to us are degenerative disc disease and low back pain. In my mind I would like for each of you to exclude all of those that have any portion of leg pain, so that we're looking at the pure, quote, degenerative disc disease low back pain patients, and please revisit the question of whether in your reviews you believe that there is any, some, or good evidence to show that there is improvement. Thank you.
DR. POLLY: That's how we initially, as the societies looked at the question, and so when we saw this tech assessment as it became public domain, we thought that that expanded, perhaps appropriately so, to look at additional evidence. But our focus on trying to find exactly what you were identifying, especially in the Medicare population, is a challenge. And I think that the anterior interbody fusion studies probably best

00158

represent that cohort, in that their entry criteria were degenerative disc disease with back pain predominant.
Now the issue of trying to remove any leg pain becomes difficult, but it was clearly back pain predominant patients, not radiculopathy patients. So I think that is the best data set that we have, and I think how we break out the best information that we have if you're specifically addressing the over 65 population, about the small subset of those patients who were included in those RCTs, and that it is a small number, but in those patients, the benefit that we saw clinically was on the order of 20 points on the ODI, which is a substantial clinical benefit by anybody's definition of it. So I think that is the best information we have on the pure discogenic disease extant with all the other items. I don't know that we can do better than that.
DR. KRIST: Yes, John.
DR. LURIE: (Inaudible, off microphone.) There are two things from the presentations that I have trouble reconciling in my own mind and hopefully somebody can help me

00159

with it. We heard from Dr. Garfin that we don't do fusions for herniated discs, that's not what we're about. And yet from the Medicare claims data, it looks like somebody does lots of fusions for herniated discs. So if we're not doing them, who is doing them?
And the related question has to do with the number of speakers who said if we're going to

study this, you know, the European criteria are too loose, we're much more selective about who we operate on in this country, we make them fail a lot of nonoperative treatment, we don't take them as early, we're talking about a very small subset of people here, and reconciling that with the much higher rate of surgery, fusion surgery in this country than in those other countries. So if we're not doing surgery for disc herniations, where are all -- you know, if we're not fusing discs, where are all the fusions with disc herniations coming from? And if we're so much more selective about who we operate on in this country, where does the rate of fusion in this country come from.
DR. GARFIN: I certainly can't answer for everybody that's a surgeon, and as you know

00160

from either vascular surgery or heart surgery, in many cases techniques vary. We have, somebody showed, I think it was Hal, all these codes that we use, so somebody may code a protruding disc or a disc degeneration, and somebody may code a disc herniation as the same thing. There is no consistency. Reoperations, reherniations, particularly in L4-5, there's a 10 percent, at least, reherniation rate after a discectomy. We tend at L4-5 to fuse those to prevent instability or a third herniation, we tend to go a little longer than 5.1, so that may be some of it, what are we coding.
And two is, there are some indications to fuse in disc herniation, and I quoted from, I think it was Turner, I don't know if it was the same Turner that was on this panel, that was a time past. But as Hal said, we're trained by our environment. There are still many surgeons out there who don't read everything or aren't paying attention, haven't really in their practice differentiated the results between laminectomy, discectomy and fusion, when in many cases they may not need the fusion. But the results are okay, so they have continued to do that for the last 30

00161

years.
When we get back to the SPORT study, there's really nothing new that Jim showed, unfortunately. We've known that from these same Scandinavian groups that we reported today, it's just not new. It's just being reported in a randomized controlled trial so all of a sudden it becomes gospel because it's an RCT, not because it's new, when other studies show the same thing. I don't know if that answers your question, but it goes to it.

DR. FLUM: How about the second part, the second part about the variation in national rates of spinal fusion, given what we're talking about, how there's a higher surgical rate here? Because I find it hard to reconcile with long waiting lists in the European countries, hard to imagine that there's a rush to do surgery without preoperative, the same kind of preoperative evaluations that we do here.

DR. GARFIN: I can't tell you what goes on in Europe, I have some good friends there, particularly in Scandinavia, and they, it does take a while to see the patient, it does take a while to get in. As the SPORT trial showed, and I

00162

don't mean to be coming back to that because not everybody may be familiar with it, but there was a huge crossover rate, and they also had a huge out-of-office treatment that they couldn't handle. So when we say nonoperative in Europe, it doesn't mean people weren't taking over-the-counter ibuprofen and going to their local massage therapist, or they got some of this stuff because they preached it on TV, and that's how you do it. So they probably do have the same degree of nonoperative stuff, but with all our physical therapy that doesn't have any science behind it to any degree, to all the injections we give people preoperatively, does that add anything except time? Are we doing any more other than waiting as long as the Europeans are? The ODI would suggest we hurt more, but we're conditioned to hurt more. A lot of countries after surgery don't take any medicine, they just don't give it to them. We just feel like we have to, it's part of our fifth vital sign, I hurt, we treat. And we tend to say we hurt more than I think other countries say, in individual patients.

Part of the fusion problem, and I had some slides but I don't think I was able to get to

00163

them, so I'll just say it. In spine fusion we have two codes, fuse with or fuse without comorbidities, and that's it. And fusion includes tumor, trauma, infection, spinal stenosis, scoliosis, every diagnosis that we fuse is in that one code or two codes. So when Hal gets up and shows regional variation, or Jim Weinstein reports it, theoretically you're saying there's the same amount of tumors around the country, the same amount of infections around the country, but that may not be. Urban centers or non-urban centers, or referral centers in the middle of Oregon may get more tumors, may get more infections, and therefore, fusion more, so it doesn't mean low

back pain fusions necessarily. We can't pull those numbers out from your data, the Medicare data. There probably is regional variation but our numbers are hard to handle.
DR. BURCHIEL: Can I just ask you one question before you step down? You touched on the issue of discography, which might have escaped some attention. We don't have a lot of tests. And you said it was 70 percent reliable and I just wanted to expand on that, because that is one of the few things that at least some people believe

00164

in, as a measure of potentially what we call discogenic pain.
DR. GARFIN: Discography, for those who are unaware, is putting a needle in a disc with the patient awake, giving them a sedative to take the sting away, and inject a dye. Some people report on discography as the volume of fluid injected, small volume in results in a positive. Some put dye in and get a CT after or a fluoroscope to see if the dye is inside and they call that positive. And others ask the patient, is it pain, not is it the worst pain you've ever had, but is it your pain. And not everybody asks the question, I assume, the same way.
Dr. Weinstein, an author in the SPORT trial, did a very innovative and creative study 15 years ago, where he videotaped patient's faces and asked them to respond, and tried to correlate their facial response with their pain response, and it didn't always correlate.
Dr. Hershey, a researcher at Stanford, has looked at a variety, done a variety of studies on discography, and patients who hurt in general report more pain with discography, the hurt is neck pain. They do a lumbar discography and have

00165

no complaint, and then oh, yeah, that's my pain, I have reproduced my pain. So we depend on patient's response which is the only thing we have, because we're operating on pain, we're not operating on neurologic deficits, we're operating on poor quality of life, so we are dependent on their response. So I use discography to try to correlate it with the MRI, correlate it with the x-rays, correlate it with my gestalt of the patient, because this isn't so much science in a vacuum.
DR. BURCHIEL: Would you say your level of confidence in discography would be intermediate?
DR. GARFIN: Yes, but my level of confidence in operating on back pain without discography is about zero. I'd like to try

something besides looking at an x-ray in David, who may have severe back pain, and you who have no, you may be identical or yours may be worse. So without just MRI or x-rays, you try to put something of the patient into the study.
DR. JARVICK: Just to follow up on that, (inaudible) psychological properties of the patients were extremely important in identifying

00166

which patients would or would not have a positive discogram, and following up on that, we talked a lot about surgery versus conservative therapy, but there is certainly no paucity of preventative conservative therapy here. And one of the things that's sort of striking about the randomized trials is cognitive behavioral therapy seems to be important in a positive result. I was just wondering about the role that cognitive behavioral therapy might play in a conservative therapy regimen, which I don't think is routinely done in this country, and have we really exhausted all the possibilities for conservative therapy.
DR. GARFIN: I was going to say, we don't do cognitive behavior therapy, it is very time-intensive, three to four weeks of almost psychoanalysis and learning and education, and behavior work and exercise physiology. It's a broad spectrum that we don't do or at least don't do well here, probably because the people who do that don't get paid for it very well for all the time that theoretically is supposed to be put in.
DR. KRIST: Barbara?
DR. BOYAN: One last question for all of you, or any of you. Every last one of you has

00167

stated clearly that there is not a lot of information on the Medicaid population, Medicare population, that can be used to, for the discussion, all these trials have been done on a lot of reasonably young, comparatively younger people, but older people do have this, especially women who have now also developed osteoporosis, as well as other things as they age, and drugs like phosphonates are taken to fuse the healing process. If your goal is fusion in the surgery, then you have to, fusion is bone formation, that's what fusion is. Is there any information that you have or that, or information that you feel that you need to adequately treat this population?
DR. POLLY: Tim Kuklow has done a very nicely designed study looking at the effects of BMP in combination with those phosphonates. In an animal model with phosphonate therapy, the healing rate is significantly better. With the addition of RHP and BMP-2, that is overcome and that's now

got some clinical experience beginning to build behind it. So I think that's the best information that we have to date looking specifically at the modern therapy techniques. I don't think we have any good data looking at forte or PTH in future

00168

mass, but for the phosphonates, we're beginning to get emerging data that the RHP and BMP-2 is seemingly involved with more common effects.

DR. FLUM: I have a follow-up for Dr. Garfin. Dr. Garfin, this is the second time you said you have very little faith in doing these operations in the absence of a positive discography. I wonder if you have a sense of how often spine surgeons are using discography before they do operations in the back and whether or not your comments about the faith in the operation working in the absence of it have implications for that.

DR. GARFIN: Let me recover myself a little bit. Number one is, I don't have an answer to how many do it. I do know that some of the, and again, you have to correct me if I'm wrong, the clinical trials, FDA clinical trials for artificial discs, some of them did not routinely include discography. So what I said is me, and certainly not a defined world view, so I can't quite answer that, but most of the studies are done without discography. The results are very close. I mean, you can say there is a range from 60 to 80, and that gets you into the orthopedic 70

00169

percent range of, almost everything we do unless we have a specific diagnosis is about 70 percent. I do a lot of revision surgery, a lot of things that have enough problems associated with them that I don't like a 20 to 30 percent failure rate given up front, this puts a certain amount of bias into the patients, they have to be willing to accept that failure rate. Others say, well, wait a minute, 70 percent of the people are going to get healthy, why not me. It may be with selection that it may be 80 percent. Some people may be clinically better than I am, maybe they can examine a patient and say yes, that's the source of your pain. I mean, we know chiropractors say it's L4-5 or L3-4, C1-2, they seem to know exactly where the pain is. I'm not that good of a clinician, but maybe others can sort of localize the pain to axial mechanical back pain. I'm just looking for another piece of help that I can to get 70 percent into the 80 percent range, if not higher.

DR. KIRKPATRICK: If I could just add to that, as a clinician who spent 14 years in

Alabama before moving to Florida, there was a number of surgeons in that community that would

00170

not use discography, and their patients would end up in the clinics with me or my partners seeing failed backs.
We don't know what the best way to find the painful disc is. At least those that are doing discography are taking as much of a scientific approach as our current technology allows to identify those patients that they think may benefit. Does that make sense?
DR. MIRZA: I think we had some data that primarily dealt with, in the study of injured workers, the highest reoperation rate was in the diagnosis of degenerative disc disease, and among those, the cognitive predicted an even higher reoperation rate. So at least in the state of Washington, these are not selective surgeon practices, but they are, for the hundreds of patients there, they are real patients and real surgeons dealing with very complicated patients such as workers' compensation patients for back pain, discography did not improve disability ratings and did not help reduce reoperations.
DR. ONDRA: Do you think the workers' compensation pool is a valid pool of patients to judge any treatment from? I just know from other

00171

studies, you could pretty much give them a magic wand, and I'm not sure it would help those people.
DR. MIRZA: The numbers are what they are.
DR. ONDRA: But is it an appropriate population?
DR. MIRZA: Well, they are appropriate for receiving the kinds of treatment. Somebody is doing the operation on these various injured workers that you might feel nothing is going to help with, but at least they are getting fusions.
DR. FLUM: And then reoperation is 25 percent, something like that?
DR. MIRZA: Right. And the other point about variation, we can be very particular about our specific skills as surgeons and our diagnostic criteria and selection criteria, but the fact is, I think variability is real, and the studies that we looked at do exclude some factors such as scoliosis, but even more practical than that, we now have a practice where there are 700 spine surgeons, and that means a patient can see two or three of us and get three different opinions. And currently when they go across the street from the university to a private hospital, they're going to

00172

get a fourth different opinion. So these are very real patients with the same findings, same symptoms, same gestalt, but they're getting different opinions.
DR. ONDRA: But I still have a question. Do you think that the selection criteria are something we're trying to get at and so forth, do you think we can get those answers in a workers' compensation group of patients that clearly have disease, but have so many other social factors that have been well demonstrated across the board, can we get any information valid from them about selection criteria for a procedure?
DR. MIRZA: I think the results only apply to the population they're studying. I would not extrapolate to noninjured workers or would not extrapolate to a 65-year-old who is not working. I think all of these randomized trials, it's interesting that we debate about the subtleties of preoperative or nonoperative treatment and all that stuff, but the fact is, the results for surgery in these trials were not dramatic. And then with certain other types of treatment, and I think it's worth exploring what exactly was the

00173

other case, and nothing that I mentioned did anything other than what was published in those papers, maybe they had a paragraph on nonoperative treatment. And in the pain article from Brox this year they do have a couple paragraphs more detailed on nonoperative treatment.
I think it's worth exploring and I think patients need to know that. I think both the SPORT trials and the European trials tell us that this at least is not a cash cow situation. It's not something that's going to create horrible outcomes unless you get surgery, and then with the proper information, these patients can choose for themselves. I think the most important thing out of these trials is to quantify that the magnitude of benefit is going to be very modest, and the patient needs to expect that. If they have unreasonable expectations, if they think they're going to be off narcotics and going back to playing better and more golf and tennis and stuff, then you need to probably moderate some of their expectations. And I think that's what the trials at least tell me the most, is that the expectations need to be realistic.
DR. KRIST: If it's a quick follow-up.

00174

Mark's been waiting to say something for a while.
DR. KIRKPATRICK: I'll defer then.

DR. FENDRICK: Even as we're winding down, I feel that the level of the evidence for benefit in these interventions starts to diminish as the morning goes on. I think it's very important for us to focus on the other part of this, the risk/benefit that four of the presenters mentioned and talked about individual patients. Three is a tremendous paucity of information of the safety of these procedures, at least presented to this point.

I remember seeing in the MRC trial that 10 percent of the surgical patients, 19 out of about 180, had reportable surgical adverse complications. And since Doug, you did mention that, as did the last gentleman, if there is a benefit to me that's modest, we really have to focus, as we have in so many of these interventions over the years, on safety.

And I would like Doug to start, and I can't imagine, although no surgeon in this room has ever had a complication, I imagine in Washington and other states that there are people being harmed from this operation either, probably

00175

not systematically, but we all know that surgery does have risks. And interestingly, the nonsurgical therapies, we don't know if there is a risk and there's never been reported a risk, but I imagine that physical therapy might see a risk now, and it's probably less than surgery, I would imagine.

DR. MCCRORY: Right. The adverse reporting in general is highly variable. We found the single best study for reporting was the FDA reporting in connection with the arthroplasty devices where they had the most complete reporting in terms of the catastrophic events, so a lot of what we described in the report was based on that study.

The comparison rates we had for the other procedures were basically reports which we felt were much less reliable in terms of how these results were ascertained and which adverse events were looked at. The adverse events in general were sort of high, a little uncertain. The duration of the various events were difficult to judge. Like one of the other presenters noted, many of the adverse events that occurred were perioperative and short lived, and don't affect

00176

the long-term outcome, and that certainly appeared to be true from my analysis.

Even though we looked at some of the neurologic complications, sexual dysfunction and some of the other effects that might be prolonged,

we did see that many of those in fact improved over time. There wasn't much data in that intermediate period after the perioperative period to the six-month to one-year data about what happened to people, it's very infrequently described. So it is interesting, but not enough of that data exists.
DR. FENDRICK: And is this always done under general anesthesia, so we get the typical effects of anesthesia as well from the operation?
DR. MCCRORY: That's my assumption, yes.
DR. FLUM: Dr. Mirza, can you comment any more about perioperative events, there are those who are bleeding, there are those who are discharged to, or not being discharged to home after discharge, and specifically in the population over 65.
DR. MIRZA: When we looked at this and chose that population for that particular reason,

00177

because we think that's a more reliable end point than some of these other things, I think it was very hard to classify or actually report complications. I think for a surgeon it's difficult to deal with complication and yet, we do focus on it a lot, we have morbidity and mortality conferences and we all change practices, so I think it's very important information, but I don't think they are terribly reliable for that. The workers' comp group that we looked at, it had better information and we could actually do chart review in the post-op, but that's not routinely done across Medicare patients.
DR. FLUM: How about the instrumentation data in patients over 65?
DR. MIRZA: The data that we have is from day one, that's prior to all instrumented fusions, certainly prior to any fusions done for degenerative scoliosis. So I think that's something we're looking at, but again, it's hard to get recent data on that. Mortality is not something many patients think about. I mean, we do talk to them and often they're so overwhelmed by their pain that they are not paying attention

00178

to that side of things. But I think it's not trivial over 70, certainly over 80, though I don't have any particular data yet.
DR. FACISZEWSKI: As a follow-up to one of Dave's questions, the question about patients who were getting fusions for lumbar disc herniation, how confident are you that based upon the interesting data that you've presented, that

those patients that are having fusions for that diagnosis are actually having fusions for the specific ICD code? In other words, the granularity of that ICD code is well known.
DR. MIRZA: I think there are limitations, and there is no way around limitations of coding errors and administrative data, but it's the one population in which it is seen, even if -- and it's a very small fraction of fusion patients that have the sole diagnosis of disc herniation. We looked at things by diagnostic scheme so that if somebody had degenerative disc disease as a primary code but no other code in addition, but somewhere in their hospital records a code for disc herniation, we would step them up to that, and with spondylolisthesis and spondylostenosis you'll see

00179

the same thing. So the primary diagnosis goes through a hierarchical coding scheme. And I think even though we don't typically do them just for disc herniations, I think, as Dr. Garfin mentioned, probably are diagnoses. When you see disc herniations, particularly third and fourth time with herniation at the same level, the patient's got tremendous back pain in addition to that, the patient's got a lot of collapse at that disc level, in addition to disc herniation. I think I've seen actually among our group and in our community, patients get surgery for that.
DR. FACISZEWSKI: Focusing on the question that Dr. Kirkpatrick had about degenerative disc disease, from an incidence perspective, would you agree with what the presenters have said regarding the impact on the Medicare population of the diagnosis of DDD and fusion?
DR. MIRZA: Could you summarize what you're referring to?
DR. FACISZEWSKI: Well, in other words, do you think it's a huge number of patients that have DDD as defined by Dr. Kirkpatrick? In other

00180

words, have all the massive increase in fusion rates, is that a big component we're talking about today, or is that a really small component?
DR. MIRZA: That's a smaller component. I think most of the increase in the fusion in the Medicare population is related to spinal stenosis, not necessarily degenerative disease.
DR. KRIST: At this time we're going to go ahead and break for lunch. When we come back, we'll finish up with our questions for our presenters and then we'll have a panel discussion.

(Luncheon recess.)
DR. KRIST: I know a couple of folks have flights to catch and what not, so it would behoove us to make some progress on our discussion. When we stopped at lunch, folks were still asking some clarifying questions of the presenters, and what I was thinking we would try to do is maybe for another 15 to 20 minutes, try to finish up with clarifying questions for the presenters, and then we can focus more on the discussion centered around these questions. So, I'll open it up to any panel member who wants.
DR. BOSWELL: Is Dr. Mirza back? Thank you very much for your presentation this morning.

00181

In conclusion, I think you mentioned that regarding spinal fusion and degenerative disc disease, the benefit based on randomized controlled trials is small to none, but it does seem to be detectable. In terms of a randomized controlled trial, do you think that that equates to value, not in terms of the degree of value, but in terms of a yes or no, there is evidence for the value of fusion for DDD?
DR. MIRZA: I think the answer is yes. Almost all the studies show the right direction, improvement.
DR. BOSWELL: Now in terms of specific diagnoses, I think you also mentioned that the outcomes are more driven by the procedure that's done and looking at the outcome, rather than looking at the cohort of patients with a specific diagnosis. Maybe that wasn't you who said that, but I think spondylolisthesis, for example, was mentioned.
DR. MIRZA: Yes, I think there are several trials that have more consistently with degenerative disc disease shown large benefits with fusion and with instrumented fusion.
DR. BOSWELL: Right. And then finally,

00182

one of the other presenters did mention, or one of the other presentations did point out that spondylolisthesis may be a good surrogate in the Medicare patient population for the benefit of fusion. Do you think that that's a reasonable assumption?
DR. MIRZA: I do. We try to, or tend to lump these things as degenerative disc disease, but older patients are more likely to have spondylolisthesis or stenosis, and less likely to have just simple disc disease.
DR. BOSWELL: Thank you.
DR. KRIST: Kim.
DR. BURCHIEL: I have a question for

you. So, with reference to Mark's question of benefit, let's go down there, what's the benefit, what do you mean, ODI, VAS? It's certainly not return to work, I think if we go straight across, there's no definitive difference between the procedures. So we're going to eventually get to this issue of what are the appropriate criteria that we should be looking for, so what is your take on the appropriate criteria for, let's say future studies?
DR. MIRZA: I think it's very important

00183
for the surgeons to report the results of their Oswestry scores, and I think it's practically impossible for patients to understand what an 8-point or 10-point or 15-point difference is going to mean, but I think one of the most important things that could come out of a panel like this is to have more clear definitions of what is successful outcome, and I think it would have to be, in my mind, something of a composite nature, like the artificial disc, where you have some component of improvement and function, probably some measure of pain medication, because again, we are recommending treatment for something that is primarily pain, and maybe some component of safety. The artificial disc studies set thresholds, I think, of 15 points on the Oswestry scale, or 25 percent improvement, plus no major medical complications. But beyond that, I think maybe pain medications should also be a consideration. But I'm not sure what we have currently in terms of, as the randomized trials do not -- where they give on average changes across groups, I really don't think they're as easy to interpret.
DR. BURCHIEL: Maybe this is a question

00184
we should ask generally, valid measures for future studies, but my impression is that taking the menu approach is a difficult thing to do statistically, and (inaudible) but in other words, taking one column, column B, is that ever going to give you a primary outcome measure that's going to be looked at, use of narcotics, VAS improvement of 20, ODI improvement of 15, and going down to check point Z. I guess I'd throw that to the panel, because I think this is not going to give you a measure that's very reliable, even though it may be more real world.
DR. MIRZA: Well, I think you could calculate a percentage and you could come up with an aggregate percentage saying, you know, 40 percent achieved a 30 percent reduction in pain. And then if you add physical function improvement,

maybe you went down to 35 percent or 30 percent. If you added no narcotic use, maybe the percentage would drop even further. If you added return to work, it would probably go to zero. But I think those are the numbers that the patients would find it easier to understand. This is the pattern on aggregate with this surgery, this is the probability we will achieve this result.

00185

DR. BURCHIEL: The problem with that, taking it to the extreme, if you look at patient satisfaction, it's probably one of the most invalid things to look at, what that means in terms of a real outcome. So I guess I, if we went down the road, let's assume for a minute that a randomized controlled trial, that's what we're going to have to have done, showing us outcome measures or, you know, a primary outcome measure, and the primary outcome measure has to be determined. And I just don't know that it could be a mixture, a blend of things, but does anybody have any comments about that.

DR. FLUM: Dr. Mirza, I want to clarify. Your point about the ODI not being an adequate measure is that it's hard for patients to interpret and hard for doctors to explain or interpret?

DR. MIRZA: Yes.

DR. FLUM: That's irrelevant. I mean, it's a validated metric that has good testing principles, good internal validity. If you ask a patient are you feeling better than you were before, that means a million things to different patients, it sounds like what we think it means,

00186

but it means different things to different people. That's why a metric that attempts to measure disease-specific quality of life is very meaningful as a measure of functional status that has been well validated, and I'm not sure I understand your problem with the ODI.

DR. MIRZA: I didn't mean to take anything away from a validated disease-specific measure, and it's very useful, in fact. But when you're trying to judge whether a treatment is successful or not, I think it would be hard to use the ODI change. I think all these randomized trials show somewhere in the range of 10 to 14 points, or 10 to 13 points improvement. I'm not sure I can really convey that to a patient as a reasonable expectation. I think the greater the magnitude, you know, if you say a 10-point change in the ODI leads to this much greater satisfaction, I mean, there are some studies that tried to look at what is a clinically important

difference, and actually maybe Dr. Glassman can comment in that area, but I think we need a simpler measure of success, so when a patient comes in and is trying to sort out whether to continue nonsurgical treatment or choose surgery,

00187

we need to be able to convey to them what their probability is, and if they're not sure of that, then they couldn't make a consent.

DR. KRIST: I see kind of two issues on the table. I mean, one is for this panel or Medicare or some other organization to decide if an intervention improves outcomes, and I think a validated instrument is a good way to do that. Another is what a doctor would say to a patient, and that's a whole different set of things, and I think it will be important for us when we kind of do our discussions to talk about what should the outcomes of this ideal study be, when we do that for Question Number 3. So, Kim, you had a question?

MS. KUEBLER: Yeah, I wanted to follow up with Dr. Boyan's comments earlier about comorbidities. We know patients with chronic pain also have depression, and unfortunately, we also know very well that depression and pain also contribute to poor function. So I mean, are those considered, is depression even considered in any of these follow-up trials?

DR. MIRZA: The European trial certainly measured depression. The results

00188

haven't really shown what the effect was but it, they said it was not different for the treatment. But from other literature, it's clear that depression or other psychological comorbidities have a profound effect, probably more so than any imaging findings, probably more so than the diagnostic pattern.

MS. KUEBLER: Thank you.

DR. KRIST: I would like to hear from some of the investigators of studies about their experiences in trying to enroll patients in RCTs, because we heard earlier that an RCT might not work in the United States. And some have done, and I would find it helpful to hear about experiences in doing this.

DR. GARFIN: Dr. Albert, who talked before, is he still here?

DR. ONDRA: Don't be shy.

SPEAKER: The SPORT study is the enrollment problem that you're looking for, where patients, you know, were having to -- hadn't had a lot of conservative treatment anywhere, and the doctor's decision to assign to the conservative

versus the surgical arm is a difficult decision to struggle with in those patients. In our fusion

00189

patients, they've had all conservative treatment and I think that enrollment is going to be difficult.
We've done a number of randomized studies that are not conservative, or not versus conservative treatment. I've done the BMP studies and we've done an INFUSE versus bone graft study that we randomize people to things where you're telling them, you know, I think these things are exactly the same and you haven't had either of them, I think the challenge is randomizing someone that has had their PT and their medication and their blocks, and to say you can do that again or you can have surgery is going to be a real challenge.
I do have a suggestion, I don't know if you'd like it or not, of a place I think you could randomize. And that is, we see a lot of patients clinically who have had therapy and who have had medication, but who have not had injections. And for the older patients with stenosis, that is a viable treatment that, you know, we generally try before we send people on to say now you ought to have surgery. And I think you could find a substantial cohort of patients who have had a

00190

fairly long period of therapy and medicine, but have not had blocks, because those patients come to surgeons at that stage. And you could randomize them to you're going to go on and have blocks, because only people who have pathology that is clearly surgical pathology, you know, probably just spondylolisthesis and stenosis, but you could randomize them to either now you're going to have blocks or you're going to have surgery, because those are both viable options that people would be recommending. And then if they cross over, they have their blocks and failed and crossed over into surgery, that would be a failure of the blocks. If they did well with blocks and didn't have to go on to surgery, that would be patients starting at a reasonable centrist position with a nonsurgical or medisurgical. And I'm sure you could find faults with that, you know, setup, but at least it is an effort to randomize people in a way that I think would fit into how we practice patient care.
DR. FLUM: I have a quick follow-up.
In the SPORT trial, it really wasn't a nonsurgical arm that they were randomized to. They either got operation or whatever, whatever the docs want,

00191

usual care. And as we saw, I think in one of the earlier randomized trials, just continuing what's failed obviously is not going to give the patient, it's not a beneficial intervention. And I wonder, as a spine surgeon, whether or not a nonsurgical intervention, truly intervention, we talked about cognitive behavioral therapy or any of the things that have been studied in Europe might be worth trying here.
SPEAKER: I want to take exception of the description of efficacious. The trials that we looked at, you know, Fairbank and Brox, those patients had 75 hours of structured therapy, far more than you would ever approve or pay for here. And they ended up with an ODI improvement of 3 to 13 points. I mean, that's a lot of input for not a lot of benefit, you know. And I think it would be important to have structure in the conservative treatment in anything that we formulate. I agree with you completely about that.
DR. FLUM: The cost of the nonsurgical intervention is relative to the cost of the surgical intervention, so just because we don't have it yet, is there a possibility for a nonsurgical intervention that is something other

00192

than what we seem to be doing? You said something about injections, and I wondered if that 75-hour approach might be something that would work in the United States.
SPEAKER: I think the patients that we see in the Medicare population who are typically coming to us because they say I can't walk through the grocery store, if you tell them, I know you've done therapy before, I know you had medicine before, I know you had injections before, but now we're going to do a cognitive therapy and more intensive physical therapy, I think you're going to have trouble holding people in that arm, and I think you're going to have the same problem you saw in SPORT with a crossover in one direction, you know, people in a much greater magnitude saw it in the SPORT study.
DR. JARVICK: But what if you were going to combine the injection with the new therapy?
SPEAKER: I think that would be reasonable. I think if you're giving them something new, the injections they haven't had before, you could entice them into a program that you could do as something else along with that. I

00193

think that after the surgery, in the surgical arms of those studies, people got no rehab. I think

they would do better if you did concentrated rehab postoperatively, and people should have that. But I think if it's something that you could sell them on, that I'm having a new treatment, and the injections might be an option, and you could add the therapy as well, if they accept that potentially.
DR. GARFIN: Could I make a comment? I wasn't involved in this RCT but I was involved in another one on kyphoplasty years ago, that really was almost experimental at the time. We didn't have (inaudible), nor did we have any data about the nonoperative arm. I mean, there were very few people back in '96 or '97 who were using Fosamax or Actonel, and kyphoplasty was brand new. So we identified 25 centers just to do kyphoplasty, set up a program, randomized it to the nonoperative arm or failed therapy in time, or kyphoplasty. After two years, we had less than 30 patients enrolled because they could go across the street and get vertebroplasty, which didn't require state approval or anything else. So this is going to be roughly the same thing.

00194
My concern about developing a study of fusion for low back pain, if that is still what we're talking about, unless we're talking about spinal stenosis now, but we're still talking about back pain, why should they enroll in the nonoperative arm when they can go across the street to get the operation if they feel they need the operation. It's not even a tool or device or a hook to bring them in on. We had that hook and we still couldn't get them in. Back pain is a different beast. People don't like it.
DR. JARVICK: I too would just comment about vertebroplasty and kyphoplasty because I am currently involved in an ongoing randomized trial of vertebroplasty versus a controlled intervention, and this was not an industry-sponsored trial, not a sponsored trial, where we were experiencing a certain amount of futility early on in enrolling patients, and there was an article in the New York Times about we had only gotten three patients into this trial after several years of trying.
Well, up through the last year, we're actually now close to 60 subjects who have been randomized, partly in this country, but most of

00195
the subjects have come from the U.K. and from Australia, and I think that there is a tremendous cultural difference between this country and other countries as far as the attitude of both patients and physicians as far as the willingness to

randomize and to give up their freedom of choice about their treatment. And whether you think of people as altruistic or whatever, that difference is real.
And it's not to say that it's not insurmountable, and somebody referred to the principle of uncertainty, some also call it equipoise, because they really have to believe, the surgeons, the treating physicians and patients have to believe in their heart of hearts that they don't know the answer to this question. And we are able to randomize patients at Mayo as one of the sites, with an equal rate of any other sites in the U.S., but that's only 20 percent of eligible patients being enrolled in the studies. In the U.K., 80 percent of eligible subjects are typically enrolled in the studies. So there are real hurdles, real differences, but they are not insurmountable, I would say, in this country if the study is designed right and there is the right

00196

hook.
DR. KRIST: I was just reminded, if you come up to the microphone, just state your name again for transcription.
DR. RESNICK: My name is Dan Resnick, from the University of Wisconsin, and my disclosures are that I participated in the cervical disc trial with Medtronic and we had no problem whatever getting people to sign up for that, because they all thought it was the best thing to do, people were eager to have that, and they were disappointed if they weren't placed in the group that actually got the prosthetic device.
I wanted to mention a couple methodological concerns regarding the Fairbank and Brox studies which may help to elucidate some of the problems that we have in the United States in terms of getting these studies done. In the Brox studies, patients all had x-rays and back pain, and had not had any sort of conservative care at the time of entrance into the study and prior to randomization.
They were entered into the study because they were given the promise that if they were randomized to surgery, they would have the

00197

surgery within three months. Otherwise, they had to wait 12 to 18 months to have their surgery through the regular channels. In the United States, people aren't going to wait 12 to 18 months to have the procedure done when it's been determined they're a candidate for that procedure.
So their goal was the ability to get people into treatment sooner than they otherwise would have

been treated.
In the Fairbank study, similarly, patients had had no pre-randomization therapy, and the only patients included were those in which the surgeons weren't sure that these patients were going to get better with surgery or not, so those were the only patients who were entered. Despite that, despite the fact that they had an intention to treat analysis with almost a 30 percent crossover away from conservative therapy to the surgery group, they still had statistically significant improvement with surgery compared to the nonsurgical group in terms of the back and leg pain.
The other thing that I wanted to mention, all these studies are done looking at patients who presented to have surgery for low

00198

back pain. These are the type of patients who many would get a discography on, but these are not the patients that are the Medicare population. I can't remember ever doing a discogram on a patient with spondylolisthesis, it is just not part of the equation. (Inaudible) anterior post, and I've seen it in patients who have other problems and are having fusion as an adjunct to their treatment of other problems, and as part of salvage procedures on people who have already been through the whole gamut of (inaudible).
On the cognitive therapy part of it, the European studies did a pretty good job in addressing depression, and it turns out that both cognitive counseling and surgery both had an almost identical effect on the depression scores, except for avoidance behavior, that was the only statistically significant difference from the Brox studies. But that's part of the patient selection, and we don't stop doing physical therapy once we operate upon them, the physical therapy continues, that's part of the ongoing care of these patients.
DR. FLUM: In the early 1990s, thoracic surgeons felt very strongly that lung volume

00199

reduction surgery was very effective in dealing with COPD, and at several centers it was being done often with varying mortality rates. Medicare made a decision to only cover that procedure in the context of the clinical trial, randomizing patients to surgery or pulmonary rehabilitation. Obviously that trial had no problem generating patients because for patients it was the only way to get the operation performed. Do you think there's enough equipoise based on these four randomized studies that we've seen here to show

really no significant dominant clinical effect of fusion surgery or one that shows a more dominant effect?
DR. RESNICK: I think if we were studying a 40-year-old patient with low back pain and an abnormal x-ray, we could make that assumption.
DR. FLUM: And less so on a 65-year-old because you said none of them had to do it?
DR. RESNICK: Well, no, because we had significant evidence that decompression, that patients with back pain in the Medicare population are a subsection of the one percent that we saw. The vast majority of fusions that are performed in

00200

the Medicare population are performed as an adjunct to another procedure or a stabilization procedure because of something that happened before, or neurologic deficit.
In the case of a sham surgery procedure, I think there are also ethical considerations when you consider that all surgeries are completely elective. If someone has end stage COPD, that is a very important thing that you need to know and that is probably worth knocking some people off to find out what that answer is. We're not saving lives here. We're just getting rid of back pain and leg pain.
DR. FLUM: The thinking (inaudible) same, almost the exact same model where there is a condition, a quality of life improvement is the goal, and there's a surgical and nonsurgical approach.
DR. KRIST: I can sense we're shifting our conversation from questions to the presenters to our panel discussions, so I want to see, do folks have further clarification questions for presenters?
DR. KIRKPATRICK: I have a very quick clarification that I want to point out happened

00201

here just now, okay? The questions that we're reviewing are based upon spinal fusion for the treatment of low back pain secondary to degenerative disc disease. In my mind as a spine surgeon, we have excluded spinal stenosis, we have excluded spondylolisthesis, okay? So we don't get that confused anymore, let's make sure that we're talking about that fraction of a percent that he said was in the Medicare population. So it's a small percentage of the overall degenerative disc disease population we might be doing a fusion on.
DR. RESNICK: It is a small percentage.
DR. KRIST: Regardless of the percentage size, that's the focus of what our

deliberations will be on, so that's the condition we will be talking about, low back pain, and not spondylolisthesis and spinal stenosis and those issues.
DR. PHURROUGH: Just to clarify too, the purpose of this meeting is to provide information to the community and not to the people who come to the meeting. Because in general, those of you who are here are appropriately selecting patients who need the procedures that you do. That's not necessarily the case for the

00202
broad Medicare population where the people with low back pain who are getting spinal fusion aren't necessarily a small percentage in the broad Medicare population.
DR. RESNICK: I would disagree. I mean, the data shows it is an isolated fraction, it is a very small proportion of the Medicare population.
DR. PHURROUGH: I think we would differ on what we define as a small isolated fraction.
DR. FACISZEWSKI: Maybe we could have some clarification. I asked the question about it before, and I believe that statement is based upon the administrative database, I would assume. And my question, and maybe you can readdress it, how confident are you in that administrative database as it relates to either a very specific group of patients or even any individual patient? Based on the administrative database, it shows the prevalence or, in this case the incidence of surgery and the reason it's performed in the Medicare population.
DR. MIRZA: The data I presented relating to fusion have to do with State of Washington operations, not Medicare databases. So

00203
this is administrative data on hospital discharges from community hospitals and academic hospitals in the State of Washington. And in those categories, certain things stand out that the rates have gone tremendously, have increased very significantly in the degenerative disc disease indications and they have also increased in older patients. In the older patients, it is mostly spinal stenosis, and in the younger patients, degenerative disc disease.
DR. FACISZEWSKI: So to answer the question which I think was queried, it is in fact true or not true that you believe that the reason or that the source of the increase in patients being operated on in the Medicare population is because of back pain and degenerative disease, would you agree with that statement?

DR. MIRZA: I don't know what to say about the Medicare population, I'm not sure.
DR. KIRKPATRICK: I'd just like to clarify this issue for the panel, because it may not be clear what we end up doing as surgeons. When we have an admission of a patient with, say, degenerative scoliosis, and they have radiculopathy or claudication, they're going to

00204

have on their coding sheet on the hospital records an ICD-9 for degenerative disc disease, for scoliosis, for neuroclaudication, and if it's seen in the leg, radiculopathy. So they have all four of those that feed into his database. So the problem for those patients, he can't tell us which, if we're going to throw out one of those affecting low back pain, he can't tell us what proportion of those were isolated degenerative disc disease and what proportion were combined with other diagnoses; is that correct?
DR. MIRZA: That is correct, there is no way to make any statement about symptoms from our database.
DR. FLUM: But in the absence of the radiculopathy or in the absence of the sciatica and the other codes you just described.
DR. MIRZA: I mean, I don't know how people code, or exactly what the subtleties are with various codes, but in general, for spinal stenosis, there are about an equal number of patients that are getting decompression as well as decompression plus fusion, and I'm not sure how that decision is being made. Now it could be that the surgical procedure is so extensive that a

00205

fusion is deemed necessary to maintain alignment of the spine or improve alignment of the spine, it could be that some of those patients have low back pain and other patients appear with claudication. But patients with the same diagnosis are getting both treatments, fusion or decompression without fusion, and we can't tell that from the administrative data.
DR. KRIST: And one helpful thing for the panel, I mean, our purpose is to discuss the state of the evidence. And so, you know, as part of our discussions we can bring up, for these indications we believe the evidence is good, and for these indications we have issues with it. So that will be something to bring up in our discussions as we're talking about this. And as we go through and we vote on these topics, we should talk about the indications, I think a discussion point is which patient criteria or which patients benefit best from these. But when

focusing on the numerical voting, we should narrow it down and make sure that we're all considering the same topic as we have been discussing here. So, yes, Jon.
DR. LURIE: I just want to make sure we

00206
have some clarity on how the European randomized trials are being represented. So, I have heard, a couple people have said that the European trials involved people very, very early in the disease, and two, that they had no nonoperative treatment. And I want to know whether anybody actually knows that those things are true or whether they're just an impression, because the fact is that the trial did not record in their manuscript all the details of the nonoperative treatment that people received, and in the Brox and Fairbank studies, these people had to have symptoms for a minimum of a year, or two years to be eligible, so they weren't early. They may not have gotten extensive nonoperative treatment, but I think the answer is we don't know what nonoperative treatment they got. I mean, I know the Europeans are hardy folks and not quite as demanding as Americans, but it's hard for me to believe that they had two years of back pain with no nonoperative treatment. I think that is a stretch.
DR. RESNICK: All I can report is what's in the papers. I happen to have the papers here, I would be happy to share.
DR. LURIE: That's the problem, they

00207
didn't list the criteria.
DR. RESNICK: Symptomatic for two years and required having failed preoperative physical therapy. That's in the papers.
DR. LURIE: That doesn't mean they didn't have it, that means that they --
DR. RESNICK: I'm not arguing that point. Another significant point from the papers, they were primarily based upon plain films, as well as the presence of disabling pain for a year. That's what it says in the papers.
DR. LURIE: But I just want to make clear that that doesn't mean that they had nonoperative treatment before they entered. We don't know.
DR. RESNICK: It wasn't required by the study.
DR. FACISZEWSKI: I have one final question. Are there any social differences between Europe, particularly the Scandinavian countries, and the United States that made RCTs easier and/or patient selection different than we may have here in the United States?

DR. POLLY: I would like to comment from my perspective of having been in sort of a

00208

captured health care system previously and now being in a different health care system. When I was part of the Department of Defense and the patient beneficiaries had a defined access, oral pathway to health care, we were able to do a series of randomized trials that were generally intervention trials, randomizing intervention A versus intervention B, but the patients were generally accepting of that. Also, I had incredibly long wait times for surgery, at one point nearly one year, and my total joint column at one point got to two-and-a-half years for joint replacement surgery once the diagnosis was made. And in some of those patients who were waiting two-and-a-half years for a hip replacement, they'd have GI bleeds, so I think there is morbidity from nonoperative treatment.

But in that health care system, a mechanism to increase access appeared to alter patient behavior would be my interpretation, as opposed to my current system where if they can't get an appointment next week, they're very upset about that and want to know why. So I think that is representative of a different sense of entitlement or sense of affiliation with the

00209

health care system that may lead to a difference in behavior, so I think there is something to that. Trying to quantify that becomes very difficult.

And I would just share a final point. We attempted to develop a study comparing operative versus nonoperative treatment in Canada for exactly this diagnosis, not the over 65 population, but just for degenerative disc disease. We spent about a year trying to put this together in conjunction with the Canadian Spine Society and we ran into a couple significant problems.

We had a very nicely outlined nonoperative regimen administered by the Hayes Back Institute, which did an aggressive combination of kinds of therapy, and we had surgeons who would agree to try to randomize patients, and then they got into the discussion about the ability to enroll patients, and their ethics boards at their hospitals which would not allow the incentivization to increase assets to the surgeon as part of that trial, and so that ultimately was I think the final straw that broke the camel's back, even in a state-run health care

00210

system, in trying to design a specific trial, which is all we're asking for, that we have not been able to figure out how to do today. So I think there are differences in health care systems that do alter patient behavior, and I think that this issue of a looking for a method to gain access is a real powerful force in patient care.
DR. KRIST: Last comment, and then we're going to focus in on our panel discussions here.
DR. WONG: David Wong. I'm the representative of the American Academy of Orthopedic Surgeons. Just to Dr. Faciszewski's point, I practiced in Canada's socialized system so I'm aware of waiting times, and I think that's a significant behavioral issue in terms of trying to attract people. If you can give them a faster track, and I think that's one thing about the European studies when there's a long wait time, is giving them a faster track.
The other thing that hasn't been discussed here is in the European socialized system, there's another track outside the trial, which is disability. And it was interesting, I was on a panel with Dr. Alan Masterson a number of

00211

years ago, and it was brought up at that point that the rates of people becoming disabled in Sweden from back pain had reached a point where they actually had to change the criteria because of the rate at which people were disabled, and it was taking up a huge percentage of the budget. So there is another track to become disabled there, as opposed to the United States, functional and still productive here in this culture, it's a different scenario.
DR. KRIST: Now to orient the panel, what I would like to do is focus on discussions that we need to have to clarify the evidence from our standpoint. When we get to voting, you'll see there's one thing in your packet and there's numbered cards in front of you. When we come to each of these questions, we're going to ask you to hold up your numbers and write it on your packet, and I guess Michelle will come by and pick these up, and then the votes will be tallied and posted on the web site.
But let me open it up for discussion here so that we can clarify these issues, these six questions before we turn to voting, or alternatively, we could go ahead and vote if

00212

people are ready.
DR. BURCHIEL: Can I ask a trivial

question. I'm still hung up a little bit on outcome measures and the idea that we focus at a point in time with, say, (inaudible) disabilities. I do think as a practicing neurosurgeon who does spine surgery including pain patients, but for a patient to enjoy some years of better life, pain relief, for example, has value. And so I'm asking a little bit about what is the data that we have that quality of life years has really been looked at for the intervention that we're talking about, and that is a measure that could be used. So it's a little harder, I think, but as a real world test, we kind of keep coming back to that issue, to me that's a real world test. The patient would say to me, I will take five years, that's a benefit to me, even if I know at the end of five years I'm going to fall apart, I would rather have those five years, rather than staying in pain for five years. So I'm asking of the panelists here their expertise on handling Question 4.

DR. KRIST: I would just like to make it, we could even discuss this more broadly, and

00213

the evidence that we've seen on the RCT studies and the cohort studies and the outcomes, what do people think of the outcomes that we're looking at?

DR. FLUM: I think that there's an opportunity to address the point you just made by using a quality of adjusted life years. That requires a certain degree of longitudinality.

DR. KRIST: I want to hold this for right now.

DR. FLUM: The quality of life years is a metric used in other studies and I think that's where you draw that kind of information from. They too are (inaudible) as Dr. Mirza was talking about. There are really two goals; one is to accurately reflect the impact of the procedure on the patient and the other is to try to inform patients about what to expect. And these operate on almost two different realms. Quality of adjusted life years speak to the health care czar who is trying to decide how much health care they want to give, and the intuitive value of the measuring quality of life is worthwhile and you have 100 percent longitudinal follow-up, and I think that clearly is an opportunity.

00214

I was going to add, now that we have Part D Medicare, Medicare has tremendous offerings of outcome assessments using health care utilization. It hasn't been a main feature of a lot of the work that's been out there, but

patients, the true way that patients use the health care system related to back pain in terms of braces and walkers and narcotic uses, those things can actually be measured, and they can provide a very meaningful objective measurement of the improvement after, maybe not any intervention, but certainly for this type of intervention.
DR. ONDRA: I think as a surgeon, I think Kim's question is what do we want to see as our entry criteria into a study and for a procedure, what are the proper outcomes, not only in terms of ODI, but what their disability improvement is, their overall health impact. You know, are these people getting older, not younger, and you know, so how does that factor in, how is improvement measured for nonsurgical or surgical means as well as overall improvement. You know, in addition to just a small slice of their pain and function, for instance, patients who are active, who have fewer other health problems

00215

relative to the population. So I think what we're really looking for is how do we design a study to answer these questions, there's going to be an entry selection criteria, and it's going to be on different treatments that we want to look at.
DR. LURIE: I wanted to raise an issue that is not so much what outcome measures per se but how the outcome measures are interpreted, and in particular this very attractive but perhaps only skin deep attractive idea of clinically important improvement, which is something we have a great desire to understand but I don't know that we have a great ability to understand it. And in particular, it is generally determined in the medical literature as a graded analysis of the individual, how much of a change in the measure is associated with a perception of the benefit for the individual. It is often misused to compare the difference in mean scores between two groups, which it is not designed to do, and we have seen people say we get an average in this group and a difference of X points, and that's not a clinically important difference. But that minimal clinically important difference might be defined by the difference between two groups,

00216

it's defined by the changes in the individual, and it has to be applied that way to make it understandable, and that's something that we need to be cognizant of when we talk about outcomes.
DR. KRIST: So you would be advocating more for a percent of individuals who have a minimally clinically significant improvement, and using that as a unit of comparison?

DR. LURIE: Yes. And the problem with that approach is that that reduces the statistical power. When you dichotomize the different variables, you lose power. But that is probably the metric we're talking about. Well, how do you measure changes in the patient so they know what the heck you're talking about, because they don't know what 14 points on an Oswestry score is. The percent of people who get at least a minimal clinically important change in that score probably is something that they can understand without having to come up with a new metric that is not validated.
DR. KRIST: I think in one of the studies, or one of the presenters showed us in one study, 68 percent of people, but it was just a pre-post looking at surgery and I think our

00217

comparative groups, or nonsurgical comparison groups, we don't have that number of results, as far as I'm aware.
DR. BOYAN: I have a question I want to ask, I think probably you. In some of these questions that are sort of floating around, none of these studies were powered to get the answer, but would a meta-analysis allow us to get at the answer? I don't understand the business of meta-analysis enough to know if you took all the studies together and combined them, is there some statistical way that we could sort of tease out what the comorbidities are?
DR. LURIE: The answer is sometimes, maybe. In the subject presented here, I think Dr. McCrory had it right and if you ask him if he did a meta-analysis, no, he didn't. Why? Because if you look at studies and you can tell that they are heterogenous, that is, what's happening in those four studies is not the same thing. Therefore, combining them is probably not the right thing to do. So that's the problem. Meta-analysis is very helpful when you have multiple small studies that are all studying the same thing about the same way, they just don't

00218

have the numbers, and typically a meta-analysis is useful when you have a study that shows a moderate or a large effect that's not statistically significant, because the problem there is that there is probably a good effect but you're not powered to see it. When you have a moderate sized study that shows no effect or tiny effect, the problem is not the power, the problem is there is no effect.
DR. KRIST: Or Fairbank, he had the ability to detect a 4 percent difference, and

that's a pretty significant power, more than minimal clinical significance.
DR. FLUM: If there were 14 studies of this type, you would have been able to do this, but four, you tend to smooth over differences more than you probably want, but there have been many worse groups of studies that have been analyzed successfully. These happen to be four studies that have really pretty clear interventions being the same. Then as we talked about, the inclusion criteria were similar if not perfect. But if these were meta-analyzed, you just wouldn't get much benefit from that. The reason I ask that is if you look at point estimates and say that the

00219

effect on these four studies when you take them all together is 0.7 with a confidence interval of .6 to .13, it is really just a way to administer precision to a science, but agree with the decision not to meta-analyze.
DR. ONDRA: I want to get back to the question I raised before, and I sort of gathered that no one is in agreement that we have the definitive answer on this issue with the current studies that we have. So given that and the difficulties with RCTs, are RCTs going to be the only way we can get at this? And if it is, how do you design that, or is there any other way, given the difficulty of doing an RCT in the United States? What are our ways to get there?
DR. FLUM: I totally disagree that the outcome (inaudible) surveyed the investigators in this field we looked at were reliable, it's been validated, it's a good solid metric and captured what we think we wanted to capture.
DR. ONDRA: That's not my question.
DR. FLUM: Right, but the outcome metric just drives whether or not -- the outcome metric often drives the randomization. I mean in fact, the practical barrier you all laid out was

00220

about recruitment. (Inaudible) unethical and totally impractical, and all of a sudden it became ethical and practical and added a lot of information.
DR. BOYAN: I might have an answer that isn't going to be friendly but, although it's how we should do it. I don't think there's enough information about anything to compare the two things we were asked to compare, and if the study has to compare usual care or nonsurgical care to surgical care, somewhere in this room we have to define some unified unit of nonsurgical care, which we haven't done. So I would suggest that the appropriate studies are to say let's accept

that we're looking at nonsurgical care, and take a variable in nonsurgical care and see if it matters, because we're trying -- I feel like I'm talking to my students. But you've got to define your question, you've got to have a single variable to have a rational study, and build up to it, get a protocol of nonsurgical care that could actually be given according to a protocol. I would suggest we take -- surgery is now an accepted thing, we're going to do surgery, but we're going to determine

00221

if we do surgery in old people, are they going to have estrogen treatment or not estrogen treatment. Make it simple so you can get a number you can use. What we've done today is talk about a lot of stuff that I'm not sure we can use.
DR. ONDRA: In the lung volume study, there was, I don't know enough about it to know if that's a fair comparison, but that was a procedure with some mortality associated with it, but again, I'm not sure that that is -- I think it's a well designed trial, I'm not sure it's a fair comparison, so we would have to look at it a little more closely to see if we're really talking about the same sort of thing. And number two, the number of people that it affects would be much, much smaller in terms of back pain.
DR. FLUM: It depends on how many people, because we really don't do all back pain.
DR. ONDRA: Well, for back pain it's very high. I was referring to degenerative disc disease.
DR. FLUM: To stay on topic, you can't look at a big group and the smaller group at the same time. I think we are not here to design a perfect trial, but I think that these are the

00222

issues that come up to address the adequacy of the outcome metrics.
DR. BOYAN: I don't think we are going to design the perfect trial, but I think we're arguing over minutia about an imperfect trial that we cannot fix with another imperfect trial, so I think we have to simplify our goals a little bit.
DR. ONDRA: I don't think we're designing anything, we're just giving CMS advice on what we would like to see.
DR. BOYAN: Exactly.
DR. PHURROUGH: Right.
DR. KIRKPATRICK: Let me just summarize what I understand this discussion was about. We are currently to weigh the spectrum of professional or educated opinion as far as the literature breakdown. We would like to be up to

randomized clinical trials for everything. And what I'm hearing, especially from my colleague here, is that we need to work towards the middle instead of working toward the other extreme, because the other extreme is not likely to be obtained with the multifactorial issues in the field. And I would agree with that pursuit of moderation, so to speak.

00223

And given that, I do think that we've heard some very valid comments about outcome measures, we have an ODI which is reliable at this point, but it is not refined enough to be very specific. We have visual analog pain scales we can use. We have quality adjusted life years we can use. I don't think we can focus on one. We have to look at a multifactorial approach of saying whether it helps the patients. And that's something that will help the surgeon say, because now I can go to the patient and say well, based upon questionnaires of how people do, many of them do well for three or four years after the surgery and do better than the ones who don't have surgery. Or I can tell them, we found that after so many years, they all do the same. It kind of depends on what the patient's questions come up as. And then if they want to balance that against the risk that they have to go through, then we get a risk table to help.

So, we need all these different factors as part of our analysis in our measures, and it may mean that there is a new one developed, and I don't want to complicate this with developing measures, but either a combination of measures or

00224

a progressive development of a new measure is important to be able to look at patient function, how they deal with life, and nobody has really talked about coping mechanisms, but that's huge and hasn't even been brought up. So you know, I don't know that we want to get into that ball of wax at this stage, but in the future it may have to be incorporated, so I think multiple different measures are important, not just one.

DR. KRIST: I'm going to ask the basic question, because if we're talking about how do you design, or what's the perfect study you want for this, in a sense it implies that there are problems that are current. So why don't we talk about, is there a problem with the current evidence in the four RCTs that we have and the number of cohorts? I have some issues with the cohort versus the RCT data, but let's talk explicitly about this.

DR. FENDRICK: I completely agree. We

have heard nothing about (inaudible) trial in the U.S. elderly population that are going to prove that it's different from these substantial variations that we've seen in four randomized trials. So we would ask, is there some

00225

physiologic mechanism, is there something that is -- I can understand that there might be a trend toward a difference, but the fact that there are four randomized trials in people without comorbidities, and I would imagine probably it would be easier to perform surgery to a higher level of specification if the differences weren't large in those people, so why would we think that they would be much larger in the people in the Medicare population, unless there is someone in this room that can say that nonsurgical therapy doesn't work well on them.

I think we know everything we're going to know for a while on surgery. I think what is quite clear that we don't know yet is the value or the impact of nonsurgical therapy in the U.S. elderly population. I think some of us who read these studies in Europe were quite surprised, not how well surgery worked, we heard from surgeons on how well surgery worked, right?

But surprisingly, it mentioned the nonsurgical therapy, and given that it was mentioned in two of the European trials, given there was no relevant intervention trial, the natural history of the disease could also be the

00226

same. People could have just gotten better at those same rates if nothing was done at all. And I think that's where a strong argument needs to be made to do another trial that's addressed to this.

DR. ONDRA: I got something completely different from listening this morning and that was while we saw that these trials were very flawed, I know one trial had a one-year follow-up, which is strikingly short considering recovering from surgery is going to take up a large part of that year, so I don't know that we have the ability to make all these answers relative to surgery. I find that very myopic.

DR. FENDRICK: Specifically, why do you think there's a difference between these markedly variables in what we saw in younger versus older, so whatever outcome measures we have, why would the marginal if not at all clinical meaningful difference that we saw in the pretrial, why do you think they'll be different in a different population, and that's a question that I don't know the answer to, but we need to do it in the U.S. because --

DR. ONDRA: Well, I think we need to do it correctly because I think these are all flawed,

00227

very flawed studies.
DR. FACISZEWSKI: Maybe I can help.
Two comments. One is that at the risk of going very far backwards, I think there is concern among the panel about, that there is a concern about whether this is a high incidence problem or low incidence problem. And from a spine surgeon's perspective, we're very much split and I think you hear that amongst the surgeons. And with all due respect, I think with the administrative database researchers, their numbers, everybody has degenerative disc disease. In surgery, very few have degenerative disc disease.
And maybe I can help define that term in my mind, and then address very briefly the tech report, which I think helps confound the confusion, not give us a solution. Degenerative disc disease is the painful syndrome, people hurt, and it can't be made on radiologic evaluation. The MRI scan does not tell me as a surgeon what hurts. And so I read the tech report and I see that in the radiology section they talk about degenerative disc disease. In my mind that doesn't compute, because the radiology evaluation doesn't tell me about disease. It has

00228

degenerative disc changes, they may have facet changes that are spondylitic, but they don't have the disease.
So I think we see some of this contamination of unclear thought, and that brings us further to the point of what's really going on with the Medicare population. I think that's what the question asks us. The ICD codes don't reflect it, the ICD codes that we as surgeons have to put down are confusing. We wrote some papers in the early '90s about presumptive coding. If you code certain things in the hospital and your code is okay, they go for it, they get paid for it. But if you have the complication of anemia, which now all of a sudden is a complication where the hemoglobin is below 30. So these things are terribly confounding.
So, to the point. I think they are very flawed studies because they don't define the patients that are enrolled. In fact, very few of them get MRI studies. In a multilevel degenerating patient, degenerated, not diseased, a different patient might had a single level degenerated disc, so we need randomized controlled studies that are all the same.

00229

And so if I could just address for one second and talk about what this means to me. If I were an internist and I was reading a study about cancer, and the study dealt with cancer patients, and cisplatin was given to all of them, and guess what, some of them got better. And the conclusion was that cisplatin actually was beneficial, but it really wasn't that much more beneficial than doing nothing. And we didn't control for bone cancer, breast cancer, prostate, they just had cancer. So my problem with these studies and in listening to the reports and reading the tech assessment was that there's no granularity, and the administrative databases don't help us with that. So I believe strongly that if nothing else happens from this panel, we have to define these terms and we can study in the future, and so when someone says degenerative disc disease in the future, I know exactly what that means, because I'm not convinced that we panel members all agree on that term itself, and that makes answering this question very easy.

DR. FLUM: (Inaudible) probably a good idea, and this is what I gleaned from this morning. One is, the study population is totally

00230

different, and we've all said we don't know how these patients would respond to any intervention. Two, we all have seen the same results, we're interpreting them differently. There's lots of different interpretations of what's positive, what's not a positive result, and mostly because the nonoperative events are being interpreted differently. So we have to design a really clean nonoperative event of the type that we discussed earlier, that would be a wonderful opportunity. Also to clarify with MRI, whatever the state of the clinical standard is right now for defining the disease process.

Those would all seem like great opportunities to pursue randomized studies, and use whatever metrics we use, but I don't think the outcome metrics is the problem. I think that we could convince the spine community and patients and the clinicians who are referring patients better if we had a state of the art study that had clean entry criteria and clean nonoperative intervention. I don't think we should say that just because right now there's no good nonoperative intervention that's paid for, I don't think that should be a limiting factor. I think

00231

we need to learn from these studies what a good nonoperative intervention may look like, and

create one. That's where I think the opportunity is.
DR. KRIST: I have a quick clarifying question for the panel. The SPORT study, is there an argument about that looking at this population we're talking about?
DR. LURIE: No, there is no back pain, degenerative disc disease or back pain fusion. There's a spondylolisthesis, there's spinal stenosis and leg pain predominant, that's the remaining arm. And again, I share some frustration in reading the tech assessment because you can't put isthmic spondylolisthesis, degenerative spondylolisthesis, axial back pain with dark discs, you can't put those things together and make any sense of it because they're different diseases, they present differently, they respond differently, the surgical outcomes are different between those diseases, the nonsurgical outcomes are different between those diseases, the long-term things you see are different for those diseases. They're all sort of degenerative, because that can mean whatever you want, but

00232

they're very different diseases, and to put them together, it's a mish-mash.
DR. KIRKPATRICK: It sounds to me like we need to stick to the specific question we need to answer. We need to design a randomized clinical trial that will probably take four to five years just to develop, and in the meantime we need to be doing prospective follow-up on everything we do.
DR. FLUM: And also get better information for what the indications of the operation are. We should have a spinal fusion for back pain, if that's the issue we want to be looking at. That's something that Medicare can influence and can add to their (inaudible) which goes out to these patients. And then as, you know, Medicare does this in many ways, where no patient gets covered unless they are on a prospective registry. This give us an opportunity to learn about these patients while we're figuring out the best way to do this study.
DR. ONDRA: And that's probably much more reasonable ground than just coverage or noncoverage, and, you know, the other thing is what outcome measures should be used for these

00233

patients. You know, are we just looking at a single one, are there a group we should be looking at, and those are things that I hope we will be looking at.
DR. KRIST: Are we ready to look at

voting?
DR. LURIE: No, I'm not. Besides that issue, there are at least two other things I need to be clear about what I'm voting about. So one is, we've sort of talked about, what's conservative care or nonoperative care in the studies. If we're voting on level of evidence compared to conservative care, what is it that we had in our minds that we're comparing it to? Is it the extensive tertiary function and rehabilitation like that provided in the Brox study, is that conservative care, or is it the physical therapy and whatever, like was provided in the Fritzell study, is that conservative care? Because the outcomes of those two things between those two studies were very different. So I would call the Fritzell study conservative care and I would call the Brox study intensive tertiary rehab. I think the outcome of those two things in the literature are different and I think the

00234

comparison between, you know, the surgery outcomes in all these trials is just about the same, but some of them showed a difference between a comparator arm because the control arm got better, and others showed a fairly big difference because the control arm didn't get better. So which of those studies am I supposed to have in my mind when you ask me to hold up a number?
DR. KIRKPATRICK: All of them.
DR. KRIST: The answer is all of what the evidence showed us, that's all we can comment on. And yes, there is a big variation in what nonoperative care is, but what we need to think about is in aggregate, the surgical decision versus nonoperative care, so we include all of them.
DR. LURIE: My next question is clearer. When it comes to the question of whether we're talking about how likely are these various procedures, fusion procedures, and we're supposed to talk about the without instrumentation and with instrumentation. So the question is for the with instrumentation patients, is it with instrumentation compared to conservative care or is it with instrumentation compared to without

00235

instrumentation? So instrumentation makes no difference, if that were somebody's world view, and if you think that fusion without instrumentation helps and that adding instrumentation does nothing, how would you answer the question about with instrumentation? Would you say it doesn't help if it's done without instrumentation, or do you say it does help if

it's exactly the same as without fusion, but it's better than nothing?
DR. KRIST: Well, I go back to, we have to think about the information that we have. Most of the studies that I saw in our tech assessment compared an operative intervention with and an operative without, and that would be one way to compare their relative importance. I mean, an ideal study to assess that would be a nonoperative control group, an operative control group with and an operative control group without. I don't know that we have that information to that level. But we have to think according to the same yardstick, so when you pick your numbers for each, it would be judging each against the same baseline yardstick.
DR. KIRKPATRICK: There is a subtlety

00236

with regard to the interbody fusion that I would like clarified. Are we saying that an anterior lumbar fusion done with a plate and BMP, is that with instrumentation or without? Because some people will put in a cage and also put an anterior plate on, which is with instrumentation, whereas I didn't recall seeing any reports of that. So I think some of these with and without are complicated concerns. My recommendation would be if there is data there, you say there is data there; if there is not data there, there's not data there.
DR. KRIST: And that can bring up a larger point about Question 4 in general, and we can talk about that now. I mean, a large amount of the data that we heard, even just looking at the difference with the four different procedures there, there was some amount of data saying that they were relatively comparable, right? I'm not saying that they are, but maybe we should just be looking at with versus without any instrumentation. So there's a number of ways to think about maybe clarifying Number 4, so what are the panel members' thoughts on that?
DR. KIRKPATRICK: I would suggest that

00237

there are several things that we've heard that show there's no difference among them, but some of those studies are larger than others, and so we might be able to say that from an evidence standpoint, we have certain levels of evidence to support it, but as far as the surgeons in the crowd, I want to make sure that we're not saying that we have a rationale to support one or the other as different based on what the evidence is.
DR. KRIST: Yeah. The question starts, based on the evidence presented, how likely.

DR. KIRKPATRICK: And I'm trying to emphasize that because that's been a hangup in prior panels that I have been on.
DR. BOYAN: I have a point of how things are done. Do we have any discussion time at the end of each vote?
DR. PHURROUGH: Yes.
DR. BOYAN: Good.
DR. JARVICK: I have a question about the patient population that we're dealing with and with respect to degenerative disc disease and imaging criteria. I heard a number of presenters say that if they had patients who have one or two level degeneration (inaudible) wouldn't operate on

00238

him. When we talk about patients who are over the age of 65, you're distinctly abnormal if you don't have any disc degeneration and it is present in virtually 100 percent of those folks, and the number of people who have multiple herniated discs, it goes up a little bit. So is this group, are we talking about only those folks that on MRI have only one or two levels of disc degeneration?
DR. FACISZEWSKI: I think your question is exactly the question that needs to be asked, and that is the question of these questions, and that is the extent on one or two discs, but we are left with two categories, there is the less than 65 and there's the greater than 65.
DR. KRIST: Well, the way the questions are worded too, the very last question is the Medicare population, right.
DR. FACISZEWSKI: But I think his question is, what do the other ones refer to.
DR. KRIST: I would read the other questions as in aggregate, overall, does this improve health outcomes for patients, not just specific to the Medicare population. And then the last question is okay, now, specifically for the Medicare population, how does what we said apply,

00239

or how does what we said earlier apply to just the Medicare population.
MR. QUEENAN: I was interpreting it actually to mean that until the last question, we're basically looking at the evidence based on the population studied by the evidence, and the last question is, is that population or those populations representative.
DR. KRIST: Yes.
MR. QUEENAN: I have a related question, though, which is on Question 3 in particular, since we had a discussion. A lot of people mentioned that there has to be certain patient selection relating to the discussion about

whether or not there was conservative treatment prior to surgery. I guess what we should assume is the case for Question 3, and I guess the way I was looking at that is that that would refer to what we glean from the evidence, but for the U.S. population, the population that would be, if not Medicare-eligible, at least in the U.S. health system, so we would interpret that in the context of treatment as it occurs in this country. That's just the way I interpreted it, but I'm really asking the question of how we should interpret it

00240

in that context.
DR. FLUM: I didn't get that. I interpreted it the other way. We only have the evidence in front of us to use. If nonsurgical treatment in Europe is what the evidence is based on, it is what it is, and that's the only evidence of conservative management or nonoperative management we have here today.
MR. QUEENAN: But then there is no question that allows us to answer the question as to whether that's applicable to the U.S. population.
DR. FLUM: Except the last thing that talks about it being applicable for the Medicare population.
DR. KRIST: I think the way I would interpret that, we're interested in thinking about the U.S. population, but I wouldn't by definition exclude studies that were not done in the United States. If you have a particularly strong reason to think that the information is not applicable to the U.S., if there is strong evidence and information from there that you don't think that this applies to patients that we're caring for --
MR. QUEENAN: There was a lot of

00241

discussion about that because the RCTs that were done were done outside the U.S.
DR. FLUM: And the nonoperative arm they doesn't necessarily exist in the United States for whatever commercial reasons. So I think those are valid issues. And you don't want to get boxed in by a question that, in answering, that's not really what the points are that are important.
DR. KRIST: Well, and I think we will make an opportunity to have comments to fill in those details, to be able to say yes, in this scenario I think if you have these alternatives, that it's not necessarily better. And that would be helpful, I think, for the group to come out with the background and thoughts behind why you feel the way you do on these various topics.

That's going to be probably the most important information for directing everyone in the room here as well as other people as they move to figure out where do we go next with this. I see people looking down and writing things on their sheets, so I'm going to take that as tacit approval to go ahead and start voting on these.

00242

DR. KIRKPATRICK: Just to clarify, are we going to take a vote and then have a discussion of each question, or do you want us to make sure we've discussed the discussion part of each question before we vote?
DR. KRIST: We'll take the vote first, but if you feel like there is a specific thing you want to talk about with the panel to clarify beforehand, let me know before the question.
DR. KIRKPATRICK: What I'm saying is the question and discussions are related, but one doesn't necessarily support the other.
DR. KRIST: Correct. We're going to do the discussion points after the votes. But if there's something that you want to talk about to clarify the question that you're voting on, just let me know before that question. Does that make sense?
So for Question 1, what level of confidence does the evidence provide in addressing the outcomes needed to determine the effectiveness of lumbar spinal fusion for low back pain due to degenerative disc disease? And if you would write on the sheets as well as hold up your numbers, Michelle will pick those up.

00243

(Panelists voted.)
DR. KRIST: And we talked some about the discussion point on this, which is, is relief of pain the appropriate primary outcome, or should it be restoration of function, return to work or something else? I have heard from the group a lot of talk about ODI. Are there other comments or things that people would like to bring out on that discussion topic?
DR. FLUM: I think we made the point about harnessing the power of Medicare in light of Part D Medicare as an opportunity of upgrading the way we understand the outcomes of surgery.
DR. BURCHIEL: But if we don't expand that to something that has more real world meaning, we're going to end up with the same results and not be able to interpret them either. I think there needs to be an additional criteria developed.
DR. FLUM: And just by way of

mentioning that, you know, in a bariatric surgery coverage decision, they were leaning towards, the Medicare coverage decision was leaning towards hospitals that were accredited by the American College of Surgeons or another, the American

00244

Society of Bariatric Surgery, and part of that accreditation process included yearly assessments of the patient-driven outcome, which is a way that could certainly be the ODI, and that would be a very nice eloquent mechanism to make sure that a yearly quality of life for the disease could be measured out. That's the type of thing that's already on line with the prospective registry, and would be a wonderful way to justify the procedure, and as a way to understand how this population plays out.
DR. FACISZEWSKI: So in bariatric surgery, they looked to a society?
DR. FLUM: The American College of Surgeons, and the American Society of Bariatric Surgery, they established a set of accreditation criteria based on structure, how they handle bariatric patients, processes, and required a follow-up at one year with a patient level outcome and then some outcomes, and that model has really marked to the power of the care as well.
DR. JARVICK: One of the things I wanted to point out is the outcome measures, we've touched on it a little bit, but you have the use of disability of function status rather than pain

00245

to be focusing on, and I think it's important to focus on those measures, and if you give more weight to that than just assessment of pain, I think you will have all sorts of problems associated with them.
DR. KIRKPATRICK: I think we pointed out that while we don't have a perfect measure, we need to improve the ODI or something very close to it that's well validated, visual analog data. For non-workers' compensation, I think it would be reasonable to look at return to work status, I don't think it's appropriate for the workers' comp population because of a number of issues that we've already brought up. And then I think somewhere either in the pain scale or somewhere else, we should include whether you're on narcotics anymore or non-narcotics, and I also agree with the idea of the quality of adjusted life years.
DR. KRIST: Okay. Move on to Question Number 2 then. What level of confidence does the evidence provide for characterizing complications, adverse events and other harms from lumbar spinal

fusion for degenerative disc disease? And we will start with Question A, short-term, and short-term

00246

being defined as up to two years after surgery.
(Panelists voted.)
DR. KRIST: And why don't we go ahead and do 2.B right now, long-term, and that's defined as more than two years after surgery.
(Panelists voted.)
DR. KRIST: Just to remind folks, this, all the averages and the vote will be on the Medicare web site tomorrow, or this afternoon, okay.
Discussion point for Number 2, what does the variability in surgical risk depend on? And let's lump these together, as this procedure is permanent, are there other potential long-term harms that have not been discussed.
DR. FLUM: My comment on variability, as we heard from Dr. Mirza about variability and utilization, undoubtedly there will be variability in outcomes, but we haven't heard too much about that. Outcomes were measured by what would be the question, I guess. Would they be due from the interventions or secondary to the spine, but that's probably the best way to look at it, at health care utilization as an outcome measure.
In the bariatric community, once again,

00247

this issue of variability was borne out, we were looking at mortality rates in that population, and the way our society approached that was to say that we don't know yet what the variables are affecting adverse outcomes, or in the operating room, but we probably know better than anybody else how to track down what's the quality of care for accreditation, some volume criteria, past record of performance, and once again, an accreditation model that looks at variability issues.
DR. ONDRA: On the long-term, it would be important to look at what is the rate of adjacent segment disease relative to the rate of fusion compared to the nonsurgical population.
DR. KRIST: We saw some outcomes on repeat surgery rates, and it would be good to have good comparison groups for that as well. It's a little difficult to put that into context because obviously the people who have surgery are at risk.
DR. FLUM: One nice way to do it would be through administrative data, and I'm sure Dr. Mirza would agree. It's not hard to get a modifier for an ICD diagnostic code, but there are ways to make it clear that this is reoperative

00248

surgery, and that would go a long way in indicating whether or not this was a quality issue, recurrent disease, or what.

DR. KRIST: Why don't we go a step further and say there is a randomized controlled trial against nonoperative techniques, and do a comparison with reoperative rates in those two groups. I think that would be very helpful. For the few studies we saw that had follow-up for up to ten years, I think that would be tremendously important.

DR. KIRKPATRICK: I'd like to add that with this question of variable risks, there is surgical risk and there is nonsurgical risk. Of great concern to a surgeon, of course, are medical comorbidities on the patient's side, and that needs further investigation, it has not been revisited in a number of years. Surgical technique, some of us talked about in terms of volume of case load and that sort of thing, or complexity of the procedure, and these can also be surgical control factors.

The majority of the complications do seem to be a risk of simply having a surgical procedure, anesthesia complications and that sort

00249

of thing.

As far as long-term issues, you know, we don't know much beyond two years, and I think that's something that may be ripe for study, but we don't know exactly what to look at.

DR. FLUM: I would like to extend on that one, I don't want to go back to accreditation for the third time, but one of the things it does is it allows the society to define the outcomes. Maybe ODI doesn't improve more than ten, you know, and I guess you guys know better than we do what the optimal outcome and risks would be, and then what patients are high risk for that and performing a risk adjust. It would allow people to compare apples to apples as opposed to apples to oranges. I think that should be part of the next generation out there.

DR. KRIST: Okay. I'm going to move on to Question Number 3. Based on the evidence presented, how likely is it that lumbar spinal fusion for lumbar degenerative disc disease improves clinical outcomes as compared to conservative treatment? And we'll start with A, short-term, once again defined as two years post surgery.

00250

(Panelists voted.)

DR. KRIST: And let's move to 3.B,

long-term, more than two years post fusion surgery.
(Panelists voted.)
DR. KRIST: Okay. The discussion point on this is one of the ones we have been dancing around a lot, and it looks like there's a series of four questions. Why don't we look at the first three and then we'll focus in on the last one? What are the causes of low back pain? Is patient important, and if so, what are the clinical and/or patient characteristics that are reliable predictors of satisfactory outcomes? And if there is an absence of evidence of long-term benefit, would evidence of short-term benefit be sufficient to justify a fusion procedure? Let's start with those and then we will readdress one last time if a clinical trial were to be done, what would it be.
DR. BOSWELL: A lot of my time is spent doing interventional pain management, and the key problem we are having right now is figuring out if any of our treatments are effective in treating the diagnosis. I'm going to say that I think the

00251

outcome studies aren't so bad, I think the problem may reside in the fact that we don't have a clear handle on our diagnoses in our patient population. So we're getting the best results we can with the outcome studies we have because we're looking at a mixed patient population. I think there has to be some emphasis on patient selection in some manner based on diagnosis.
DR. FLUM: And that point has been raised several times. In a randomized trial, if there is classification basis, and undoubtedly there is, in a randomized trial, we have to look at it as one of the arms of an appropriate powered randomized trial.
DR. BOSWELL: That's right, but what it means is that we have to determine that before we can tell what the difference is.
DR. FLUM: Right, but we can say for the hodgepodge of people that are called back pain, if you look at the criteria in the European studies, if you look at the people who meet those criteria, albeit with mixed diagnoses, it would highlight the results that we would see. And that's why I voted one for these, because I think if you have four randomized trials and three are

00252

telling you that there is not a significant comparative difference, that both arms got better, then the question is what's different about the last, the fourth study. And that's where I think this body of evidence falls down a little bit, on

the comparable efficacy. I think clearly both arms improved, but the question here was did one arm, did the surgical arm improve more than the nonoperative arm.
DR. KRIST: But there is a differential response to the different therapies for different diagnoses. One group of diagnoses might improve, or one group or the other might improve. I mean, one might improve for surgery and the other might improve for nonsurgical outcomes, so the average looks the same.
DR. FLUM: Well, that's accounted for by the fact that they're randomly allocated and you have enough of each group in each arm.
DR. KRIST: You have to know the diagnosis at the beginning, though.
DR. JARVICK: And that's precisely the problem. One of the things we're hearing is that we don't have a specific way of separating out the different diagnostic categories. MRI, we all

00253

agree is not great. Discography certainly has its problems. We simply don't have the diagnostic sophistication, and I'm somebody that reads these things for a living, and I'm the first to admit that we're not there yet. So patient selection, undoubtedly it's patient selection going in to do trial to get the appropriate intervention, but in the absence of the tools to appropriately select the patients, or the success with what we've got, and, you know, it becomes an effectiveness trial, which is what our experience has been. And the results are what they are, that there's no clearcut difference between the two groups.
DR. BURCHIEL: I want to make sure we're (inaudible) NIH panels, low back pain, I'm not sure they shed a lot of light on origins of back pain except to say it's complex and there's a menu of possibilities. I think that was mentioned a couple of times today. So if we're going to have a real etiologic basis for therapy, I think we're a long way from that, it may still be five years, or maybe not even in five years.
So on a measure of granularity, I think what we're talking about is yeah, we can define a few things. I think radiologically we can say

00254

there is degeneration or not degeneration, but the fact is that most of these patients come, at least in my experience, with several diagnoses, and that's basically the way they come. Stenosis, other degenerative disease, so we have to get away from this pragmatic classification, to not lump them all together, and ultimately we're going to separate degenerative facet joint disease from

degenerative spondylolisthesis, and we're going to have to put some of these definitions down that's reasonable. Otherwise, we'll never be able to power our studies.
DR. JARVICK: But the problem is they're getting these diagnoses based on radiologic criteria, and it may have nothing to do with their pain, or it may have something to do with their pain.
DR. BURCHIEL: Right. Whatever criteria we use, we're going to have to settle on a reasonably small number of discrete medical conditions that are very common to an observer, and not try to go beyond that, and we'll never get to the granularity of simply degenerative disc disease.
DR. ONDRA: This gets to entry

00255

criteria, and some of that is, you know, you have an analogy with cancer, you don't say you have cancer, you say you have a lung carcinoma, or maybe adenocarcinoma (inaudible, off microphone.) So you get somebody with loss of this kind of, you know, for DVT, no stenosis, no leg pain, no arthropyosis, however you want to clarify it, no spondylolisthesis, that's the population, but some entry criteria included that's really designed to not be granular to the point of undoable, but not a garbage bag.
DR. KIRKPATRICK: Did I miss it, did we go to the last part of that four-part question?
DR. KRIST: We are starting to slip into the trial.
DR. KIRKPATRICK: We were going to talk about the first three and then go full bore into the fourth, right?
DR. KRIST: Yes.
DR. KIRKPATRICK: Number one, to paraphrase the argument, we don't know, 90 percent of the time we don't know the cause of back pain. Number two is yes. Every speaker we heard said in properly selected patients, but we don't know what the selection criteria are, so that's another area

00256

for further study. I think avoidance of workmen's comp is a key element to evaluating multiple medical comorbidities, and there may be other issues that have further refinements of the diagnosis such as discography, which is still controversial. And then number three, I think we saw a nice slide that showed over a five or ten-year period, the two curves end up converging at the end. So the surgical arm came down in six months quite well with ODIs being reasonably well satisfied, and I believe they would ultimately

meet the quality adjusted life year benefit that Kim is looking for, so I think there is in short-term a reasonable expectation that the patients may do better.
DR. KRIST: Yes.
DR. LURIE: (Inaudible, off microphone.) You have to be clear about what it means for not having a good outcome or what predicts a difference in treatment value. So in the studies where we can look at it, the big lumbar spine study being the best one, the workers' compensation patients didn't do as well as the non-workers' compensation patients did. They didn't do as well in the surgical arm, they

00257

didn't do as well in the nonsurgical arm, and the difference between surgery and nonsurgery was the same. Actually, it was a little bit bigger in the workers' compensation population.
So when people say the workers' compensation patients don't do well, we don't want to look at them, we have to be clear about whether they don't do as well as everybody else or whether there is something about workers' compensation that affects what we see in terms of treatment, and the evidence that we have is that they never do as well as the non-workers' compensation population, but the treatment effect is probably about the same, if not a little bit bigger in that group, and we have to be careful about that.
DR. KRIST: I think that's the advantage of the design when you have a comparison group to look at differential change, as opposed to just pre-post.
DR. KIRKPATRICK: To clarify, I'm not saying we shouldn't operate on those patients or study them at all. I'm just saying we should exclude them if we're trying to figure out what's appropriate for the Medicare population, because when under workers' comp they're not getting

00258

Medicare. They may ultimately be, but in trying to sort these issues out, I think it's muddying the water more than anything.
DR. KRIST: Now taking up the last question, we've talked a lot about it, and I want to try and be concise on what we have to say, but if one clinical trial were to be done, what should it be?
DR. BURCHIEL: I know even though we have all said it would be an extremely arduous task, I personally can't see a way around a randomized controlled trial. I think we talked about validated measures, we talked about standardization of conservative therapy, we talked

about the consensus issues, we talked about the granularity issues. I think that one of the pitfalls of the field right now is it is still a maturing or dynamic field. We're going to have to define what it is at some point because if the target continues to move, we will never get an answer on this. So we're going to have to draw a line in the sand and that's what's going to be used and that's what it's going to be, and that will be the distribution of the trial, and not be left up to the surgeons.

00259

I think the issue of sponsorship is important. I know we're in a big trial right now where the industry is donating equipment but they have no role otherwise, and I think if we have an industry involved in the interpretation of the data and in any with the reporting of the data, it is going to be a worthless study.
DR. JARVICK: I would completely agree with that. I think just because an RCT is difficult to do and is expensive and will take a long time doesn't mean that it shouldn't be done. I think a problem which is as critical and has potentially as high an impact, not necessarily talking just the Medicare population, I think, but all patients with degenerative disc changes and back pain, that we don't have a definitive answer yet, and the best way to get the answer is an RCT.
DR. FLUM: I think one of the features of that trial should address bias, and we all know about bias, it comes in many forms. One form of bias is observer bias where the doctor is a cheerleader and physical therapy is used as an outcome. And blind observers would be a nice way and should be a key component of any kind of evaluation that's done.

00260

The second thing that we talked about, or the second bias is that patient expectation is a huge driver of outcomes. Let's go back to the European studies where both the operative and nonoperative groups get better, so we're talking about perhaps a small difference, a small comparative difference. Well, even if you have 15 to 20 percent placebo effect, and although the New York study on arthroscopy is controversial, this could be even more controversy because the spine surgery is genuine for the most part. There are ways to get around this, there are people working on this issue, but it will be the only way to disentangle the effect of the patient's expectation on outcome, which we know from that New York arthroscopy study is a huge driver of outcome.

Just by way of review, in the New York arthroscopy study, patients were randomized with knee arthritis and had to go through arthroscopy on that knee and had IV sedation and they made a few cuts on the knee, and they were watching their own arthroscopy being performed, whereas in the other arm they came to the operating room, got the IV, got a few cuts on their knee, and were

00261

watching somebody else's arthroscopy being performed, they simply got the three cuts on their knee. The outcomes in both groups were identical at every time point after the surgery. It's very telling about the role of the operation and the patient's expectations about the operation. I think that should at least be a component of the discussion as we move forward.

DR. KIRKPATRICK: I agree with the spirit of an RCT. I think the selection criteria and crossover issues, the nature of nonoperative treatment issues, as well as the clarification of outcome issues that have already been discussed make that very complicated. So I would ask you to rephrase the question, do you mean tomorrow or do you mean in five years? Because if you're talking about five years out, could we develop a reasonably good randomized clinical trial, I think in the meantime we could do some very good longitudinal follow-up on prospectively enrolled patients to be able to define some of the other issues that have been raised, like complication rates, comorbidity issues.

DR. ONDRA: I agree that I would do both, and some of these issues would be a tough

00262

sell for a first line surgery.

DR. FACISZEWSKI: The randomized controlled trial should be designed with the outcome of surgery (inaudible) United States. And it's industry-sponsored with limitations, but they were prospective consecutive series of cases, and actually they're a very good consecutive series, and it gives us a benchmark for use in the future. And I agree with the comments about some limitations in a randomized trial. As a surgeon, it's impossible to keep patients from being operated on for a year and not working, they'll go to another expert for the surgery. Long-term perhaps, but short-term I think we need to at least look at the prospective series. Lastly, I think we need to know what effects, if any (inaudible). When we looked at the nonoperative treatments, they weren't consistent either. The question is what power did they have or what effect, and I think we spoke

about this earlier, but I think this is the place to actually make a statement about that as well, because the fusion patients, as they were compared to the nonoperative group, they weren't compared to no treatment at all, and where I think we're

00263

giving credit for nonoperative care, it may not be any, I'm not sure what that effect is yet. I think we need to study it.
DR. JARVICK: And I think we talked about the problem with a case series with uncontrolled data, and while it's very good and very useful for complication risks, for looking at outcomes, comparing one group to another, it's totally, I wouldn't say totally useless, but it definitely has its limitations. Getting back to the issue of a sham that was mentioned, there may be some compromise ground that one could take. We may potentially be able to bring them into an angio suite, what we talked about earlier, having some sort of needle intervention, give them anesthesia so they don't really remember what happened, and drape them and prepare them, and then make the intervention sort of as sexy as possible, make them think they're having something done, and it may have a benefit on those patients, and we don't really know at this point. So I agree that placebo effect is potentially important and that is something we should try and get at somehow.
DR. KRIST: I think we have gotten a

00264

lot of good information for all the spine surgeons in this room, so why don't we move on to Question Number 4, and I'm going to try to go quickly on this.
DR. BURCHIEL: Before we start, I looked at this form and I'm a little puzzled because if I look at B, C and D, the without instrumentation doesn't make sense, so I would throw that back to the spine surgeons.
DR. KIRKPATRICK: Gutter fusion is appropriate to have both columns, because the gutter could have a noninstrumented posterolateral fusion in it. As far as posterior lumbar interbody and transforaminal interbody, some people will do a posterior lumbar interbody without instrumentation, some people -- I know few people would do a transforaminal without instrumentation, but some people do, so those are relevant, but they weren't really separated out well for us today. Anterior lumbar interbody, I agree with you. My interpretation is going to be that, you know, the with instrumentation is actually the 360, that's how I would view an

anterior lumbar interbody fusion with instrumentation, because I don't know that's there

00265

enough data or I've seen enough people put plates across an anterior lumbar interbody fusion, so that would be the with instrumentation category. So that's how I would look at it. I hope that's reasonable.
DR. KRIST: So you would have with instrumentation, and C would not have with instrumentation?
DR. KIRKPATRICK: In other words, C, I was going to leave out the with instrumentation, and then still have all the others to vote on.
DR. JARVICK: Alex, I just have a clarification. I know we discussed this already but I wasn't quite sure what your answer was. When we talk about improved health outcomes for lumbar degenerative disc disease, is that in comparison to nonoperative care or not?
DR. KRIST: Yes, I think it's overall. I know that's difficult, but we're looking at it overall, and so I'm not sure you can compare one to just the other.
DR. FLUM: I read this too as efficacy of data, in other words, all the observational and other series are going to apply, and is there evidence that these things improve health

00266

outcomes.
DR. KRIST: Yes.
MR. QUEENAN: In contrast to Question 3, this is not compared to conservative care.
DR. KRIST: Well, no, I think for both of them, I think some of the interpretation has to be the quality of the evidence, and just if you believe that the time period, that they improve over time, I think they're looking to see if it's the procedure that's resulting in their improvement, and not just do they improve. They could be subject to all the biases that we talked about. I think the purpose of Question 4 is does this procedure itself result in the improvement.
DR. JARVICK: So if our answer to Question Number 3 was we didn't think there was good evidence overall, then --
DR. KIRKPATRICK: Question 4 should be the same. Question 4 is stratifying Question 3.
DR. KRIST: Yes, for each specific procedure compared to -- so yes, if you put it in the context, if you didn't think that 3 was particularly helpful, then your vote should mirror that on some level in 4. Is that what you're

00267
looking for, Steve?
DR. PHURROUGH: Yes, although you could find that there was a lot of evidence for one.
DR. KRIST: So you might say one of these is particularly good, so maybe one would get a five, and the other three wouldn't be overall. Is everyone clear on this now, before we move forward?
DR. LURIE: Because I didn't understand what the answer was, is with instrumentation as compared to the same thing without instrumentation, or is with instrumentation and incremental benefit to without instrumentation, which one?
DR. KRIST: I think the premise, the concept --
DR. LURIE: If you have posterolateral fusion without instrumentation or posterolateral fusion with, you want us to somehow vote on these two things, right? The question is if we're comparing, I think I just hear we're comparing posterolateral fusion without instrumentation to conservative care. Then when you say with instrumentation, are we considering the incremental benefit of adding instrumentation to

00268
the posterolateral fusion, comparing with or without instrumentation, the way most of the studies do, or are we comparing with instrumentation to conservative care?
DR. KRIST: To complete things here, what we're asking you to vote on is not necessarily what we have evidence on, okay? So think about it from the standpoint, I'll just use concrete, and I'm going to make the scenario, if you think that gutter fusion without instrumentation is slightly effective, but with instrumentation is more effective, you might say three for one and four for the other. But it doesn't necessarily mimic what we have evidence on. It's, does this procedure improve outcomes, and so it would be does the procedure with instrumentation improve outcomes, and it would be does the procedure without instrumentation improve outcomes, not necessarily the relative, although the difference between the two will tell us that.
MR. QUEENAN: Both compared to conservative care.
DR. KIRKPATRICK: But I thought what you said was if I felt that the efficacy was better, not the evidence, and I thought what CMS

00269
was asking for was the evidence.
DR. KRIST: No, it is the evidence, and

I apologize for saying efficacy. It's the evidence, okay? All right. Let's go through with 4. Under short-term, and I'm going to do without instrumentation first and then with instrumentation. So we'll look at posterolateral gutter fusion without instrumentation, so you can vote on that one first.
MR. QUEENAN: Alex, I apologize. This is based on the evidence, what we think the evidence is, not a judgment on the evidence itself.
DR. KRIST: Correct, based on the evidence.
DR. KIRKPATRICK: Accepting that it may not be great.
DR. KRIST: Right. So for 4.A, posterolateral gutter fusion without instrumentation.
(Panelists voted.)
DR. KRIST: Now posterolateral gutter fusion with instrumentation.
(Panelists voted.)
DR. KRIST: Now for posterior lumbar

00270

interbody/transforaminal interbody without instrumentation.
(Panelists voted.)
DR. KRIST: And now with instrumentation.
(Panelists voted.)
DR. KRIST: Now anterior lumbar interbody without instrumentation.
(Panelists voted.)
DR. KRIST: Now anterior/posterior combined without instrumentation.
(Panelists voted.)
DR. KRIST: And now anterior/posterior combined with instrumentation.
(Panelists voted.)
DR. KRIST: Now we'll move to long-term, meaning more than two years post fusion surgery. So posterolateral gutter fusion without instrumentation.
(Panelists voted.)
DR. KRIST: Okay. Now with instrumentation.
(Panelists voted.)
DR. KRIST: And now for long-term posterior lumbar interbody/transforaminal

00271

interbody without instrumentation.
(Panelists voted.)
DR. KRIST: And with instrumentation.
(Panelists voted.)
DR. KRIST: Okay. And anterior lumbar

interbody without instrumentation.
(Panelists voted.)
DR. KRIST: And anterior/posterior combined without instrumentation.
(Panelists voted.)
DR. KRIST: And anterior/posterior combined with instrumentation.
(Panelists voted.)
DR. KRIST: Okay, good job. Does anyone have anything unique about the discussion point on 4? It's similar to 3 but specific to procedures. I'm assuming we can move on to 5.
DR. KIRKPATRICK: If I could suggest, a refinement of the indications for each should be explored. In other words, if there is a benefit to doing one of these techniques in certain (inaudible). Maybe one with degenerative disc disease and facet arthropathy needs a 360, whereas if it's just degenerative disc disease, a (inaudible) some sort of project looking at that

00272

sort of question would be the ideal thing.
DR. KRIST: Okay. Question Number 5, what level of confidence does the evidence provide that radiographic interpretations are correlated with clinical outcomes of lumbar spinal fusion due to lumbar degenerative disc disease?
(Panelists voted.)
DR. KRIST: And then the discussion question, is there uniform agreement regarding terminology for radiographic interpretations? And I mostly saw ones and twos, so I doubt that there is much of a discussion with that.
DR. JARVICK: Actually, I think the fact that there were ones and twos is because the studies don't easily correlate the outcome after spinal fusion isn't the same as is there a standardized nomenclature. There in fact is a reasonably standard nomenclature for describing degenerative disc changes that all the major societies have signed on to, this was published four or five years ago. I mean, there is a standardized nomenclature, but how well it predicts or correlates is a whole other issue.
DR. FACISZEWSKI: As one of the co-authors of that paper, not many people use

00273

those terms.
DR. KIRKPATRICK: So my suggestion on is there agreement, no, because many clinicians don't agree.
DR. BOYAN: I'm saying it's so bad that every time there's an FDA panel, they have to bring in an imaging expert to explain it to everybody in the room.

DR. JARVICK: Fair enough. You know, people should be using these terms, put it that way.
DR. KRIST: It will probably support their use if it's linked to clinical outcome, then there would be a motivation.
DR. FACISZEWSKI: There is a great deal of misconception, and I think that's why, part of the reason we're here is because people are talking about degenerative disease and low back pain, and pretty soon everyone thinks it's a huge problem, and I'm not sure that's the case.
DR. JARVICK: Well, until people do start using the standardized nomenclature, it makes the research very hard to do and the patient selection classification hard to do. So again, I think the nomenclature is there and should be

00274

used.
DR. KRIST: It probably goes beyond just the radiographic nomenclature to the diagnostic nomenclature as well.
All right, Question 6. Based on the evidence presented, how likely is it that the results generalize to the Medicare population, and then for A, relief of pain?
(Panelists voted.)
DR. KRIST: Okay. And then B, for complications, adverse events and harm?
(Panelists voted.)
DR. KRIST: And the discussion point here is, do studies need to be done in the Medicare population to strengthen the conclusions, and what is the impact of age and comorbidities?
DR. BOYAN: I have been waiting calmly, because obviously I think you have to do studies in the Medicare population, but I think there are things that need to be said to people here. And that is even though old people do heal, they don't heal the same as younger people do. They heal more slowly because they have issues related to age that are not, I wouldn't call them comorbidities, it's just the fact that they're

00275

older. They have fewer defensible stem cells, they heal more slowly, they may have other defects that are not the same as young adults. So you can't just assume that if something is working one way in a younger population, that it's going to be working as well in an older population. So we have no question that it has to be effective for the Medicare population. And then when we get to the comorbidities, obviously the incidence of disease is greater, and these are also true for chronic diseases, autoimmune issues, and those

things are all going to impact the outcome, so it has to happen.
DR. KIRKPATRICK: With all due respect to Barbara, I agree with her on a basic science level. There are clearly differences among ages. However, my concern is that the expense and logistical complications of trying to do such a study in an over 65 group may not be enough benefit to warrant those hassles. And so, you know, that's my major concern.
I do think that we need to bring up the issue of physiologic age as opposed to arbitrary chronologic age, because it does appear that many of our population is maintaining their health

00276

longer now, as we see by the rising mortality ages, things like that. They're more active and this all may translate into changes in the way that you in fact heal. So I think that's another concern to bring up. And I think another key thing with comorbidities, we just don't know how much that affects with the results and that would be a better concentrated effort on comorbidities than on the age factor.
DR. FACISZEWSKI: If I could just add that my understanding is that not all Medicare patients are over age 65, some are actually in the disabled group, some of which may have disabling back pain and be under age 65, and therefore may reflect more equality with the cohorts that were presented in the research. So my vote was related to the over 65, and I'm assuming the question was related largely to over 65 in a percentage basis, but certainly not to diminish the Medicare patients who are under age 65, and I would be very happy to learn about that component of the Medicare population.
DR. FLUM: I would like to add to that, because specifically the point that's been raised about the lack of nonoperative interventions in

00277

the United States, I think it would be very hard to imagine the randomized trials from Europe would have similar results in the United States when comparing usual care and nonoperative care. But in the absence of a designed and well reimbursed nonoperative intervention, how likely it is generalized to this population here, I think that's problematic.
DR. KRIST: Okay. Now one of the things we'll do is go down the table and have folks make comments and sum things up, but we've done a lot of talking as we've gone through these questions, so maybe I will end with that, and if anyone feels like they have anything to say on

this topic that they haven't had a chance to say already.
DR. JARVICK: I would like to make one comment, that I think there is a real opportunity here for CMS to play an active role in helping to gather the evidence that seems to be lacking. And I think we talked about the hurdles of doing randomized trials in this country, a lot of which center around incentive for patients to enroll and paying for the procedure, and production of trials, and CMS could play a vital role in that.

00278
DR. FLUM: And just to build on that point, if we are going to change the way we reimburse for surgery, I think it's a great opportunity.
DR. BURCHIEL: I think there is one thing we haven't talked about, or indirectly, that the bulk of the patients who get the surgery are not in the Medicare age range or beneficiaries, but insurance companies look to CMS for leadership, I think that's why everyone is here, because what happens here has import across the board in the marketplace. So I for one don't understand why CMS or NIH has the awesome responsibility for a study that's going to be very difficult and expensive, and largely relevant to patients outside the Medicare population. I think it's a broader issue because we talked about consortium that might include industry. This is a massive issue not just for CMS.
DR. KRIST: We appreciate all the expertise in this room and thank you for taking the time today to come here.
DR. MANCHIKANTI: I think in the elderly population, we are missing a point. If you look at the diagnosis, many of them have facet

00279
joint pain, but probably one of the things we can do is eliminate the facet joint pain before going to fusion, and that would be the proper case.
DR. KRIST: Thank you.
DR. PHURROUGH: All right. Thank you all, particularly the panel for your time and effort. This is a lot of work, a lot of stuff to read and do. Many of you will want to know what our next steps are. I lied, and we are going to do an NCD.
(Laughter.)
No. We will produce a fairly substantial set of minutes from this discussion. We are interested in proceeding with sort of outlining the data selection that needs to occur both in terms of a long-term discussion around what a good trial should look like versus some

ongoing data collection observational type of data, what can we as an agency do to assist with that, are there coding or claims issues we can work on to assist with that.
And we have taken, as you've heard today, taken the opportunity over the last couple of years to use different tools and techniques to stimulate data collection, and we would like to

00280

have continuing discussions around how we can best utilize those tools in this particular arena. I don't think we will mimic LDRS where we were concerned with 18 percent of mortality where we stopped covering the surgery and required it only in a trial. I think we would have a difficult time to say we're no longer going to pay for fusions, that would be a challenge. But we are interested in continuing interaction that will not stop here today, but will assist the community, you, the providers and patients in understanding what are the best treatment for low back pain.
So thank you, panel, again, and the audience for assisting us today. The meeting is adjourned.
(Whereupon, the meeting adjourned at 2:54 p.m.)