Transcript of April 16, 2002 Meeting

Please Note: This transcript has not been edited and CMS makes no representation regarding its accuracy.

00001

CENTERS FOR MEDICARE AND MEDICAID SERVICES
Medicare Coverage Advisory Committee
Executive Committee Meeting

April 16, 2002

Baltimore Convention Center
100 West Pratt Street
Baltimore, Maryland

.00002

Panelists

Chairperson
Harold C. Sox, M.D.

Vice-Chairperson
Robert Brook, M.D.

Voting Members
Leslie P. Francis, J.D., Ph.D.

John H. Ferguson, M.D.
Alan M. Garber, M.D., Ph.D.
Frank J. Papatheofanis, M.D., Ph.D.
Thomas Holohan, M.D.
Daisy Alford-Smith, Ph.D.
Wade Aubry, M.D.
Barbara McNeil, M.D., Ph.D.

HCFA Liaison
Sean R. Tunis, M.D., M.Sc.

.00003
Panelists (Continued)

Consumer Representative
Linda A. Bergthold, Ph.D.

Industry Representative
Randel E. Richner, M.P.H.

Executive Secretary
Janet Anderson

.00004
TABLE OF CONTENTS

.00005

TABLE OF CONTENTS (Continued)

.00006

TABLE OF CONTENTS (Continued)

.00007

PANEL PROCEEDINGS

(The meeting was called to order at 8:38 a.m., Tuesday, April 16, 2002.

MS. ANDERSON: Good morning and welcome, committee chairperson, members and guests. I am Janet Anderson, executive secretary of the executive committee of the Medicare Coverage Advisory Committee. The committee is here today to discuss and act upon recommendations from the MCAC Diagnostic

Imaging Panel regarding the use of positron emission tomography (PET) scanning technology for the diagnosis and patient management of Alzheimer's Disease (AD) and other dementias. The committee will also continue the discussion from the last executive committee meeting regarding the use of decision modeling in policy making.

In evaluating the recommendations presented to you today, CMS encourages the committee to consider all relevant forms of information, including but not limited to professional society statements, clinical guidelines and other testimony you may hear during the course of this committee meeting.

The following announcement addresses conflict of interest issues associated with this

.00008

meeting and is made part of the record to preclude even the appearance of impropriety. The conflict of interest statutes prohibit special government employees from participating in matters that could affect their or their employer's financial interests. To determine if any conflict existed the Agency reviewed all financial interests reported by the committee participants. The Agency has determined that all members may participate in the matters before the committee today.

With respect to other participants, we ask in the interest of fairness that all persons making statements or presentations to this committee disclose any current or previous financial involvement with any firm whose products or services they may wish to comment on. This includes direct financial investments, consulting fees, and significant institutional support.

I would now like to turn the meeting over to Dr. Sean Tunis, who will give his opening remarks. Then Chairman Dr. Hal Sox will ask the committee members to introduce themselves and to disclose for the record any involvement with the topics to be presented today. Dr. Tunis.

DR. TUNIS: Thanks, Janet. I just wanted

.00009

to welcome everybody to another executive committee meeting and thank you for taking the time and trouble to be here today. We have a modestly full agenda so hopefully there will be some time at the end for the committee members to tell various stories. I'm hoping Dr. Garber can tell us about his experience in running the Boston Marathon yesterday. He's alive, apparently he finished, and apparently carried his cell phone along the way just to stay in touch. And I think Randel Richner has a story to tell as well about some adventure yesterday in Baltimore.
In any case, I do want to turn the meeting over to Hal Sox to describe the charge for the first part of the committee meeting today.
DR. SOX: Good morning. I think we will start by asking the members of the executive committee to introduce themselves and when they do that, when you do that, would you please state any conflict of interest that you have relative to this particular topic. And we'll start with Linda Bergthold.
MS. BERGTHOLD: Linda Bergthold, consumer representative, no conflict.
DR. HOLOHAN: Dr. Tom Holohan, chief of patient care services, Veterans Health
.00010
Administration. No conflicts.
DR. GARBER: Alan Garber, Department of Veterans Affairs and Stanford University. No conflicts.
DR. ALFORD-SMITH: Daisy Alford-Smith, director of the Summit County Department of Human Services in Ohio. No conflict.
DR. McNEIL: Barbara McNeil, Harvard Medical School and the Brigham and Women's Hospital. No conflict.
DR. FERGUSON: John Ferguson, neurologist and consultant in private practice. And no conflict.
DR. BROOK: Robert Brook, Rand UCLA. I don't think I have a conflict, but I do come from UCLA and a large chunk of the material is from scientists at UCLA as opposed to Harvard, so Barbara has no conflict.

DR. AUBRY: Wade Aubry, internist and endocrinologist, the Institute for Health Policy Studies, University of California, San Francisco. No conflicts.
DR. PAPATHEOFANIS: Frank Papatheofanis, department of radiology, UCSD. I am a practicing nuclear medicine physician who performs PET scans. I guess a marginal or potentially apparent conflict
.00011
would be that our department negotiates with some of the manufacturers who may be involved in the deliberations today, on behalf of the university for the purchase of equipment. Other than that, I have no conflict.
MS. RICHNER: Randel Richner, Boston Scientific. No conflict.
DR. SOX: Well, our charge today is to either approve the recommendation of the imaging panel or to send it back to the panel for reconsideration. So really, those are the two options. We're here not really to hear new evidence, but to decide whether the process used by the panel meets the expectations that we have set up in our MCAC processes, and insure that there wasn't some serious miscarriage of logic in interpreting the data and reaching the conclusion.
So, that's really how we judge the panel, and make a decision whether to approve or not. It's mostly a check on the process of the panel and the notion is that the panel are the people closest to the data and that, the ones that were in the best position to decide what kind of evaluation to provide the coverage group. Our job really, we're experts on process, and we just want to be sure that in coming
.00012
to their conclusion, they used appropriate processes both with respect to the consideration and to the interpretation of the evidence. Are there any questions about that? Yes, Wade.
DR. AUBRY: I'm wondering if that's a change from the previous process or procedure. It seems to me in past meetings, there have been substitute motions regarding the previous panels'

work. Are you saying that basically we either accept the recommendation or send it back, is that true, and is that a change from previous panels?
DR. SOX: It's true in the past, I can remember at least one instance where it was my panel that we were evaluating, where we tweaked a word or two, but we're anticipating the time when we will not approve the panel, and simply evaluate the process, and so I would say that we'll have to consider the circumstances as they arise, but in general we are in the process of moving away from really any second-guessing the panel at all other than to evaluate the performance in evaluating and interpreting the evidence. So, that's a bit of a wishy-washy response but we'll just kind of see what happens.
DR. AUBRY: Thank you.

.00013

DR. SOX: Any other questions? In that case, Frank is going to make the presentation of the process, the data and the interpretation. So, I will turn it over to you, Frank.
DR. PAPATHEOFANIS: Thank you, Hal.
Hopefully the committee has copies of the summary of the meeting minutes as well as the report by the Duke evidence based practice center, which is this document.
I wanted to review very briefly the events of January 10th by starting off with a discussion of process, and then I have a few slides that deal with the actual question at hand and the recommendation that emerged from the deliberations.
Before I describe anything else, I wanted to mention that CMS through the guidance of the panel, the Diagnostic Imaging Panel members, was able to obtain the expert advice of three consultants from Harvard, who had expertise not only clinically in the treatment of patients with AD, but also had expertise in the methodologies that were used to arrive at the decisions the panel made. And I was very grateful for their participation and I think that everyone who was present was very appreciative of their thoughtful review of the Duke report, and I think that their

.00014
contributions to that session were very meaningful. As you know, AHRQ commissioned the Duke center to prepare a technology assessment. The assessment was performed by a group headed by Dr. David Matchar, who is here in the audience and who I hope we can tap if we have any concerns regarding the details of the model beyond our deliberation. But as you can see, there was a whole host, on the cover page of this document, there was a whole host of individuals who participated in the formulation and in the completion of this study. Likewise, at the rear of this document, there is another list of consultants and peer reviewers that went through the document on behalf of the Duke group, and I think that from my perspective and I think Barbara will speak later, it seemed that they did a very thorough job in assessing the questions that were posed by CMS.
So when we met, we started off by using this technology assessment as basically a reference document from which we considered Dr. Matchar's group's analysis and considered the conclusions that were formulated by the analysis. Some of the points that came up during the discussion were specifically related to the notion of the American Academy of
.00015
Neurology guidelines for the evaluation or workup of patients with suspected AD. And again, referring to this reference document, you can see where the Duke group broke down the various options that existed for the management of a patient who presented with either early signs and symptoms of the disease, or with first degree relatives or other dispositions that might suggest the likelihood of them developing Alzheimer's disease at some point in the future. The deliberations that we had focused on whether or not that was a reasonable tack to take. We relied on experts, physicians who actually treated patients with disease or suspected disease, as well as the review of the literature, the evidence tables, and the conclusions that came out of the report. So I think that the committee had at its disposal quite

a bit of information upon which to make the decision concerning the questions at hand.
Finally and not at all unimportantly was the testimony that we heard from members of the audience. There were several presentations that were made by speakers who signed up in advance and had been granted specific time slots to present their evidence on behalf of the use of PET for patients with AD, and there were also individuals who
.00016
volunteered and were given permission to address the panel at the time. So I think we had a very comprehensive mix of information sources in the expert testimony of the consultants that were used by our panel, again, both clinicians and methodologists. We had the evidence that the Duke EPC provided in the form of a written report, and finally, we had the input of the members and experts at hand.
And so, I think that pretty much wraps up the process of how we arrived at our discussion and how we arrived at the conclusions that we made.
I want to flip over to my very brief slide presentation. I don't know how that's going to work. I can see the yellow light already going off.
(Laughter.)
DR. PAPATHEOFANIS: What I wanted to do was take perhaps five minutes to review in summary the deliberations of our panel and again, review the specific question at hand that we considered. The question that was posed to the contractor, in this case the Duke group, is as stated there: Is the evidence adequate to demonstrate that PET has clinical benefit in evaluating patients with AD?
.00017
And this model, as is noted on the slide, was developed by Dr. Matchar and his group and then validated by Dr. Peter Neumann, at the Harvard School of Public Health, and Dr. Neumann was one of the outside experts that we had present at our meeting January 10th. He participated very actively in the discussion and I think did a fine job of addressing

our questions regarding the methodologies that were used.
To give you a very brief background, once again, we took as our starting point the existing American Academy of Neurology guidelines that recommend treatment with cholinesterase inhibitors for patients with suspected AD. Again, these are patients who have a family history or who have signs and symptoms on physical examination. The meta-analysis that was performed by the Duke group showed the true positive and true negative rate for PET that you see there, approximately 86 and 87 percent.
We also had very specific wording, and this guideline appears in the technology assessment provided by Duke, that the AAN guidelines do not recommend the use of PET and finally, there is a reference there that is available.

.00018

By the way, the minutes of our deliberations are also available on the web, so whoever wants to delve into the details, they are posted.
To continue with the background, treatment of patients with AD delays progression by six months, which we felt and we thought was a significant interval, so there is justification from a clinical perspective for the use of the cholinesterase inhibitors. Current drugs are generally believed to have a low risk of side effects, and we spent quite a bit of time discussing these side effects, and the committee was convinced that either the side effects are transitory or they are of such low grade that they don't preclude their use in the majority of patients, I think in over 90 percent of patients.
Finally, there are no good data on the impact of PET on the management of outcomes of patients with this disease.
To review in a little greater depth our approach, a decision analytic model was developed by the Duke EPC. It considered three different aspects of the disease as I alluded to earlier, that is, patients with mild to moderate dementia from presumed

AD, who based on a AAN recommendation, or I'm sorry,
.00019
an AAN guideline workup, were thought to be candidates for treatment; patients with mild cognitive impairment or MCI; and then finally, the patients with no symptoms at the time of evaluation but with first degree relatives with a history of AD. The sensitive analyses that we used, and you can see multiple sensitive analyses throughout the document, tested PET and treatment characteristics. The specific end points were the correct diagnosis, and the diagnosis or misdiagnosis that led to the identification of false negatives. The sensitivity analyses were done very methodically and basically there were one-way and two-way sensitivity analyses, wherein the one-way, one variable of interest was kept, was modified whereas all the others were kept constant, but I think the next level of sophistication was in the two-way sensitivity analyses where we see both complication rates and diagnostic efficacy. We found that to be especially useful.
So, let me go over a couple of these two-way sensitivity analyses very briefly. The first one covers effectiveness in treatment from 0 to 100 percent, and also looks at the complication rate. Now the complication utility is in the Y axis, and
.00020
basically the complication rate increases as the number decreases. The treatment efficacy also is displayed the same where the treatment efficacy increases as the number decreases. And what you have here is a baseline or base case, which is very well described in the technology assessment, of a patient who is 76 years old and undergoes diagnosis and presumed treatment according to the AAN guidelines and according to the existing patterns of care that have been expressed to us by clinicians and neurologists for treating these patients.
And so what you see here is the result of this two-way sensitivity analysis, and is that the base case falls squarely in the zone corresponding to treat all patients without further testing.

The other two-way analysis that I want to very briefly touch on, and obviously this is all explained in great detail in the reports you have reviewed, deals with effectiveness and complications over the current ranges that we're interested in. And again, what you see is the base case or the base line falls well within the treatment of all patients who present with mild dementia. And the complication utility is over 0.99 for the duration of treatment, which means that it's extremely, extremely low.

.00021

Basically, the position of the base case in this scenario also confirms that the model, I think, fairly represented what was evident in the published literature and what was derived from the evidence tables. And so our group, the Diagnostic Imaging Panel concluded that PET is not or would not be useful in any of the three diagnostic groups, if you will, that were identified and were carved out in the assessment. The treatments actually become more effective and complications from them are worse in very very few patients.

With PET and current treatment options, avoiding unnecessary treatment is associated with denying treatment to patients with AD or who will be developing AD who will benefit from it.

Currently, false negatives are worse than false positives, and the basis for that statement is the notion that there is in the cholinesterase inhibitors a relatively effective treatment for delaying the onset of signs and symptoms of disease, and for ameliorating current symptoms. And missing those patients and denying them the opportunity to be treated is tragic from a social ethical human perspective, and based on the 89 percent sensitivity of PET in the situation, the committee felt that

.00022

avoiding therapy unnecessarily in these false negatives was not acceptable.

The question then for the contractor, is the evidence adequate to demonstrate, and you know that, and you know what our answer is, it was a unanimous no. And there was quite a bit of

deliberation by the committee on all points that are associated with this analysis. I can honestly say that we took this fairly carefully. It was the first time that a decision analytic model was being used to really drive the process, and I think we were cautious and really tried to cross our Ts and dot our Is, and I can't overemphasize the importance of the outside contributors.

So rather than babble on, if that's adequate, Hal, I think I will stop here and give Barbara a chance.

DR. SOX: Great. Barbara, would you like the chance?

DR. McNEIL: I would like to just say a word. Frank, I think, summarized it well. A couple of points.

The first is that the committee took very carefully two criteria the executive committee had developed initially in its lifetime, that emphasized

.00023

that the evidence had to be adequate to determine whether the test provides more accurate diagnostic information and affects health outcomes. And because there had been no data on the impact of tests on health outcomes, the only approach dealing with that was through a model based activity that David and his group did. So that was our underlying modus operandi.

And the other point worth mentioning is that the sensitivity analyses that Frank just mentioned, we spent a lot of time discussing the utility scale there which goes from 0 to 100, and this was particularly important. We benefitted a lot from Peter Neumann's expertise in this area. It turns out one of his graduate students had just completed a magnum opus on the summary of the utilities that had been assessed from patients and from their caregivers for a variety of disease states. So he brought to the panel that summary literature which dated back from the late '70s, and it was clear that on the basis of that literature and more current literature as well, that the utilities even for the complications that we're talking about,

were very much in the upper 90s, that they in no way could fall much lower than that.

.00024

So our conclusion is as Frank said. We also considered carefully the Table 10 document which looked at some nonquantifiable issues that I know one of the outside people addressed.

The final comment I would make is a very important one, I think, and it's a little counter-intuitive, and that is that those analyses showed that for testing to be warranted or to be recommended, treatment would have to be considerably better, that is, moved to the left, and complications would have to be really worse. And it was only when that combination, lots better treatment and lots worse complications, that we would be in a situation to flip from the recommendation to one that was test and then on the basis of information from testing, treat.

So my summary statement would be that I thought, I have been part of only a couple of panel meetings that we have had and this executive committee as well, but I thought the process was really quite good, Hal.

DR. SOX: Perhaps somebody could comment on test performance. You didn't tell us about the studies of test performance, what gold standard was used, the degree to which the results of the PET scan

.00025

might have determined whether the patient got the gold standard test.

DR. McNEIL: There was a number of studies, Hal, and some of them were flawed with verification bias as you're implying, but the AAN guideline talked specifically about several that used histopathology as the diagnosis. And because of the possibility of verification bias, which David addressed in his report, the importance of sensitivity analyses was really quite critical. But I think that the meta-analysis showed the data that Frank summarized and the bottom line is, it would be very hard to get to a sensitivity higher than the high 80s or low 90s even if you

pulled out selectively, and there are some ROC curves that show a few individual cases without sacrificing a lot on the specificity side.

DR. SOX: If verification bias were present, it should tend to lead you to overestimate sensitivity and make the situation look better for testing.

DR. McNEIL: That's right, so that 87 percent that was on the slide, at least from the literature today, is probably pretty good.

DR. SOX: So, is there any value of

.00026

sensitivity where routine testing would have been appropriate?

DR. PAPATHEOFANIS: Not from what the literature showed here. There is also, Hal, to Barbara's point, in figures 5.B through 5.D, a progressive increase in the performance of the medication. It's expressed as a treatment efficacy variation in this case, where what you see is basically as we progress to greater treatment efficacy resulting in higher expected values for these patients. The highest gain is on the order of two to three weeks, which I think very dramatically brings the point of moving the curves around a little bit.

DR. SOX: One last question. Did you assume a single value for the pretest probability of AD given the presentation with mild to moderate cognitive impairment and are there some pretest probabilities, relatively low pretest probabilities where testing might make a difference? How did the prior probability --

DR. McNEIL: Maybe David could answer that, Hal, but I think it's best summarized actually, without going, well, I guess going to other literature, in the AAN guideline which actually talks

.00027

explicitly on page 115, about the impact of pretest probabilities in various situations, and how a variation in the pretest probability would not have changed the conclusions, but maybe David wants to say something.

DR. SOX: Did you want to say something, David?
DR. PAPATHEOFANIS: There is also the base case; were you going to get into that?
DR. MATCHAR: I was just going to comment that the base case, we actually used a relative low prior probability of about 50 to 55 percent, reflecting the fact that even after patients are evaluated clinically, they may still not have AD ultimately, so over that range and even a wider range, there was really no, there was in fact no impact.
DR. SOX: Bob?
DR. BROOK: Did you look at a base case scenario which basically says that the system right now is incompetent in diagnosing anything that looks like a dementia, let alone AD and if a person just wanted to know, you know, a person could recognize, hey, I'm doing something different, I want a PET scan because I can't get anyone even to do a mini-mental,
.00028
find me, follow me, do anything because the average level of care for this constellation of symptoms is just terrible. Did anyone model out the worst case scenario where sort of what is going on in primary care in the United States at the moment?
I mean, you were assuming sensitivity over something, and I couldn't get from the material that you sent me what is the assumption about the average level of quality of diagnosing this condition at the moment.
DR. MATCHAR: Well indirectly, you can take that from the analysis, because the basic idea is that there is no evidence that PET scanning differentiates a more responsive patient to treatment and that the treatment is relatively benign so at a relatively low or very low prior probabilities, treatment would be a reasonable strategy. We are not dealing with cost issues and some other issues that were outside the purview of this analysis, so if you take that as a given and then you take a low prior probability as we did in the sensitivity analysis, that is that clinicians may be really lousy at

evaluating patients with cognitive or possible cognitive impairment, that evaluating them poorly admittedly, we're not advocating that, but evaluating
.00029
patients poorly, simply patients who have some cognitive impairment, and then proceeding from that point and saying okay, let's try treatment, empiric treatment. Not a great idea, we're not recommending it, but if you were to take that strategy, then the empiric treatment strategy would be dominant.
DR. BROOK: So, let's go over this. If you have an 80-year old person --
DR. SOX: Bob, what I would like to do now is just focus on issues of clarification of the analysis and --
DR. BROOK: That's what I'm trying to do.
DR. SOX: Let me just continue. And to leave other questions really for the discussion period. So continue your line of questioning, but not too long, so we can give a chance for the public to have their say. Go ahead.
DR. BROOK: I'm fine.
DR. SOX: Alan, a question of clarification?
DR. GARBER: David, don't step down yet. I was just a little unclear about how the utility of complications was calculated and I'm not sure I read the report correctly, but it sounds like the reason you get a dysutility is because you no longer get the
.00030
benefits of treatment mainly, or was there an additional penalty due to the complications themselves, and how did you model that?
DR. MATCHAR: Well, first of all, in generating the model we allowed there to be all kinds of bad things that could happen with treatment. We included the possibility that first of all, that not everybody would have a complication, that there would be a certain probability of having complications. Complications could lead to decreased quality of life, complications could lead to mortality, and complications could lead to actual progression of disease, and we considered that as a possibility only

because one never knows that a future treatment may be effective for some people but actually may make some people worse.
And so, we really did look at complications as a more general phenomenon, not just this issue of decreased quality of life. But if we only focus on the base case, we're talking about say the 15 percent of individuals who have the complications and then two things happen, one is that they have a transient decrement in quality of life, which in the scheme of utility is relatively minuscule. And the second is the indirect phenomena,

.00031

which is that people who had complications were assumed to stop treatment.
DR. SOX: Are there any other questions from the panel?
The next step then in our deliberations is to hear from members of the audience, and we'll start with scheduled public comments, and Janet is going to say a few words about the ground rules here.
MS. ANDERSON: Today's first speaker is going to be Dr. Gary Small, who will be followed by Dr. Silverman, and then the rest of the list is available out on the front table if anybody wants it.
I just want to remind our speakers that for the record you will state your name, your affiliation, and disclose any conflicts or other professional information that we should know, and that we are keeping everyone to 10 minutes.
DR. SMALL: Thank you, Janet. My name is Dr. Gary Small. I am director of the imaging and genetics at UCLA, part of the Alzheimer's Disease Search Center. I am also a member of the American Academy of Neurology practice parameter committee on dementia, and I chair the PET subcommittee for the neuroimaging work group for the Alzheimer's Association.

.00032

I can add that that subcommittee is in favor of the use of PET to assist in the early diagnosis of dementia, and also that the Alzheimer's Association is a national organization that

represents the millions of caregivers and family members who are dealing with this devastating illness.

In terms of apparent conflicts of interest, I have done research on PET scanning, taken care of Alzheimer patients for nearly 20 years now, and also used PET in my clinical care of patients, and also consult with the several companies that make cholinesterase inhibitor drugs for AD.

Dementia and AD affect about 8 percent of people 65 years or older. Four million patients in the U.S. have an annual estimated cost to the United States of over $100 billion. There is compelling evidence that patients with probable AD benefit from cholinesterase inhibitor treatment, and they benefit in terms of their mental function, their behavior, the activities of daily living and the caregiver burden. There is also no evidence to treat patients without a specific diagnosis with any pharmacologic agent in this area or certainly in any area of medicine.

.00033

A definition for dementia, acquired syndrome of decline in memory and at least one other cognitive function such as language or other mental domains sufficient to affect daily life in older persons. So in a sense, a dementia diagnosis is a severity diagnosis, because we know these are very subtle changes that progress gradually over the years.

If someone has dementia, there is a long list for the differential diagnosis, but just to summarize it as we heard earlier, about 55 percent of cases are going to be AD and about 45 percent will be some other condition. It could be a neurodegenerative disorder, it could be other conditions, it could be thyroid illness and so forth. If we can accurately diagnose this 55 percent, then there is going to be a 45 percent savings in terms of side effects or other risks in treating people who don't have the disease that we're concerned about.

As we have heard just a moment ago, there are issues in terms of what's going on in the

community with primary care physicians in the challenges with the dementia diagnosis. Primary care physicians care for most patients, about 65 percent, who have dementia. Usually they do not use the

.00034

standard American Academy of Neurology diagnostic criteria. A recent study found that only 40 percent of primary care physicians, compared with 97 percent of experts, even knew that AD was the most common cause of late life dementia. Misdiagnosis is prevalent.

Callahan and associates did a recent study where they found that 75 percent of patients with moderate dementia were not diagnosed, and when you get to mild dementia, it's as high as 97 percent. We know that under recognition of dementia leads to consequences, higher hospitalization rates, emergency room visits, motor vehicle accidents, medication errors and mortality.

The current diagnostic assessment is summarized here and there are different iterations of this, such as what we see with the American Academy of Neurology, where there's a history, talking with caregivers, looking at the course of the illness and exam, standardized mental tests, functional assessment and laboratory assessment, primarily used to rule out other treatable illnesses aside from AD.

What would be some of the benefit of early diagnosis? This is just one example from a treatment study of one of the cholinesterase inhibitor drugs

.00035

and there are similar data like this. If you look at the vertical axis, this is improved cognition going up, and this is number of months in the study. This is a double blind controlled study and this is the active drug group showing better cognition after six months, the placebo group showing a decline. And after six months, they put all the placebo patients onto active drug, and you can see there is some improvement but never quite to the level that patients might have been had they started earlier on the cholinesterase inhibitor therapy.

Now certainly there are issues about trial

design and dropouts and so forth, but this is something that has been replicated, and it suggests one explanation and that is, if we start people earlier, we're going to have better benefit. We've heard about some of the data as far as clinical assessments, and the American Academy of Neurology panel after looking at 7,000 articles in the literature, found only one study with Class I evidence looking at early diagnosis of dementia using the so-called conventional clinical approach. And this is multiple clinical assessments over years. They had 130 autopsy confirmed outcome cases. Diagnostic accuracy included sensitivities in the 83

.00036

to 85 percent range, and specificities in the 50 to 55 percent range. The study that Silverman, et al., did recently looking at just a single baseline PET scan in 284 patients and of these, a comparable number had autopsy outcome. We can see higher sensitivities and specificities, sensitivities in the 93 to 95 percent range, and specificities in the 73 to 78 percent range. So PET assists with early dementia diagnosis; AD is prevalent, it's costly but can be treated, especially early on. The current approach to the dementia diagnosis involves multiple costly assessments performed over years. PET provides early positive differential diagnosis for Alzheimer's and other dementias, and the classic Alzheimer's PET pattern can appear years before the disease can be confirmed clinically. Here's just one case example, we have many. A patient who came to our UCLA memory clinic, she had been diagnosed with depression after over two years of multiple evaluations including serial MRI scans. We finally gave her a PET scan and you can see the deficit in the parietal area. We started her on a cholinergic treatment and actually both her mood

.00037

and her cognitive symptoms improved within a matter of a month. We have heard about the treat all scenario

and that there is no downside to cholinesterase inhibitors. I would just like to raise a couple of points in that matter. First, the data from clinical trials do show side effects that you can see; for example, rates of nausea ranging from 5 to 50 percent in active drug groups, compared with 3 to 28 percent in placebo groups. The dropout rates in the clinical trials due to adverse events range from 7 to 32 percent in the active drug group versus 1 to 8 percent in placebo groups.

And these are clinical trials where patients are relatively healthy. In clinical practice our patients have multiple medical conditions, which increases the side effect rate. And there are certainly multiple case reports of more serious side effect such as convulsions, violence, mania, hepatitis, and other conditions.

What would be the practical consequences of improved diagnostic accuracy? We'd have more accurate diagnostic information education, that would reduce family and caregiver burden. It's not just getting people on the right treatment. There is a

decreased likelihood of repeated diagnostic assessments and testing. We have heard about this idea of AD labeling and the problem with that. A recent study actually showed that when patients and family members got the diagnostic label, that it actually improved the caregiver attitudes and the interactions with patients.

Also, information about the disease improves quality of life for the patients, for the family members, and Mary Mittelman's group actually showed that it delayed nursing home placement by as much as a year.

Other practical consequences, early accurate diagnosis and treatment will maintain patients at higher levels of functioning, leading to fewer physician and hospital visits. It will reduce caregiver burden, it will delay nursing home placement, we've already heard about that. It reduces the use of other psychotropic medications.

The diagnosis of dementia is missed in a

large proportion of our patients. Current clinical approach to dementia diagnosis is often inaccurate. It involves multiple examinations over years. The current treatments are effective but they do have side effects, so we want to find the right patient
.00039
group to treat. PET adds to the current clinical approach by improving this early diagnostic accuracy and reducing the need for repeated clinical examinations.
Finally, when should PET be used? I think it should be used to assist in the early diagnosis of dementia. And what would be the effect of using PET? We would have more accurate earlier diagnosis, there would be better treatment outcomes, there would be fewer unnecessary clinical assessments, and we would have earlier treatment when drugs are most effective.
Thank you very much for your attention.
DR. SOX: Thank you, Dr. Small. We have time for a couple of brief questions and brief responses by way of clarification. Barbara.
DR. McNEIL: I have one question and it may be moot. Is somebody presenting from the Alzheimer's Association? Can you reconcile your statement that the Alzheimer's Association recommends PET with the written statement that we just got staying that it doesn't?
DR. SMALL: My statement was on behalf of chairing the PET subcommittee for the neuroimaging work group. This is a committee that right now is meeting, we're deliberating. We have a draft of a
.00040
report that should be available very soon, before the World Alzheimer's Congress when we're going to be reviewing that.
DR. McNEIL: So it's not discordant then, it just hasn't bumped all the way up to the top.
DR. SMALL: It's a more recent deliberation.
DR. SOX: Other questions for Dr. Small?
In that case, thank you very much, and we will move along. Our next speaker is Daniel Silverman, from UCLA, and you can introduce yourself and remind us of

any conflicts you have.
DR. SILVERMAN: Thanks first to Ms. Anderson and the committee for the opportunity to present here today. I am Dan Silverman, from the UCLA school of medicine, and I have no financial conflicts of interest in any firm representing anything to do with today's topics.
So the question before the committee today is, should this executive committee ratify the recommendations of the Diagnostic Imaging Panel? And I'm going to be discussing numerous failures in the process of the Diagnostic Imaging Panel that would preclude that from happening.
We would all like to have had an outcome

.00041

that we could be confident in from the Diagnostic Imaging Panel's proceedings but unfortunately, the truth of the matter is that we can't be because of many of these failures. I will focus first of all on failures in procedure, and these fall into three major categories, two of which I will talk about now.
First, the imaging panel actually failed to vote on the question that they were asked by the executive committee to address and that they had agreed to address, and so whatever they ended up deciding actually is only minimally relevant to the question that was posed to them by CMS and the executive committee, and the question they actually agreed to address.
And then secondly, that they failed to follow the approved guidelines for evaluating diagnostic tests as have been explicitly described by the MCAC executive committee earlier in that year.
So, I will begin with the first of these failures, that is, on the voting question issue, and I will trace back to where this first began to be formulated somewhat concretely, and that really was in May of 2001 when CMS submitted a formal request to the Agency for Health Care Research and Quality to produce a quality assessment, quote, on the use of

.00042

PET and other neuroimaging techniques in the diagnosis and management of dementia.

And then secondly, a month later this executive committee actually met to discuss, quote, the contents of and framing the questions for a future presentation of neuroimaging for dementia, to be presented to the Diagnostic Imaging Panel. So it was clear that that was supposed to be the proceeding that led, gave specific direction to what it was that the Diagnostic Imaging Panel was supposed to address.
And then during the executive committee meeting, the question was actually extended somewhat as the refined their review of what the question should be, and so they also addressed the issue of patients who do not yet meet the clinical criteria for dementia. And one expert on AD was called to provide input, and that's Dr. Marilyn Albert who we also have the benefit of hearing her expertise this morning.
She stated with regard to the potential value of the preclinical patients being tested, quote, people who have complaints and concerns about their memory problems are going to clinicians for evaluation, and those people are very difficult to

.00043

evaluate. And she went on to say, even the treatments we have now do slow up the disease, and it's pretty clear that the earlier you take them, the more beneficial they are. So if you could identify people in the preclinical phase of the disease, then treatment intervention would be beneficial and there would be a great worth in that.
And that was actually concurred upon by the executive committee and it was done so despite that it was acknowledged that there was no direct evidence for treating these patients, because all the studies had actually been done with patients who had probable AD and in one case, also possible AD, but that based on the strength of the logic of the connections that had been established between getting to patients who are earlier in the process and the more benefit the earlier they were, that it was considered to be acceptable if you could find those patients, that that would be accepted as patients who

could benefit from treatment.
And so, the Diagnostic Imaging Panel convened on January 10th to carry out its charge, and the voting question which was set forth in the panel's information packet and stated at the outset of the meeting was, as Dr. Papatheofanis also pointed
.00044
out, is the evidence adequate to demonstrate that PET has clinical benefit in evaluating patients with suspected AD, which is a very good translation of what CMS and the executive committee had asked them to do, as long as you recognize that suspected AD includes a wide range of people who had been discussed in those meetings, that is, patients who have dementia; patients who are preclinical both with MCI and also as was discussed, who don't meet the criteria for MCI but have age associated memory impairment; and patients who have dementia can fall in categories of those who have possible, probable AD, and those who don't have possible or probable AD.
But this Diagnostic Imaging Panel neer voted on the voting question. Why not? Well, in the final minutes of the meeting on January 10th, the voting question was withdrawn and another motion was made, you can find this on page 184 of the transcript, to vote on a substitute question. But it was unclear even to the person making the motion, exactly what that substitute question should be.
And after what amounted to seven or eight pages in the transcript, about five to ten minutes of discussion, the discussion culminated in the following way, with the chairman, Dr. Papatheofanis
.00045
saying, are you recommending then in the wording of the voting question that we change the word suspected to possible or probable AD.
Dr. Albert responded first. She said, well, I think so. And Dr. Papatheofanis said Kim, because he was the person who made the initial motion, and he said I'm just trying to differentiate this from MCI so we don't sort of overstep what has very little information which you notice, goes directly against essentially the spirit of what the

executive committee had asked them to address. But then Dr. Papatheofanis says, would you prefer possible probable.
And I just wanted to take a moment to say what was really happening here is after a year and a half process and intensely over the last 11 months, where there was much work being done by many people: By CMS; by the Diagnostic Imaging Panel; by the group at UCLA who initiated the coverage request and provided documentation; by the Duke group who actually did a very elaborate model on this, for a carefully crafted well-defined question about what it is they should be addressing. They ended up in the final minutes of the meeting basically turning it into what the preference was of a person who made a
.00046
motion, disregarding all that had happened before that point.
Okay. Since the time is short, I will turn now to a second category, and that was failure to follow approved guidelines for evaluating diagnostic tests as has been described by the executive committee. As stated during the executive committee's meeting on June 14th, the MCAC criteria is applied to this issue, and these criteria in broad form, which were outlined in the early part of 2001 and ratified in May of 2001, would involve examining the test accuracy, examining the impact of improved accuracy on management, and then the impact of change in management on health outcomes.
And as you heard several people discuss, there was no question about the high accuracy of PET. It was estimated by our group to be 85 percent, it was estimated by the Duke group to be 87 percent overall, given the prevalence that was estimated, and it was in fact one of the four outcome measures of the Duke group that was looked at was diagnostic accuracy that PET was a clear winner on, if you included that in your evaluation.
Also as you heard many people testify today, there was no question on the impact of change
.00047
in management on health outcomes. If patients are

properly treated, they do much better keeping them out of nursing homes, for example, anywhere from six months to 30 months depending on which studies you look at.
So the only question remaining was to look at what the impact of improved accuracy on management would be, and that was never examined by the Duke group. There are several ways in which you could do that. One of them would be to actually consider empirical data on the effect of diagnosis on management that actually occurs. You could look, for example, when physicians get a diagnosis of possible AD, how often do they actually treat patients with the cholinesterase inhibitors. And likewise with probable AD and likewise when they have mild cognitive impairment, and you could either get empirical data or model how that would change if in fact they had additional information from say a PET, when they had evidence that actually in the brain, the Alzheimer's type changes had already begun to occur. But they didn't do that.
Another thing they could have done is to consider what the impact should be based on the professional organizations' guidelines as Dr. Small

.00048

reviewed, in which they state that in patients who have probable AD should be treated, and that there is no evidence to support treatment of any other group of patients. But they didn't do that either.
And they didn't do what often is done, which is just to make an assumption that you will treat patients according to what their diagnosis is. This for example happened when PET was being considered and passed a couple years ago for staging lung cancer, where if you have a patients where you start off with a Stage I lung cancer, and the PET finds a distant metastasis, which makes them Stage IV, you don't treat them like a Stage I anymore, you treat them like a Stage IV, and if you don't find that distant metastasis, then you do the surgery to cure them by treating them as a Stage I. But that wasn't done either.
What was done actually is that the Duke

group substituted a hypothetical approach without providing any literature documented empirical evidence to support it.
And then finally, I will turn to failures in analysis -- well, actually my light is turning yellow, and since I have time to make a conclusion, I will just do this very briefly and say that although

.00049

there was a two-way sensitivity analysis, almost half of the cases that you see in green here that could occur, were cases in which PET would win. And you did hear Dr. McNeil and Dr. Matchar at the time say that you would have to have a substantial dysutility to increase, but what they didn't point out is that it would take actually a very small dysutility if you decreased estimate of efficacy.
And if you look at the Duke model and how they actually estimated the baseline efficacy case, this occurs on pages 33 and 35 of the model, there is only one paper that was used to estimate the efficacy. And if you look at the 99 percent confidence intervals in that paper, the efficacy actually would range anywhere from .98, all the way at this end, to .58. So if we just take something like halfway between, it would take a very small efficacy change to create a very small dysutility that is needed to actually get to the area that PET wins. And this committee, the Diagnostic Imaging Panel just did not consider fully the green area, to put it in simple terms.
So, conclusion, although the central question that CMS and the executive committee asked the Diagnostic Imaging Panel to assess was properly

.00050

slated by the panel staff to be the voting question, it was never voted upon and a different voting question driven by different assumptions, assumptions that were contrary to what the executive committee had charged them with, was substituted in the final minutes of the meeting, resulting in a failure to adequately address the central question.
Secondly, the Duke center model didn't meet the needs of the Diagnostic Imaging Panel for

addressing the question fully because they didn't address that central middle issue, and as a consequence, the panel felt that it had insufficient evidence available to it with which to support the Medicare coverage request. The evidence does exist, it just wasn't given adequate consideration. So in answer to the question we started with, should this committee ratify the findings, the answer is no, and I will refer you then to the handout to what it is that we recommend actually that CMS do. Thank you for your time.

DR. SOX: Thank you very much, Dr. Silverman. Why don't you stay up at the mike in case somebody has some questions for you. Does anybody have any questions of clarification for Dr. Silverman? Yes, Tom.

.00051

DR. HOLOHAN: Doctor, in the slides you presented, you quote Dr. Albert, and tell me if I'm reading into this an implication that's not there. But the quotations that are provided give me the impression that you believe that she supports earlier diagnosis. For example, quote, if you could identify people in the preclinical phase, treatment intervention would be beneficial, there would be great worth in that. That's difficult for me to understand considering the April 16th letter from Dr. Albert, who clearly states that speaking on behalf of the AD Foundation, she believes that at the present time, PET is an experimental procedure. So, it seems to me this is not exactly the position that she took when she spoke to us, nor that she has taken in writing.

DR. SMALL: These are actually easy to reconcile. The second point you made is what she is saying is the position of the Alzheimer's Association overall. The first point, the point that is actually in my presentation, is the proceedings of the executive committee to decide what would be considered to be reasonable evidence on which to base whether PET would be considered to be valuable, and it was being acknowledged that although it would be

.00052

very valuable to find patients who are at an earlier stage than you can make the diagnosis accurately where you are just relying on clinical evaluation, that there was no direct evidence because randomized control trials haven't been published yet looking at those patients, but the strength of the evidence was so strong that if you treated patients earlier in the course of the disease that they would do better, that it was worth evaluating patients to see if you could find those patients who are early.
And there is at least six years of very strong evidence in PET that you could find those patients years before you could make the diagnosis accurately by clinical evaluation alone.
DR. SOX: Yes, Frank.
DR. PAPATHEOFANIS: Very briefly. Just to sort of clear the record a little bit for the panel, if you refer to the minutes and the final panel recommendation, you will see the items that we voted on, and I think the items we voted on were clearly the ones that we were charged to vote on. We also had Sean sitting there beside us, clearly asked and I'm sure the minutes will reflect this, whether or not we had addressed the Agency's concerns and whether our deliberations had met with their

.00053

agreement.
So I think clearly the questions that we were posed were addressed and were voted on.
DR. SILVERMAN: Can I comment on that?
DR. PAPATHEOFANIS: The other comment I want to make, and this is just a very brief one, is that we have seen two speakers now who are making a strong point that the earlier we treat patients, the better the diagnoses for addressing patients with presumed disease, the better off they will be. I don't think anyone in this room could agree any stronger than I do. I think that what we learned from the evidence was that PET isn't there and it's still, the evidence wasn't there to say that that is the technology that will provide that.
DR. SOX: Okay. We'll have a chance, Dr. Silverman, to get into a discussion during the

discussion period. I would like now to try to keep things mostly to questions of clarification. So, Bob.

DR. BROOK: I'm trying to sort this out in clarification. The first question is whether this is a routine thing that ought to be done, and I think that's what the voting panel has done, but I was curious when I heard your comment here. This is a

.00054

patient who obviously presents because there is a problem with memory or something, and the doctor, presumably a good doctor, doesn't think there's adequate evidence to justify prescribing for some reason and wants to do a PET scan on that patient, which let's say is one-half of one percent of all people that doctor sees, because that's where we are at right now. Did the panel think about this case, is this one of the cases that was considered?

DR. PAPATHEOFANIS: It was, and it's in the language that precedes the description of the motion.

DR. BROOK: And even after considering this case -- so you considered explicitly this problem?

DR. PAPATHEOFANIS: Not explicitly, but something in the vicinity of a patient like this.

DR. BROOK: And where the PET is being used not to rule out AD but rule it in in a confusing patient, like the clinical example that you saw of somebody that is confusing even the best diagnosticians in the land and PET is being used to rule it in.

DR. PAPATHEOFANIS: We didn't have the best diagnosticians in the land but the average.

.00055

DR. BROOK: Okay.

DR. SMALL: Actually, the committee explicitly failed to look at this case. You can find this in the final minutes of the transcript, in which it was pointed out that they never even looked at the case and never voted on the case of patients who had a clinical criteria for dementia but were thought not to have AD, that the evidence wasn't strong enough to

diagnose them with possible or probable AD. And the question was raised at that point, what are you going to do about the fact that we're charged with the committee to answer that question and you never answered it, and then, a movement was made to adjourn at that point.

DR. SOX: Okay. I just want to remind everybody, just so we try to keep things orderly, let's try as much as possible to stick with questions for the speaker, and then we can get into discussion amongst ourselves at the appropriate time. Leslie, did you have a question?

DR. FRANCIS: This may be that I'm just foggy from a night in the Atlanta airport so, I wanted you to say again for me what you understand the difference to be between your slide that was the language actually voted on and the language you

.00056

wanted the panel to have voted on, the difference, that is, between suspected and possible or probable. Exactly what group of patients would have been handled but wasn't, or could you just put up the two slides.

DR. SILVERMAN: That is a very important question and I am glad you gave me a chance to clarify that. So, the question that was supposed to be voted upon was suspected AD, and it was clear from the proceedings of the executive committee last year that that included several groups. It includes people who have dementia and it includes people who don't yet have dementia but probably have or may have incipient AD, they just don't meet the criteria of severity yet enough to call it a dementia. And in mild cognitive impairment, at least 50 percent of those people, according to the American Academy of Neurology papers reviewed, would actually have incipient AD.

The Duke center actually estimated 80 percent of them have incipient AD. It would also include, and Dr. Albert talked about this group too, patients who don't strictly meet the definition of mild cognitive impairment but who do have memory complaints, what we sometimes call age associated

.00057
memory impairment, and at least 10 percent of those people after the age of 65 have incipient AD in their brains.
And it would include people who have dementia, of course, who have the diagnosis of possible or probable AD, which based on the inadequate sensitivity and specificity of clinical evaluation, about 75 percent of them actually have AD. And it would include people who have dementia who aren't thought to have AD, and based on the inadequate sensitivity and specificity, actually about 15 percent of them do have AD.
And that entire question ended up being narrowed down to one voting question initially, which was possible or probable AD, which of course is the group that you would least expect to benefit from PET, because they already have gone through the two to three years of making that, that typically occurs to make that diagnosis accurately, and time has already elapsed that we are trying to save.
And so then at that point, and Dr. Papatheofanis alluded to this, Dr. Tunis essentially said wait a minute, guys, what about this whole group of mild cognitive impairment patients that you were charged to look at and you haven't even voted on
.00058
them. And then a second motion was made to vote on mild cognitive impairment which, by taking it apart in this was, was led to say well, we don't have enough evidence in treating mild cognitive impairment that is useful, and so we'll dispense with that.
And what they still haven't voted upon was the patients who don't meet mild cognitive impairment definition, like the age associated impairment, and most importantly, the patients who are demented but are thought to have possible or probable AD, and a number of them actually do. Which means that if you treated patients according to the diagnosis, they would fail to get the treatment that everyone here agrees could benefit those patients tremendously.
DR. SOX: Okay. I think it's time to move on. Thank you very much, Dr. Silverman, and we will

now hear from Dr. Peter Conti. Would you please introduce yourself, state any conflicts that you may have.

DR. CONTI: Good morning, Mr. Chairman and members of the panel, ladies and gentlemen. My name is Peter Conti. I am associate professor of radiology at the University of Southern California. As far as conflicts are concerned, I have a number of consulting opportunities for equipment manufacturers

.00059

and service providers for PET radiotracers and PET procedures. I am also funded by several government agencies to conduct research in the field of PET, but I am now speaking on behalf of the Society of Nuclear Medicine.

I would like to offer our strong support to the addition of AD as a CMS reimbursable indication for FDG PET. We would like to reiterate our position presented to the diagnostic imaging panel on January 20th, 2002.

Right now more than 19 million Americans are estimated to be caring for someone with AD, and home care for a person whose disease has progressed is estimated to cost about $47,000 per year. By the middle of the next century, as many as 14 million of today's baby boomers could have AD. As you know, the standard wisdom is that there is no definitive way to diagnose AD other than by brain biopsy or autopsy. The information compiled by the UCLA group and presented to CMS from studies all over the United States strongly supports the value of PET as an alternative diagnostic approach for this devastating condition. You will hear in a few moments additional material from the University of Michigan which will be presented by Dr. Kirk Frey, independently

.00060

supporting these conclusions.

We believe that there are compelling reasons why PET is a valuable tool for physicians attempting to determine whether the memory lapses and behavior patterns seen in these patients are due to Alzheimer's disease or to some other process. Number one, since FDG PET is more effective than clinical

examination for the differential diagnosis and identification of various dementia causes, the greater diagnostic accuracy provided by PET early in the course of dementia illness will lead to more effective disease management.

PET enables physicians to clearly identify and differentiate between different types of dementia. This can be critical, not only for the treatment of these other diseases, but for the initiation of Alzheimer's specific medication.

Third, notwithstanding the potential for therapeutic intervention, the usefulness of FDG PET is important for patient quality of life. Specifically, additional certainty with respect to the diagnosis will help the patient and family make more appropriate life decisions. In addition, the increased certainty may help family members cope with the condition; for example, depression affects more

.00061

than half of primary family caregivers and uncertainty about the diagnosis may contribute to family and caregiver feelings of depression and helplessness. A negative study would be of value to patients as well, as it can predict the absence of further cognitive impairment with fairly high certainty, which could well affect the decisions the patient and family make about the future, such as retirement, moving or staying near home, not taking a cholinesterase inhibitor, et cetera.

In short, the radiopharmaceutical FDG with PET can be used to assist with the characterization of early dementia in geriatric patients for whom the differential diagnosis includes one or more kinds of neurodegenerative disease associated with the dementia process.

We believe in particular, this is particularly helpful in this patient population when there has been a change in cognitive status where the etiology is not apparent or when symptoms are not reversed in a reasonable amount of time. Providing families and physicians with the means to better manage those with this disease would seem a more cost effective approach to care. We believe this approach

should include access to and reimbursement for PET
.00062
scans.
We disagree with the conclusions and
recommendations offered by the Diagnostic Imaging
Panel on January 20th, 2002. The reasons for this
disagreement include the following: Number one, the
analysis performed by the outside Duke University
team corroborated the fact that PET is in fact more
accurate compared to standard methods of detection of
AD; perhaps that's an area of agreement rather than
disagreement. Two, the analysis that includes the
treat-all strategy used by the Duke team has no
precedent, at least none that was offered to the
community during the debate; this potentially limits
the usefulness of the results obtained.
A more traditional approach such as
looking at equivalents or superiority to the existing
methods should have been applied. In fact, time does
not permit me to list the numbers of procedures that
are done in practice today which are covered by CMS
that do not affect the outcome of patients, or
diseases that have no treatment.
There was no accounting, third, of the
extra burdens that's placed on physicians and
patients in determining who should be treated and
when. The treat-all strategy is not practical from
.00063
cost, drug interaction and time to delay diagnosis
perspectives. Elderly patients taking multiple drugs
for other disorders with limited budgets cannot be
expected to comply. Clearly, more accurately
identifying patients with disease would lead to a
more informed decision as to whether to treat such a
vulnerable population. In fact, it would likely be a
requirement for CMS, given the latest information
from the White House to develop appropriate
diagnostic strategies for identifying patients that
have this disease.
Fourth, therapeutic options will change
over time. Any policy decision should have the
flexibility to withstand the flow of changes in
healthcare, especially treatment options. A serious

problem has now been introduced regarding treatment development for these patients. From the data presented by the Duke team a large position of patients may be misdiagnosed with AD if traditional diagnostic methods continue to be employed. This means that because the entrance criteria for investigator or drug company sponsored trials will permit a large mixed patient population, results will be significantly biased. Certainly tightening the criteria used by using PET, which they

.00064

admit improve diagnostic accuracy, is more appropriate.
We therefore urge you to reject the conclusions reached by the Diagnostic Imaging Panel and add AD to the list of reimbursable indications. In the event you consider approving the indication, we have provided written language available for use in formulating the indication, and I will just read that very briefly.
This is the proposed language for use of FDG PET in dementia. An adjunct test to assist the diagnosis of early dementia in geriatric patients for whom the differential diagnosis includes one or more kinds of neurodegenerative disease; for example, AD, and frontal temporal dementia. The criteria for appropriate use of PET FDG in the evaluation of dementia should include, one, presentation of cognitive impairment or behavioral problems representing a change from the patient's normal level of functioning. Two, the etiology of symptoms not apparent or symptoms not reversed with reasonable amount of time following initial evaluation and management emanating from standard workup. And third, patient is not suffering from a state of advanced dementia at the time the PET scan is

.00065

acquired.
I would now like to take the opportunity, if there are no questions, to introduce Dr. Kirk Frey, from the University of Michigan. He is the president of the Brain Imaging Council of the Society of Nuclear Medicine. I will take some questions

before Dr. Frey does take the podium.
DR. SOX: Questions of clarification for Dr. Conti? Any questions from the panel? In that case, proceed.
DR. FREY: Thank you for the opportunity to address the committee this morning. I'm Kirk Frey, from the University of Michigan. I am a professor in the Departments of Neurology and Radiology, and I come to you wearing two hats, one as an investigator involved in designing and applying tests of brain chemistry and function, particularly with an interest in neurodegenerative disorders like Alzheimer's and Parkinson's diseases and second as a practicing neurologist in a geriatric referral clinic where patients with these kinds of problems present for evaluation.
Let me share with you briefly some data from our laboratories that were not available to the committee for review because they had been abstracted
.00066
but not yet formally published in the reviewed medical literature. These support the conclusion regarding the diagnostic impact and accuracy of PET in the evaluation of patients with suspected AD but they come from an unselected clinical population recruited in our university.
This study was a prospective design. We went to our clinics and enrolled patients who were presenting at their initial visit for a newly discovered cognitive complaint or abnormality. We included two clinics to select these patients, one a specialty clinic in geriatric neurology where I practice, and the second was a general neurology clinic at our adjacent Veteran's Hospital, which is more representative of primary contact general neurology.
All patients underwent a standardized clinical evaluation as recommended by the AAN for evaluation of suspected dementia. We then employed a two-year follow-up period to establish an operational clinical diagnosis. And we distinguished patients into groups of progression of their cognitive complaint versus no progression. At the entry into

this study we performed an FDG PET scan, but this was not utilized in formulating the patient's clinical
.00067
diagnosis.
We analyzed this PET scan objectively according to the presence or absence of a pattern, which I loosely refer to here as the AD pattern. This was done with an objective tool to look at the metabolic deficit in association cortical areas relative to primary sensory motor cord disease. This cartoon shows the pattern of FDG that we used. I won't spend your time here other than to say there are a number of recognized patterns and we combined all of those that are associated with degenerative dementias into those we considered positive. Yes.
DR. FRANCIS: Were the results of the PET scan known to the people doing the clinical evaluation?
DR. FREY: No, they were not, and in terms of making the clinical distinction here, we abstracted from the patient's chart the objective neurological findings rather than looking at the treating physician's clinical diagnosis. So we enrolled 116 patients. Of these, 90 completed their two-year clinical follow-up. There were 26 where incomplete data were available to us, and to look at whether those might have influenced
.00068
the result due to dropout, we found that of those 26 patients that didn't complete the follow-up, 10 were felt to be not initially demented on the basis of neuropsychometric and clinical testing; 15 were demented on that initial exam.
And let's look at those. If we took the PET data from those 26 patients, those who were not initially demented all had a PET pattern which did not indicate the presence of AD. 9 of the 16 who were demented on their initial examination did have a positive AD pattern. And the reasons for withdrawal are listed here on the right. Largely these were patients who came from a distance to our medical center; many died without autopsy or were unable to

be located for follow-up.
With regard to the patients who did complete, these are the findings. The numbers here reflect the fraction of PET scans that were interpreted by our objective estimator as positive for the AD pattern and you see first amongst those with a progressive cognitive decline over two years, we had an 80 percent sensitivity.
We found 7 of 31 patients had an AD like pattern but did not have a progressive dementia. It turns out that these represented two areas of false
.00069
positivity. First, there were patients who had relatively advanced multiple sclerosis. This was evident to the referring physician on the basis of magnetic resonance imaging but this was not available to us in terms of interpreting the PET data. So this was done in an isolated and blinded fashion.
The other patients in this group suffered from alcohol and polysubstance abuse. When one looked at the fraction of patients who actually satisfied after two years the clinical research diagnostic criteria for probable or definite AD and compare those against patients who ultimately were neurologically normal, that is, these are patients with an unsubstantiated complaint or who had depression or other psychological problem which resolved, we found that the sensitivity of the FDG PET was 89 percent and the specificity 100 percent. Overall, the diagnostic accuracy here was 90 percent and data agree very favorably with the data that the diagnostic panel had for review, but this now is an unbiased prospectively defined clinical population, which is exactly the situation where we would argue the diagnostic impact of PET might best be identified.
So as I have already told you, we found
.00070
some false positives. For the most part, these would have been clinically evident by combining the PET data with the routine clinical information collected by the neurologists at the time. False negative studies, I should point out that none of the patients

who progressed actually had a completely normal FDG scan. Instead, they had patterns that we were not willing to call diagnostically positive, and it turns out that as a result of this, and let me just skip forward, we actually identified that we could increase the sensitivity. That is, if we looked at the patients who had mild cognitive impairment at the time of entry, rather than limiting ourselves to the neocortical changes here in the parietal lobe and here in the frontal lobe, which were our a priori assumptions of areas to look at, it turns out that the posterior singular cortex is actually the most sensitive area to look for very early changes in dementia, and all of these cases that we identified as negative who later progressed showed a decrease in this posterior singular region.

So, if we were to modify further our objective criteria as to how to interpret FDG, we think that its performance would be even better than what we reported here.

.00071

Let me now change hats from that of the imaging scientist to the clinician in the geriatric neurology clinic dealing on a daily basis with patients who come with a new complaint. Patients come to us expecting two results; they come to us with hopes and with expectations. They hope that the problem they come to us to evaluate is something where we can have a positive impact and resolve for them, but they expect to receive an accurate diagnosis and an accurate prognosis as to what the future holds for them and what the medical explanation for their complaint is.

The assessment done by the Duke EPC, I think addresses the issue of patient hope very nicely. Given the somewhat limited impact of available treatment for AD at the present time, it can be concluded based on the data manipulations that were made that the treat-all strategy is somewhat better than the treat-none strategy, and because the impact is relatively difficult to demonstrate, particularly with more advanced Alzheimer's cases, the change in terms of life expectancy or quality of

life may not be a large measure. On the other hand, the impact of actually making an accurate and correct medical diagnosis when
.00072
a patient presents with a complaint is not at all weighted in the Duke analysis. In fact, what we see in our clinics in Ann Arbor is a very large portion of patients who have come from an initial evaluation where a diagnosis was not made an a medication was started. The patient then fails to achieve what they expect in terms of medical benefit and it referred on to yet another evaluation by another specialist. This is a very frustrating situation for the patient. It's a very frustrating situation for the referring or for the consultant physician as well, to tell the patient that they may or may not have received an adequate diagnostic evaluation, and that you may or may not be able to answer the question as to what the underlying cause for the complaint might be. So I would encourage you to consider the value of diagnosis in a patient with a complaint that may or may not represent the early signs of AD as an important and unweighted value in your assessment so far. Thank you.
DR. SOX: Thank you. Are there any questions, clarifications for Dr. Frey? Yes, Randel?
MS. RICHNER: You said that this information was in abstract form. I'm not sure, were you aware that the Duke EPC study was going on in
.00073
evaluating this, and were you given the opportunity to provide this information to Duke?
DR. FREY: No. I became aware of it at its publication on the Internet in December.
MS. RICHNER: Okay. Because one of the things that we talked about as an executive committee a few years ago was that we would use unpublished data for evaluations of technology assessment, and I remember that discussion well with Dr. Brook and others. So I think that this to me is very compelling information that wasn't necessarily considered in this evaluation, and that should be part of the process. Once again, we're still having

problems with process. The questions again is an issue and then now this, getting data after the fact that was this very very interesting data. So I think as a point on the record, I know that this afternoon we're going to be talking again about process issues with the EC, but once again, we need to clarify what we're doing when and how and what kind of data we are going to be using in our evaluations. And also, the questions, again.

DR. FREY: But I would point out that your analysis of the published literature led to largely the same conclusions about the diagnostic performance

.00074

of this test.

DR. McNEIL: Actually that was my point, Randel. I think we as a group wanted to get all of the data that we can, and in this particular case, nice as that study is, it doesn't impact at all on the data that we used in the Duke analysis, and it essentially comes down to the same point. But it does raise the issue which we should discuss later about abstract data, and how much we include and where we get it, but I don't think that's relevant for the process issues here. I think it's a little bit of a red herring at this point.

DR. SOX: There will be plenty of time for this sort of discussion when we get after the break, so let's try to stick with questions for Dr. Frey. Frank.

DR. PAPATHEOFANIS: You described in your study at Ann Arbor that you obtained a baseline PET and then tracked the patients for two years with a clinical correlation there. Were any other noninvasive studies performed at baseline, like MR or HMP or SPECT?

DR. FREY: Yes. We performed a stereotypic evaluation based on the then current AAN guidelines, which included either a CTR and MR,

.00075

depending on whether the patient had already received neuroimaging or not, a battery of blood testing and neuropsychometric testing to document complaints that weren't objectively obvious on bedside confrontation

or neurologic exam.
DR. PAPATHEOFANIS: Were you able to evaluate their predictive strength, if you will, the MR and so forth, two years down the line?
DR. FREY: We could do that, I have not looked at it, but suffice it to say that the vast majority of these were negative, other than for the presence of additional medical conditions such as demyelinating disease in 5 of the 7 false positive studies.
DR. SOX: Bob.
DR. BROOK: If I heard your testimony correctly, you seem to agree with everything except the point of the accuracy of the diagnosis, in terms of giving patients better information.
DR. FREY: Yes, but I think that --
DR. BROOK: I mean, the treat-all versus -- I mean, you concluded by saying the only real issue is that Duke didn't consider the expectation that the patients have, the multiple visits and all that kind of stuff.

.00076

DR. FREY: And I consider that to be an absolutely insurmountable shortcoming.
DR. BROOK: I understand that, but given what you know about sensitivity and specificity, and what you heard, how would you change your message when you have somebody who presents with dementia and you have diagnosed this as -- you have done the memory tests, you know that this thing only has a sensitivity of .9, what do you tell the patient differently after you do the PET than before you did the PET?
DR. FREY: So the impact here would be --
DR. BROOK: I'm asking you as a clinician, I'm coming to you and I just couldn't understand the logic between the two parts of your testimony.
DR. FREY: The example would be a patient who has a memory complaint, who comes --
DR. BROOK: Yes. I mean, that's presumably who you're evaluating, has a memory complaint, you did everything possible, you don't know exactly what's going on, you're the best

clinician in the state of Michigan, and you now do a PET that has a sensitivity of .9, it turns out to be normal. Do you really tell me that you know any more than what you knew now? Because you're already up at
.00077
a level of this curve, and I just don't understand your testimony.
DR. FREY: The critical missing value here is that it takes two years of prospective follow-up in order to properly categorize the patient, whereas at the time they initially present to me, if I have the neuroimaging data from PET, I can achieve virtually an equivalent diagnostic specificity and specificity as if I follow the patient in the clinic telling them I don't know what you have, come back in six months, we'll see if you're worse. The analogy to this would be to turn the clock back about 20 years in the diagnosis of multiple sclerosis, where patients would shop from physician to physician to physician with a complaint that we would all objectively find on neurologic examination but not be able to make the diagnosis of --
DR. BROOK: So you're --
DR. FREY: -- demyelinating disease.
DR. SOX: Don't interrupt, Bob. Give him a chance.
DR. BROOK: I don't understand. You're using it in a positive way.
DR. FREY: Yes.
.00078
DR. BROOK: You're saying that at baseline you don't know the diagnosis, and if it's positive, you are going to tell the patient that basically they have AD, is that how you're using it?
DR. FREY: I would use it in both ways, so that if you come to me at a time I cannot objectively show that you have a progressive dementia, that you have a complaint that fits in the Alzheimer's sphere, if I have this additional side data and your scan is unequivocally abnormal, I would tell you so and we would manage your case expectantly, as though you have early incipient AD and you should have available

to you all of the healthcare resources that you have heard may have an impact, and particularly a greater impact when applied earlier in the course. Conversely, if your scan was absolutely normal, together with the other clinical information I have, I would counsel you that we do not find evidence that projects that you are going to experience a significant decline in your cognitive abilities over the next several years, and we should work behaviorally to try to maximize your abilities, to treat your depression, your anxiety and so on, and not assign you the diagnosis of probable AD and do you and your family the disservice of assuming that

.00079

you're going to have short-term further loss of abilities that are medically managed with cholinesterase therapy.

DR. BROOK: May I ask one follow-up question?

DR. SOX: Yes.

DR. BROOK: I'm still confused, and I'm really sorry. You have done a really superb workup and you're a superb doctor. And you can't tell whether this is progressive, that's what you're telling me?

DR. FREY: And that's true when you see the patient for the first time 90 percent of the time.

DR. BROOK: Absolutely. And the PET scan in that situation turned out to be normal. Would you not treat the patient?

DR. FREY: I would not treat the patient expectantly with a cholinesterase inhibitor, who's shown not --

DR. BROOK: Would you treat the patient with a cholinesterase inhibitor?

DR. FREY: No, I would not.

DR. BROOK: So you would use the PET scan and all you know about sensitivity and specificity,

.00080

and the Duke analysis, and you as a clinician would not treat any patient whose PET scan was not definitively abnormal in the absence of cognitive

decline, a history and physical exam cognitive decline. Is that what you're testifying, and that's how you think it's valuable to you as a clinician?
DR. FREY: I would qualify a little bit, and I would say this is in a setting where I am not able to make a definite clinical diagnosis, that's correct. I think this is valuable side information with independent impact.
DR. BROOK: And of all the patients that you see, what proportion of them fall into that group at the first visit?
DR. FREY: It's difficult to get an accurate estimate, but I would say between 1 in 10 and 1 in 5.
DR. BROOK: So for 1/20 to 1/10 of all of the patients you see, you would not start them, who had problems with memory, but you can't deal with memory change over time, they clearly are there and they have some memory impairment.
DR. FREY: No. They have a memory complaint. I mean, here is the difficulty. The objective tools we have for clinically evaluating
.00081
patients are relatively crude until the deficit is self evidently obvious to all in the family and everyone around. Patients are introspectively very much more sensitive to the presence of a change from their baseline than many of us in the office are able to substantiate. Most of us, at least in our clinic, do not intervene with the treatment which is established for AD until we are able to make a clinically defensible diagnosis that indeed we think the patient has AD.
DR. BROOK: I would just add one little comment. As an internist and geriatrician, if I take a sensitive enough history, I can get anybody over the age of 65 to indicate that they have a memory complaint compared to when they were 20, and if I follow your logic correctly, what you're recommending is everyone ought to get a PET scan that I as a primary care internist -- you didn't, you know, if I follow you correctly, that I as a primary internist ought to give every person literally that comes in

and says I have a memory complaint, because I surely can't, even if I send them to you or make a diagnosis of dementia on this group because they don't have any positive testing. And the bottom line is, every one of those people deserve a PET scan.

.00082

DR. FREY: Well, remember, these are patients that were referred for a neurological consultation to evaluate the problem, so it must have --

MS. RICHNER: They had already been through it.

DR. SOX: Bob, are you done?

DR. BROOK: I'm finished.

DR. SOX: Thank you. John?

DR. FERGUSON: Were all the PET scans read blindly so those who read them and interpreted the PET scans did not know the clinical course?

DR. FREY: That's correct. They knew that it was part of this study so that the patient had a cognitive complaint or concern, but they did not have access to any other clinical data at the time of the PET analysis, that's correct.

DR. FERGUSON: And were there follow-up PET scans besides the initial one?

DR. FREY: There were not unless the patient elected to enter a research protocol that might have entailed additional imaging.

DR. SOX: Thank you very much, Dr. Frey. We will now hear from Dr. Marilyn Albert, from Massachusetts General Hospital. Could you provide

.00083

any further introduction that suits you and tell us about any conflicts.

DR. ALBERT: My name is Marilyn Albert and I am professor of psychiatry and neurology at the Harvard Medical School. I am also the director of the gerontology research unit at Massachusetts General Hospital and I am co-director of a clinic, the geriatric neuro-behavioral clinic at Massachusetts General Hospital, where we see patients with cognitive decline. But, I'm speaking to you today in my position as the chair of the medical and

scientific advisory committee of the Alzheimer's Association, and I am here just to present the position of the Alzheimer's Association with respect to the issue that's before you this morning. I have no conflicts.

The Alzheimer's Association, as you probably know, is next after the Federal Government, the largest funder of research in the field of Ad, but it also carries out many other important activities, including trying to provide support and family care to patients, trying to provide information. And in the scientific domain, one of the things that it sponsors is working groups among scientists in the area that are a continuing activity

.00084

of communication among scientists and we have working groups related to diagnostic imaging. And we have working groups related to diagnosis per se, and the position of the Alzheimer's Association that I bring to you this morning comes from the activity of those working groups that have considered what the current status is of diagnosis, and what the current status is about imaging in relationship to diagnosis.

The general consensus of the members of these working groups is that the clinical diagnosis as it is currently used is the current state of the art and that while PET scanning and other forms of imaging modalities are extremely important potentially for diagnosis and for monitoring disease and evaluating the potential impact of medications, that right now they are at the experimental level and are, there is insufficient evidence to suggest that they should be substituted for the standard clinical diagnostic workup. That's really the basic position of the association and I would be happy to answer any questions you might have about that or anything else.

DR. SOX: Perhaps you could respond to this issue about treating patients with cholinesterase inhibitors when you have a 20 or 30 percent probability that the patient has a condition

.00085

that would respond to that drug, whether it would be appropriate to treat those patients.

DR. ALBERT: You're talking about very very mild patients, or anybody at any stage who has a possibility of having a disease?
DR. SOX: Patients like those in the trial.
DR. ALBERT: I was obviously correctly quoted this morning as saying that I think that anybody, if we could identify people accurately who have early disease, there would be enormous benefit. Myself and many other people around the country are doing a lot of research in specifically that area. That happens to be my own scientific area of interest, to find out how to identify people as early as possible in the course. The motivation for that work is that we believe that in the future there will be more effective treatments than there are now, and it's likely that they will have negative side effects, many more negative side effects than current treatments currently provide, and at that point it will be really essential to identify people as early as possible.
Currently the way that patients are treated in my clinic and in clinics around the
.00086
country as best I know, is that not only patients who meet the criteria for probable or possible AD are treated, but some patients who meet criteria for so-called MCI or mild cognitive impairment who don't yet meet the clinical criteria for dementia are also treated. The reason that's the case is that as you have heard, the downside of treating them is minimal and the upside is possible, although it hasn't been proven.
DR. SOX: So it sounds as though you don't have a quarrel with the analytic strategy that the Duke group took.
DR. ALBERT: No, I don't.
DR. SOX: Wade.
DR. AUBRY: I just had a question regarding the language. In your statement, you said substituted for a clinical evaluation. Is that what you meant to say, or did you mean an adjunct, or to assisting the diagnosis?

DR. ALBERT: The position of the association is that are certainly a subset of cases where doing PET scanning will improve the diagnosis but as I understand the question before the committee, that's not the question being addressed.
DR. SOX: Other questions for Dr. Albert?

.00087

Don't go away. Leslie, did you have a question?
DR. FRANCIS: I guess I would like to ask you whether you agree with what some of the other people said. I think you did, about the downside of the treat-all strategy, that is, the Duke model assumed that there is very little downside. Do you agree with that, or with the people that argue there is more of a downside, and do you manage differently patients who you start on therapy and they do show some side effects, would PET be useful in that group?
DR. ALBERT: With response to the second question, to my knowledge, there are no data that PET can identify patients who might particularly benefit from treatment, so I can tell you what's happening in clinical practice and certainly in our clinical practice is that we are more and more likely to treat people at mild or in milder stages. In our research study we have people who meet criteria for so-called MCI who are not yet clinically demented, and we feel sufficiently strongly that even though there is no proof yet that these medications are helpful that we ought to offer them and present them to people as a possible intervention.
So we provide it, and the reason that we provide it if people are interested is because it's

.00088

not likely to make them ill and it has the possibility of benefit, but we would certainly not do that if we thought that people could become acutely ill, so we would completely reevaluate the strategy if we had medications that had a larger downside. So I think we act in accordance with the model that the Duke group developed.
DR. SOX: Bob.
DR. BROOK: I would like to ask the clinical question again. The recommendations and the

process is sort of everybody, and are there any subgroups that you can think of given what you know about the sensitivity and specificity of PET, that this test would add something to the clinical care of those people, either the group that Leslie talked about, which was people that developed side effects and you want to know whether to push it over the side effects, or other diseases for which these drugs may interact with, or anything else. Is there any single subgroup that the Alzheimer's Association has looked at that believes that this thing should be standard practice, not overall, but a clinical subgroup?
DR. ALBERT: I don't believe that the Alzheimer's Association has identified a subgroup where they think this approach would be beneficial.
.00089
I can tell you from my own personal opinion clinically that if I have a patient who I think might have frontal temporal dementia, I would be very likely to try to do some functional imaging study, either PET or SPECT, but that's not the position of the Alzheimer's Association.
DR. BROOK: And why do you do this and how do you use it? I mean in sort of the model, the process that the committee used, what's the value of it in this patient?
DR. ALBERT: There is some evidence that patients with frontal temporal dementia are very likely to have frontal hypoperfusion or more unilateral hypoperfusion than is typical in the average Alzheimer case. And so, if we suspect frontal temporal dementia, and the clinical case is very confusing, then we might be likely to do a functional imaging scan. But as you were suggesting, certainly in our clinical practice we put a great deal of emphasis on getting a good history, and the best way of identifying frontal temporal dementia is still a good clinical history.
MS. RICHNER: I think the question again is, if you do the PET scan, will there be any change whatsoever in terms of how you treat the patients,
.00090
because you really don't have a lot of choices in how

it would be treated; is that right?
DR. ALBERT: We don't typically do PET scans as part of our workup.
MS. RICHNER: But in those cases where you say you would use a PET scan, would there be any change in the treatment?
DR. ALBERT: I actually don't think so, and part of the problem is that we don't have good treatments for frontal temporal dementia.
MS. RICHNER: Right, so there's no change.
DR. ALBERT: And it's possible that the cholinesterase inhibitors might help them, we don't know, so it wouldn't change treatment.
MS. RICHNER: I see, thank you.
DR. SOX: Tom.
DR. HOLOHAN: Let me just clarify what you said, which is not what's being repeated. You said PET or SPECT, did you not?
DR. ALBERT: Yes.
DR. SOX: Any other questions? Thank you very much. Now, we are running a bit behind but I'm going to press on before the break with open public comments, and I would like anybody who is interested in making comment during the 15 minutes that we have
.00091
allotted to this to raise their hands. One person, two people. Well, I'm not going to give you seven minutes each. Three minutes. So why don't you who put your hand up first, go ahead. Introduce yourself, tell us about any conflicts, and then go ahead.
DR. JOHNSON: My name is Keith Johnson. I am a neurologist at Brigham and Women's Hospital and Massachusetts General Hospital in Boston. I have specialized in the diagnosis and management of AD for about 15 years. I am also engaged in research and the practice of nuclear medicine, and I have no conflicts.
I would like to raise a couple of points that may represent benefits for the use of PET in the diagnosis of AD that may not have been adequately considered. The first is the length of time required to arrive at a confident diagnosis of AD or dementia,

or MCI. The second point is the value of knowing the diagnosis for patients and their families. First, many elderly patients present with a clear history of progressive cognition and functional decline. They satisfy criteria for AD and are begun on a cholinesterase inhibitor. However, for a substantial number of patients and an

.00092

increasing number who are in the early stages of impairment, there is a significant time lag between an initial evaluation of memory trouble and a confident diagnosis followed by appropriate treatment. This is true primarily because the diagnosis is to a large extent focused on proving that there has been progressive deterioration, i.e., observed decline in cognition and function over a period of six months to one year or more. This specific limitation of the clinical evaluation would in many cases be significantly reduced with the use of PET, because diagnostic accuracy would be increased at an earlier stage of the illness, permitting earlier initiation of drug therapy. And we now have evidence, as you've heard, that the benefit of cholinesterase therapy is reduced when the initiation of treatment is delayed until later in the course of AD. We also recognize that a less certain diagnosis is very often associated with over utilization of medical resources by individuals who are fearful of the diagnosis. The majority of dementia patients are evaluated not by specialists who are skilled in the diagnosis of AD, but rather by busy primary care or internal medicine physicians for whom an efficient

.00093

and accurate diagnostic aid would be particularly beneficial. Secondly, it is often the case in my clinical experience that major life decisions are initiated on the basis of a diagnosis of AD, including changes in place of residence, property ownership, financial planning and employment. Patients and their families need to anticipate future healthcare needs, changing requirements for mobility

and transportation, as well as safety measures. Individuals who may be more confidently given a diagnosis of AD and those who are reassured by a negative PET scan may both derive direct benefits from a test that will significantly enhance the certainty of the diagnosis.

DR. SOX: Thank you. Any brief questions for Dr. Johnson? Thank you very much, sir. The next speaker, please introduce yourself and --

DR. SMALL: Yes. I'm Dr. Gary Small and you heard about my conflicts earlier and my titles and so forth.

I just want to, I appreciate having another opportunity to come up and make a few other points after hearing the other comments, and I would reiterate that it does, it takes us several years to

.00094

arrive at an accurate diagnosis without the use of PET and by using conventional methods, and this is pretty typical around the country.

I was interested in hearing Dr. Albert's experience at Harvard and Mass General, and my own experience at UCLA and also in talking with leaders around the country, other centers, that this is actually not going on, that most experts I've talked to are not routinely treating MCI with cholinesterase inhibitors. In fact, right now there are clinical trials of cholinesterase inhibitors, comparing them to placebo in this group, so we are still waiting to see what the outcome of those treatments are going to be. So that in my own experience and talking with leaders around the country, that's not the case.

And I showed some of the data in terms of whether in primary care, this is actually being done or recognized and in fact in the real world with the primary care physicians, the dementia patients are not even being treated, let alone the MCI patients being treated and diagnosed. So, I just wanted to make those few points.

DR. SOX: Thank you very much. Follow-up questions? Yes, Sean.

DR. TUNIS: Just to take an opportunity.

.00095

You're actually a member of the practice parameters committee for the American Academy of Neurology guideline that's been mentioned a fair bit today; is that correct?
DR. SMALL: That's correct.
DR. TUNIS: So, I'm just wondering if you can kind of update us. I'm assuming that you all have been in touch with the folks on that committee and the chair of that committee to talk about the relationship of their nonrecommendation for PET in and the evaluation of suspected dementia. I'm just wondering, can you update us on where the Academy is regarding this?
DR. SMALL: Yes, and thanks for asking about that. As it turns out, the Academy at the time that we made on our recommendations, the issue of data that was available was relevant, because the paper that has been referred to several time, Silverman, et al., showing the high sensitivity and specificity of PET, had not been accepted for publication, so that wasn't considered during those deliberations. The committee has not formally met recently.
However, Dr. Jeff Cummings and I, along with another American Academy of Neurology leader,
.00096
Dr. Meader, wrote a letter to the committee making many of the points that I have made today about the importance of PET, the usefulness of PET to assist in the early diagnosis of dementia. So the committee has not formally met but several of us on the committee are extremely supportive of the position that I have made today.
DR. SOX: Thank you very much. Yes, sir.
Please come to the mike, introduce yourself and -- whoops, you have been here before, Dr. Conti, so just say who you are for the benefit of the reporter.
DR. CONTI: Peter Conti again. I have listed my credentials and conflicts earlier. I actually have a question for Dr. Albert and I would like -- perhaps I'm the only one in the room that maybe is misunderstanding the issue or maybe there is a conflicting statement here, but from what I

remember about this type of request of class drugs is that these are in fact investigational, they are off-label uses of unapproved pharmaceuticals. Is that still correct?
And if in fact we're dealing with an investigational drug, and we're classifying PET as an experimental procedure, why is the Alzheimer's Association pushing an investigational drug in an
.00097
experimental procedure they's not pushing? And then why is it that they use PET, or she uses PET, which is still considered investigational, in a certain subset of patients? So, I would like her to clarify that.
DR. SOX: Dr. Albert, would you like to respond and then we will go ahead with break as soon as you're done.
DR. ALBERT: I was trying to make a distinction between the position of the Alzheimer's Association and what we do clinically, and the position of the Alzheimer's Association is that we feel that the standard clinical procedures for evaluating patients is the current way that patients should be evaluated, and that all of the imaging modalities that are currently available are still in an experimental phase with respect to diagnosis. What we do clinically, what I do clinically in my own setting is a separate issue.
DR. SOX: That's fine, thank you. You can have the discussion during the break if you like.
DR. GARBER: Maybe Dr. Albert can respond to the investigational. Are these drugs investigational?
DR. ALBERT: There are currently three
.00098
medications that are on the market for the treatment of AD and memory disorders, and there is a good deal of evidence that people who have substantial cognitive impairment but don't yet meet clinical criteria for dementia have underlying pathology of AD, and that's the reason that some clinicians throughout the country that I know of offer those medications to individuals. Not because we know that

they will definitely be beneficial, but because it's possible that the may be, and there are currently trials underway, as was already mentioned.
DR. SOX: So at this point, has the FDA approved --
DR. ALBERT: It has approved three drugs for the treatment of AD.
DR. SOX: Approved for that purpose.
DR. ALBERT: Absolutely.
DR. SOX: Thank you very much.
DR. HOLOHAN: And to reiterate again, Dr. Albert, you said PET or SPECT.
DR. ALBERT: Yes, I said PET or SPECT.
DR. SOX: Thank you. We will now take a ten-minute break. Don't go too far, please.
(Recess taken from 10:33 to 10:47 a.m.)
DR. SOX: The next step in the process is

.00099

discussion among the committee, which will lead to a motion and a vote. I would like to remind everybody in the room that this, we can ask for people to help us with our discussion but it's not in the ground rules for people to volunteer to help us in our discussion. So if we want you, we will call upon you, but it's basically the business the committee is transacting in public, which is a little different than what has been has happened before.
And so the plan goes something like this, that we will first of all, I'd like to have Barbara and Frank address the issue about the question that was considered, and I will read the original question, the revised question. You can just kind of explain what your thinking was, and I will also ask Sean if he wants to comment on that.
Then I am going to ask Dr. Matchar to comment on any of the issues that were raised by the speakers in the previous session. Then we will discuss any evidence about the interpretation of the evidence, although I want to point out that mostly what we will be doing there is teeing up issues for the coverage group itself to take into account when they make their coverage decision, and only under exceptional circumstances would we remand this issue

.00100
back to the panel on the basis of interpretation of the evidence.
Then I'm going to ask everybody, starting with Randel and going around the room, to make any last comments they would like to make on this issue and then I will ask for a motion and we will vote. So that's the plan. And there will be no lunch until we're done.
(Laughter.)
So, I'd like to read from the minutes of the Diagnostic Imaging Panel. First I will read the original question. A motion was made and seconded to vote on the following question: Is the evidence adequate to demonstrate that PET has clinical benefit in evaluating patients with suspected AD? And the underlying word is suspected.
The question that they actually voted on was exactly the same, except that instead of suspected, it said possible or probable AD, and then afterwards, as defined by the current American Academy of Neurology guidelines. So that's the difference between what was teed up and what you finally voted on.
So, Frank and Barbara, could you tell us what it was that led you to make this change? And I
.00101
think the real issue for us as an executive committee is, was this an error, an important error in process to have made this substitution. That's what we really have to judge, so perhaps you can explain why you did that, and we can form our own judgments.
DR. McNEIL: Hal, I will explain what I think is the non-difference between the two areas, and then the issue of why the change was made that is in the minutes, and maybe Frank will comment on that.
First of all, I'm glad Leslie asked the question and I'm glad that Dr. Silverman clarified what he viewed as the areas of suspected AD, and as I listened to his remarks there were four, four components to that. One was dementia, with or without a documented diagnosis of AD, and these are the words he used. The second was possible or

probable Alzheimer's. The third was MCI, which 50 to 80 percent of the time led to incipient AD. And the other was memory complaints, which 10 percent to X percent, I don't remember the upper bound, went to incipient AD.

Now as I look at those four components and look at the definition of the words possible and probable, I think possible and probable includes all of those things, particularly since suspected AD

.00102

disease is not a disease with a diagnosis of 100 percent, it's either possible or probable. So if it were definite, then I would agree that there might be a mismatch, but there's no such thing as 100 percent positive predicted value of AD from what I gather from reading this literature and hearing the testimony today.

So that, I would say that the difference between the two is nonsubstantive.

DR. SOX: Frank, do you want to comment?

DR. PAPATHEOFANIS: I really don't have anything more to add. I think that covers it.

DR. SOX: So, Barbara has asserted that this change is nonsubstantive. Is there anybody on the panel who would like to challenge her on that point? Because, if not, I don't think we have a quarrel with the process. Anybody want to challenge that? Leslie.

DR. FRANCIS: I don't want to challenge it, but I just want to understand that -- I took it that the commentator was saying well, you just looked at a narrower class, so you know, you looked at the class where it was less likely to be helpful as an adjunct. But you're saying no, we understood the larger class and we made that clear.

.00103

DR. McNEIL: That's exactly right, Leslie, and I think that the words in fact in the four components document that.

DR. SOX: Now I would like Sean then to comment on -- I'm sorry, Bob, I didn't see you.

DR. BROOK: If that's the case, would it be proper to amend the recommendation that you guys

made, maybe just put in a footnote that says this includes suspected, whatever the language was, AD?
DR. PAPATHEOFANIS: Sure, that's reasonable.
DR. BROOK: I mean, because if the committee believed it to be the same and we are voting on the same recommendation, is that within our purview of process, to put in parentheses that this means, this includes suspected AD?
DR. SOX: Alan.
DR. GARBER: Well, I share the sentiment that it's important to document what was meant, but I but I would propose that be reflected in the minutes and it will certainly be in the transcript, what Frank and Barbara just said, so I don't think it's necessary for us to try and go back and change anything the panel said.
DR. SOX: Since it's not a substantive
.00104
change, I agree with Alan. I think that the minutes provide adequate documentation, or will provide adequate documentation of their thinking, and it's almost a legalism.
DR. BROOK: Which minutes?
DR. SOX: The minutes of this meeting.
DR. BROOK: So the minutes of this meeting will include explicit statements that the panel included suspected, the chair people believe or whatever, the panel included a belief that would state firmly that suspected was included as a subcategory of possible or probable.
DR. PAPATHEOFANIS: Yes.
DR. BROOK: Thank you.
DR. SOX: Sean, I would like you now to comment on the second question that you raised about MCI. Can you tell us what your thinking was there?
DR. TUNIS: In hearing this discussion and then also in reviewing the transcript of the diagnostic imaging meeting, we did spend a lot of time on this taxotomy issue of what exactly we were voting on. We had a lot of input at the time from Dr. Silverman as well, and we seem to have done a vote on the possible and probable dementia, using

that phraseology. Then we had another separate vote
on mild cognitive impairment that came out unanimously the same way.
Then there was a supposed hole that was left of something that wasn't addressed that I think that I'm just going to read from the transcript, Dr. Silverman's comment, and see if we can at least get clarification if we're leaving anything out here, because if we are, I know it will come back to trouble us when we try to actually develop our policy on this. And since Dr. Silverman is here, maybe he can help us work through this.
I wouldn't use the words probably Alzheimer's. That has a very specific definition as assigned by, I don't know if it's NIH or AAN, which. American Academy of Neurology. But they actually probably have Alzheimer's is what we would say.
There's also a hole that's being left here if you consider just MCI and just possible and probable AD, because there are many people who have dementia who would qualify by DSM-III or IV criteria as having dementia who still wouldn't have possible Alzheimer's or probable Alzheimer's.
So before going beyond that, maybe Dr. Silverman, if you could help us define, what is the hole that's still being left after we have
covered possible probable Alzheimer's as well as MCI, what else is left?
DR. BROOK: And suspected.
DR. SILVERMAN: Yes. This goes first of all to the error that Dr. McNeil made just a few minutes ago when she said that she thought that the possible and probable would be encompassing to suspected. In fact it was explicitly said and it was added to these words, possible and probable as defined by the American Academy of Neurology current recommendations, and so that explicitly does not include people who have dementia but don't have possible or probable, and it explicitly does not include people who have age associated memory impairment and so forth. It was very clear that it

was meant to apply to this focus group. And then to have an ad hoc say, you know decision that we will go back and say it really applied to everybody when the debate didn't focus on that is I think a very unfair thing to do as well. So the hole that's being left very explicitly in the case of dementia, are people who do have dementia, that is as several people defined it, that they have more than one cognitive domain being affected, that they have evidence of progression that's seen over a

.00107

period of time, and that it involves them to the point that they have functional impairment. It will include some people who will then go on by the American Academy of Neurology recommendations to use the ADRDA criteria to include people who have probable AD, and it will include some people who have possible AD, and it will include some people who have dementia but don't meet the criteria for either possible or probable AD, and those people were never even voted on, and we explicitly asked them to be voted on, and there was a refusal to vote on them at that time.

DR. HOLOHAN: I'm unclear, I'm sorry.

DR. SOX: I am too. Again, what is the hole that wasn't plugged?

DR. SILVERMAN: That people who meet the criteria for dementia as defined by DSM-III or IV criteria, but don't meet the criteria, which are more profound for AD as defined by the ADRDA and NIDCS criteria, which is the criteria that the American Academy of Neurology uses to define possible and probable AD. In other words, there is a whole host of things that can cause dementia, Alzheimer's is just one of them. It's the most common, but there is -- some of the other ones were referred to today.

.00108

Dr. Albert talked about frontal temporal dementia. Those patients wouldn't meet possible or probable AD criteria necessarily, depending on exactly how they present clinically. There is dementia that can be caused by a whole host of disorders that may or may not get detected by other tests that are being done

in order to try to rule them out and if they are detected, at the very least, it will turn a probable into a possible. But depending on what they find, it may even knock it out of the possible and probable category all together. So, AD is one of about 25 different disorders that are known to cause dementia, and those other patients are not being voted on.
DR. TUNIS: I guess the question is, what is the data that's relevant to that patient population that would have been the subject of a vote? What's the information base that people were supposed to be considering for that subgroup which is at least somewhat hard to explain, and I'm wondering if it's been defined enough to actually bring some data to the table. And I don't know if Dr. Matchar has any thoughts about this, because I know Dr. Matchar struggled with these taxotomies as well in doing the model, so maybe he has some way of helping out with this after you.

.00109

DR. SILVERMAN: The data is the great danger that everybody in this room agrees upon, or at least that's been spoken to, and seems to agree upon, of having a patient who actually might have AD, their brain is actually going through deterioration, of cholinergic impairment and so forth, they actually might have AD but they don't meet the criteria for being diagnosed with AD, and so they don't get properly treated --
DR. BROOK: They don't meet the criteria. You have to use these words specifically; otherwise, you're not helping your case. They don't meet the criteria for even possible AD.
DR. SILVERMAN: For even possible AD as defined by the American Academy of Neurology's acceptance --
DR. BROOK: Where is that definition, or what is it? I can't find it in any of the materials we have.
DR. SILVERMAN: That definition appears as a consensus recommendation of the National Institute of --
DR. BROOK: What is it? Can you specify

it? What are the components of it so I can understand precisely the group of people. Because

.00110

they voted on mild cognitive impairment, and they voted on possible Alzheimer's. Now what's between mild cognitive impairment, which they unanimously voted no, and possible Alzheimer's, which they unanimously voted no, what falls in between those two categories?

DR. SILVERMAN: People who have any of the other 25 possible diagnoses that you think are causing their dementia. Maybe Dr. Small can clarify that a little bit.

DR. SOX: Do you want to speak, get to the microphone and try to help us here?

DR. SMALL: Let me try to clarify these points. As I said, dementia in a sense is a severity diagnosis, okay? So you have, before dementia you have mild cognitive impairment. These are people who have --

DR. BROOK: They voted no on that.

DR. SMALL: Well, I'm just trying to clarify the definitions for the committee, okay? So mild cognitive impairment is people with memory loss that may have other cognitive impairments somewhat, but it's not interfering with daily life, so that's not quite dementia. As that progresses, people get dementia. Now if you have dementia, as I pointed out

.00111

in my pie chart, about 55 percent, I think we all agreed on a figure like that, have AD. When we say AD, we mean probable AD as well as possible AD. Now if you include the possibles, maybe it's 60 percent or some confidence interval there, but basically probable and possible AD were clearly defined in 1984 by the so-called MACAN criteria or the NIND/CDS/ADRDA criteria published in Neurology in 1984 where you need impairment in memory, impairment in other cognitive domains, interferes with daily function and it gradually progresses. It is very well laid out. If it's probable AD, it's very clear.

If it's possible, it means that there is

something atypical about the course, there's a secondary illness, there's a bit of a question, but that's quite different from suspected dementia where there you are including, or dementia in general is that other 45 percent, which could be depression, it could be frontal temporal, it could be Lewy body. And then if you're going to suspect a dementia, then you might even include MCI as well. And you know, the other issue I just wanted to mention is that the AAN committee, they looked at those criteria, and also the AAN committee

.00112

is against treat all, clinicians are not treating all, it's not happening, it's against the standard of care. David Matchar, help us out.

DR. MATCHAR: You want me to make comments about this specific point alone?

DR. SOX: This specific issue, and then we need to get through this issue and then we can talk about other issues.

DR. MATCHAR: Well, again, we understood up front that patients who are going to be going through a clinical evaluation of some quality, hopefully guided by the American Academy of Neurology recommendations but perhaps not exactly, were going to be evaluated as having suspected, possible, probable, about 55 percent of those we assumed were going to actually have AD, and that was actually the base case analysis. We acknowledged that about 45 percent of these people are not going to have AD and so the analysis that you all saw and evaluated was based on the notion that if patients were going to be evaluated as well as clinical clinicians currently do in the community, which perhaps is about 55 percent positive predicted value, that it makes sense to go ahead and treat those individuals based on a full

.00113

typical clinical evaluation. Not to treat them simply because they walked in the door and said you know, I have a problem with my thinking, but because you've evaluated them to the level of your ability as a typical clinician based on your reading of the

American Academy of Neurology guidelines.
So that's really it. I mean all of these, the discussion about what does it mean to be possible, probable and so on, we never were able to -- you know, we have those listed as appendices in the technical report, and I frankly don't really understand what subgroup is being discussed here. But I think that we subsume this question under the sense that the baseline analysis of which a large minority of individuals don't have AD.
DR. SOX: What I heard was that there is a spectrum of cognitive impairment, from MCI to dementia.
DR. MATCHAR: Right.
DR. SOX: And once you get into the dementia category, then you could have possible or probable Alzheimer's as the cause, and that sounds like it would cover everything from the zero probability of Alzheimer's to 100 percent probability of Alzheimer's, and then you could have other causes

.00114

of dementia besides AD.
DR. MATCHAR: Well, one specific thing that was assumed a priori was that secondary and particularly reversible causes of dementia were being evaluated, and that is part of the American Academy of Neurology recommendations. What's left over then is things that are not treatable. I don't know if that's answering the question.
DR. SOX: I guess the question that's been raised is have we failed in using the words possible and probable to describe the spectrum of likelihood that the patient has AD? Have we somehow missed some group? Deb, do you want to comment?
DR. ZARIN: Could I try to clarify one thing? Clinically the model started, there was an MCI and an asymptomatic group, but talking about the group with mild dementia, it started with a diagnosis of mild to moderate dementia based on the AAN workup. So once you diagnose someone as being demented, which has to do with cognitive impairment and functional problems, then there is a differential diagnosis of dementia. One of the diagnoses possible is

Alzheimer's. But then it was just said, there are 25, whatever number of possible causes of dementia there are were subsumed in the other 45 percent.

.00115

So the model dealt with this by saying okay, we know you have mild dementia. We are assuming based on sort of a common clinical workup that I think it was 56 percent of you will be ultimately proven or if you could prove it, would have Alzheimer's. So you have that other 44 percent who have all those other, so that was how we dealt with it in the model.

I don't know if that helps, but just think about the diagnosis of dementia separate from then trying to find out the cause of the dementia.

DR. SOX: Bob.

DR. BROOK: I understand what you did and I think I understand what I'm hearing. But we have to figure out some wording here to reconcile this. Let me try to explain it. What I here the clinicians testifying is they want to know whether Frank and Barbara's group explicitly considered those people that meet the diagnosis of dementia, they have more than mild cognitive impairment but the clinician is not willing to say they even have possible Alzheimer's. He is absolutely fully confident they have frontal temporal disease, or depression or something else, Alzheimer's is not even possible by the neurology definition.

.00116

And he wants to do a PET scan for those people, presumably I would guess to find out whether they have Alzheimer's patterns, not to rule out Alzheimer's, and then treat them, and they would not offer therapy for this group where clinically the evaluation at the end of it is, well, clinically the evaluation at the end of their evaluation is not even possible Alzheimer's, so there is no justification for treating those people clinically. But they want to do something else and if the PET scan is positive, they would treat them.

I think that's what I'm hearing. Does that make any sense?

DR. ZARIN: Well, the problems with that is that the treatment studies as I understand it, have used essentially the AAN workup to select people to go into the treatment trial. So they say, let's say, we have 200 people who we have done this workup on, and they meet the criteria for what I'm calling mild dementia, okay? We are putting 100 of them into some standard treatment arm and 100 of them into this cholinesterase inhibitor arm. And on average the cholinesterase inhibitor arm did better than the other arm. We don't know if a different way of selecting those patients would have led to a better

.00117

outcome, so it's a non-answerable question. The argument being that if we could figure out which 45 didn't have Alzheimer's, that they would be better off not getting the treatment is unanswerable, that hasn't been done.

DR. BROOK: I just want to rephrase. I want to know whether the panel explicitly considered those people where the clinician, and we have heard only from tertiary centers, where after a tertiary center referral, the doctor says, Dr. Small says this person has no -- he has dementia but doesn't even meet the possibility of having Alzheimer's, and he would like to do -- and I don't know what proportion of his patients fall into that group, but he has no possibility of having Alzheimer's but he would like to do a PET scan for that group of patients. That's what he said.

You explicitly voted out the impairment before they get to dementia, and you explicitly voted out possible or probable. But I'm asking, that other group of patients, did the committee explicitly make the recommendation about them? And Barbara said yes, and it was refuted on the basis of the fact that they don't fall into this.

So, is this for anyone? Can we get rid of

.00118

this? Is this possible or probable Alzheimer's, or do we mean anyone with suspected dementia? What did the committee decide?

DR. McNEIL: I think this is a question of

wording and that we are really getting caught up in the words that we're using. I think that the committee functionally looked at a combination of scenario A and scenario B, and those two scenarios together encompassed both.

DR. BROOK: So the minutes of this meeting ought to reflect that not only did we look at possible or probable Alzheimer's, but the committee also looked at suspected dementia, not suspected Alzheimer's, and concluded from their analyses that all three of these categories, mild cognitive impairment, selected dementia of any type, and possible or probable Alzheimer's, that PET scan was not worth doing.

DR. McNEIL: I would say that was the intent of the committee.

DR. BROOK: Okay. Can the minutes so show that, of this meeting?

DR. SOX: The minutes will reflect that.

DR. BROOK: And that plugs the hole.

DR. SOX: I think that deals with the

.00119

issue that we have been struggling with for the last 10 or 15 minutes. Okay.

Now I would like to move the discussion on and give David a chance to comment on any other methodologic or interpretive issues that were raised during the public presentations.

DR. MATCHAR: Well, I did just want to thank you, Frank and Barbara, for presenting both the model that we did and obviously the discussion during the previous meeting very succinctly and well, and I'm glad I wasn't responsible for doing that.

But, I think a few points that I just want to reiterate. One is that it really was very important in this analysis that we separated out these levels, the demented patients, the mild cognitive impaired patients, and the patients who are concerned because of a strong family history, and that the focus of the group clearly was on the symptomatic patients because that's the area in which the clinical research has been done regarding treatment. That is, patients who have been in

clinical trials, as Dr. Zarin mentioned, are all patients who are symptomatic, not mild cognitive impairment, although there are ongoing trials, and not patients who have first degree relatives,

.00120

although their lifetime incidents of dementia is extremely high; if people live long enough in that population, they probably will develop AD. So these were very important that we separated them out. The primary conclusion, again for the symptomatic patients, was driven by the point that Frank made earlier which is, if you have a patient with dementia who you might otherwise based on AAN criteria, based on clinical treatments have treated, if you impose a test on those patients and place that as a criteria for whether they receive treatment or not, then some number of those patients will not get treated, and that number is probably on the order of about 10 percent, and I think that the data that was presented earlier, the new abstract based data that was presented earlier, I think also is consistent with that, that about 10 percent of patients who had, who ultimately developed dementia will be PET negative.

So what you have done is guaranteed that a certain number of people are not going to get a treatment that's been shown to be effective. But that's for the symptomatic population.

For the other populations, again, the key element is that we don't know that treatment works in

.00121

that population and I think in more general terms with regard to PET, that there's no evidence even in the dementia population that a PET scan predicts treatment responsiveness, and I think that's really key, because that's part of the logic that I was understanding, that we want to treat people who do not satisfy AAN criteria for dementia because there is a very strongly held feeling that if you had a positive PET scan, you will respond to treatment because those people have AD and treatment works because you have AD. It may indeed by true that the only reason people respond once they are symptomatic

is because they are symptomatic; it may not be that these drugs will work to help people in delaying the onset of dementia.
So, that is the testable hypothesis that is being tested. Certainly, I don't know if indeed it's being tested in the first degree relatives of Alzheimer's patients, but certainly it would be worthwhile doing.
So, I think that there were three things that we don't know that would keep us, that I felt based on the analysis, would keep us from being able to be optimistic and positive about the use of PET scans in the current situation. One is of course, as

.00122

I mentioned, that we don't know whether PET positive patients who don't yet have dementia are in fact treatable, that's unknown. We don't know whether PET predicts treatment responsiveness even among patients who do have dementia, so that's an unknown. And the issue that was raised about whether primary care physicians would do a better job if they used PET scans, and one of the suggestions was perhaps we ought to just forget about sending patients to a neurologist or to the cognitive impairment center, but rather just sending them for a PET scan.
By the way, CMS does pay for these evaluations by neurologists and by Alzheimer's specialists, so you want to use this presumably as a substitute. Well, we don't know that if primary care physicians had a PET scan available, that they would treat any more quickly or they would treat a larger number of people. I would wonder if indeed it was true that a PET scan would lead a primary care doctor to more appropriately treat symptomatic patients who should be getting treatment anyway, is there potentially a less expensive was of doing it than by sending them to receive a PET scan. So in other word, if primary care docs aren't already doing the right thing, can we get them to do the right thing

.00123

short of spending the $2 or $3,000 to do a PET scan.
And I apologize for raising the issue of dollars, because we specifically avoided that, but

it's certainly a resource intensive thing to do. So, that's my comments.

DR. SOX: Okay, questions? Leslie.

DR. FRANCIS: This is actually a question both for the diagnostic imaging panel and for you, and it goes to the treat-all strategy. When I was sent this list of general discussion questions, there is a lot more texture to them than what the panel finally voted on. The general discussion questions ask, are there specific groups of patients, what other issues which haven't been identified, and so on. And as I tried to work out the logic just as a layperson, it seemed to me that your biggest worry was the false negative possibility as against the treat-all strategy, that is, that the downside of adding PET was that some people wouldn't get appropriate care and that roughly assumed sort of a balance between people who get care when they shouldn't and people who don't get care when they should, which rests on the idea that treatment actually is relatively benign.

Now, suppose you had a group of patients

.00124

who had been identified by the standard clinical workup and started treatment on the treat-all strategy. They had been, say, being treated for several months. It isn't yet clear what the effect of the treatment is going to be on cognitive status, but side effects are clear. I mean, there are some folks who do have side effects from these drugs. Would it then make sense to try to do a PET scan, saying oh, look, we're less worried about a false negative here because of the fact that this is a patient who is experiencing side effects. I'm just trying to figure out whether there might have been groups of -- I don't know what other groups of patients the panel looked at, what other considerations the panel raised, because that didn't filter through to me in the summary here.

DR. MATCHAR: Well, this is a clinical question that I think I should defer to those who are involved in clinical care of patients with AD and who use these drugs. But my understanding from our

interviewing experts in the area was that individuals who are having side effects, the side effects we're talking about are the ones that are sufficient to cause someone to want to stop the drug, and that there didn't seem to be sort of a very thin line that

.00125

needed to be discerned between those who should stop and those who shouldn't stop. And indeed, it's something that could be evaluated as a separate question, is there a subpopulation for whom their dysutility or loss of quality of life because of side effects counter balances whatever improvements in quality of life they are going to receive from therapy.

DR. SOX: Bob.

DR. BROOK: David, let me pose another clinical scenario. A large number of Medicare patients don't have drug coverage. This drug costs $1,200 a year. If Medicare covered PET scan, and since you can't tell -- I mean, this drug doesn't improve function, it just prevents it from declining slightly, and the side effects and others. Are there a group of patients who if you really did a utility, would say that hey, I'm willing to, if the PET scan is negative, and the government paid for it, I would not put out of my pocket $1,200 a month, a year, for medication.

And did the panel consider -- this is not explicitly considering costs, but it is a question of whether you believe, you know -- what I'm trying to get at, I mean, I believe that one of the problems of

.00126

this field is that immediately upon approving anything, every person over the age of 50 is going to get a PET scan but the question is -- just like they do with heart scans on everybody -- but the question is, on the other side, are there groups of people from what you know about utility analysis, decision analysis, where you can postulate that there would be people that would see the sensitivity and specificity of this test, and did the panel explicitly consider those subgroups of people that might use it to make a different decision? Why should they put out of their

own pocket $1,200 a year in medication.
DR. MATCHAR: Well, perhaps you ought to turn to Alan Garber about this, but I think that my response would be that you're actually asking us to explicitly incorporate the dollars because you're dealing with an issue of willingness to pay, and there's a question of what somebody considers, and it's an economic issue.
DR. BROOK: Not for the test. I am talking about in the utility function as an outcome.
DR. MATCHAR: What utility are you talking about? Is it the --
DR. BROOK: The patient makes a choice that they would rather buy food than pay for Aricept.

.00127

MS. RICHNER: Did you consider that in your decision analysis?
DR. MATCHAR: Well, I think -- would you be willing to respond to that question?
DR. GARBER: Well, I will tell them what I think your study did. You were not instructed to look at that question, and therefore, you didn't, and I don't think any of us could say what the answer would be to your question, Bob. I think it's a question, it's a reasonable question to ask, but it's sort of irrelevant to our deliberations today, because I don't think whether a patient might want the test for some reason has ever been the basis for us making a decision. So Bob is just saying well, if they have to lay out this money for the drug, would it be maybe, you're asking would it be cost effective to do the test in that context, but that's not rally the way that we judge any of the technologies that come before us and that's not what we asked David's group to.
MS. RICHNER: One of the concerns that I was thinking about, if the person had a positive diagnosis with the PET scan, it was approved, whatever, would they be willing to tolerate the side effects of the drug regime, you know, be a little

.00128

more aggressive and a little more willing to stay on the therapeutic intervention for a while knowing the

trade-off and knowing that they had a confirmed disease.

DR. SOX: Well, if some day there is good solid evidence to that effect, then we can reconsider, but right now there is no evidence other than sort of anecdotal and common sense reasoning, which doesn't -- Deb.

DR. ZARIN: But it also, I think that your question assumes, at least if you were thinking of a rational patient, that the test is a better way of predicting whether you are going to get benefit from the drug or not. In other words, if you knew that the test actually could predict whether you were going to respond to the drug better than the clinical workup, then that seems like a logical question to ask, could it help me decide to save my money, whether it's relevant for this committee to ask or not. But we don't know that the test is better at predicting who is going to respond to the drug or not. That experiment hasn't been done. There is reason to believe that some people without Alzheimer's but with other things in the differential respond to cholinesterase

.00129

inhibitors, so it is not at all clear that the test is a good way to predict who is going to respond.

DR. SOX: Bob.

DR. BROOK: I'm not asking that question. I'm asking a person, I'm not saying it has to be better. I'm saying that I could as a rational person say this neurologist is really a fake and his sensitivity and specificity is not that great, and I want both. If I'm going to put out the $1,200 a year for this drug, I want both a combination of a good clinical workup that says I have Alzheimer's, and a PET scan that says I have Alzheimer's, because either one of them has error in it, and that's what I want. And I just want to make sure that, I mean, it sounds like these kinds of subgroups were not included in -- I mean, that's really a very rational thing to do. So it's like saying that if I do an angiography and I know that one reader has a 70 percent reliability, I want two or three readers, and

if I put those two or three readers together, I may increase my specificity and I may miss a case or so, but I'm willing to do that, I'm willing to make that trade-off before I get my chest cracked. So that's what I'm asking you, David, are there cases where these things would both be positive where some

.00130

rational person would say if I have to pay for it, that I would be only willing to pay for it under those circumstances?

DR. MATCHAR: I think that what you are describing is a circumstance that is entirely, it is a utility or a psychological perception, it's something that somebody decides in their mind that if they were to receive something that they perceive to be objective as opposed to being something that --

DR. BROOK: Isn't it because --

DR. MATCHAR: (Inaudible, multiple speakers.

DR. BROOK: Let me just ask the question this way. Isn't it the case that if I'm positive on a good clinical exam and positive on an independent PET scan, that the probability that I really have AD is higher than if I'm positive on either one of them? Isn't that known?

DR. MATCHAR: No.

DR. SOX: I just want to say, we really need to wrap up this conversation and get on to other people, so let's try to answer and then let's move on.

DR. MATCHAR: Well, just to answer your question, the simplest part of your question which

.00131

is, did we include this explicitly, and the answer is no.

DR. SOX: Okay. Barbara.

DR. McNEIL: Well, I guess one of the things as I was thinking about this, Bob and others, all of the issues that have been raised by Randel and by Bob, really in the simplest form boil down to some kind of utility. I mean, that's really what they're coming down to, however we could integrate all of these factors, they come down into some kind of

utility assessment. And the issue there is, is there any way that we could believe that all of these would get to a utility so low that it would flip people into the test strategy rather than in the treat strategy. I mean, that's really, if you boil it all down, that's what it comes to.

So what I did actually in preparation for this meeting is review the article by Chapman and Neumann, which many of you may have seen, that summarizes the utility literature to date across consumers and caregivers and the lay public. And in fact it's very hard to get very low. I mean, if you have death and you've got total paralysis and stuff like that, you can get down to the zeros, but most of the things that we're talking about don't really hit

.00132

that low. So when I read this, and I actually requested the article so I'd be sure I was up to date on these utilities, I think it would be hard, I really think it would be hard taking into Randel's concern and your concern that we pull the utility down so that we cross that threshold line. And we're never going to know, because we don't have those data specifically for Alzheimer's patients. The data that I'm quoting are for a whole slew of patients, so if we wanted to make these utilities Alzheimer's specific, we'd have to go out and do several really quite large studies.

DR. SOX: Wade, you have been waiting to ask a question, so go ahead.

DR. AUBRY: My question related to Leslie's, and that was the clinical situation, and it was just sort of a comment Leslie had said, if the patient had been treated for several months and was having side effects. And I think speaking as a clinician, generally you would know within a shorter period of time whether there would be some clinical improvement based on the drug, and then there would be a clinical decision about whether the side effects outweigh the benefit. And so, that sort of gets back to the usefulness of the clinical evaluation.

.00133

So, I guess the question in regard to PET

would be, you know, if you make a decision to stop the drug, would a PET scan be useful in that patient? You have a patient who's had a clinical improvement, and that is really sort of beyond the scope of what we're doing here.

DR. SOX: Yeah, beyond the scope, no evidence. Other issues that people would like to raise for David?

DR. TUNIS: David, I think it was Dr. Small, I believe, maybe Dr. Silverman today, talked about and actually some of their slides showed some of the studies that suggested, you know, delay in nursing home placement or other outcomes associated with at least treating, I'm not sure about testing, where there was potential improvements associated with an accurate diagnosis and treatment. And I don't think those kind of outcomes were included in your model. So for example, better caregiver arrangements, delayed admission to nursing home, fewer repeat visits for evaluation and testing, those sorts of things, and I'm just wondering, was that a body of literature that you looked at in terms of preparing the model or if not, why not, or does the model in any way address those sorts of outcomes?

.00134

DR. MATCHAR: I know of no literature and we didn't identify any literature that demonstrated that PET scan results in some way allowed people to make better plans for their lives beyond what could be done by having a patient evaluated by a neurologist expert in this field. But there is evidence that that kind of advice is very useful, but not that PET scan will influence the use of those interventions.

DR. SOX: What I would like to do now by way of focusing the discussion or perhaps finding that we're ready to vote is to have each person kind of state what their thinking is on this issue, where they are, recognizing again, and just to remind everybody that our job is to decide whether the panel followed good process, interpreted the evidence that was in front of them in a reasonable way.

And so, Randel, would you start please?

MS. RICHNER: I have my five seconds in the sun here, but I just wanted to say that I think what concerns me more than anything else, and Bob briefly touched on that, is the emphasis on whether an adjunctive diagnostic test, and this is very generic in a sense, must lead to a change in treatment pattern or a change in health outcome, and

.00135

it's a fundamental concern. I think that if all tests had to go, diagnostic tests had to go through this level of scrutiny, the radiology labs would be pretty quiet.

So, I think that this is very fundamental here in terms of how we're evaluating this but given that, I know we have to get back to our question here, what we're evaluating, PET. Considering what the question was, and that's also a concern in the process, how we're evaluating the question, how we're deciding the question, what evidence we're going to need for that question, I would say that I have to agree that at this point there is little evidence that supports there would be a change in treatment associated with this.

DR. SOX: Frank, you can second guess your own panel if you want.

DR. PAPATHEOFANIS: I guess I'm not going to comment on our process. I think you know what the committee decided and I think I defined that.

DR. SOX: Wade.

DR. AUBRY: I think that the question that was framed was appropriate, and I think the concerns that were raised by the speakers about the actual question that the panel reviewed was adequately

.00136

answered, so I feel comfortable with that. I also feel that the approach used by the assessor from Duke was sufficient and more compelling than the evidence that was presented today during the presentation. So, I feel that in terms of addressing the questions and in terms of the process of the imaging panel, that I would be in favor of upholding their vote. I also think it's significant, I don't think my decision hinges on this, but I do think it's

significant that the official statement of the AAN does not recommend PET in the routine evaluation of these patients, and that the official position of the Alzheimer's Association is that this is an experimental procedure. Again, that's not the major determining factor in my view, but I think it is a significant piece of additional information.

DR. SOX: Leslie.

DR. FRANCIS: Since we're advisory, I actually have a question that I just want to be sure that I've got the answer from the folks from the Diagnostic Imaging Panel, and that's their answer to question two of the CMS questions. Were there any other issues not addressed in the model that you all thought in the panel discussions might influence the decision to use PET that we haven't heard about? I

.00137

mean, have we gotten out on the table everything we need to know in answer to that question from your deliberations for CMS to now listen to?

DR. PAPATHEOFANIS: Right. I think the questions that you're referring to, the genesis of those are through the Agency, and they were intended to stimulate thought when folks received the package, and were really I think very useful in that context. But to answer your question specifically, no, I think everything was on the table and the minutes reflect that.

DR. McNEIL: I would agree, Leslie. I think in particular you might be talking to some of the psychosocial of legal issues --

DR. FRANCIS: Well, yeah.

DR. McNEIL: -- that were addressed specifically in Table 10, and were specifically commented on during the course of the deliberations.

DR. SOX: Bob.

DR. BROOK: In general, I agree, yes. I would reiterate a position that I have stated many times before. I am sure that a well done up-front sort of appropriateness analysis where we really looked at subgroups would find subgroups where decision analysis would support using a PET scan in

.00138

these groups of people with the current evidence that's available, and I think we drew out a few of those cases. I think it's probably in this case small, but I hope that the process in the future could become more sophisticated in identifying these subgroups explicitly and explicitly modeling or using judgment to figure out what to do. But given where we started from, I think the subcommittee did a commendable job.

DR. SOX: And if there were such subgroups discovered and published in the peer reviewed literature, then the coverage group could modify their coverage policy.

DR. BROOK: I disagree. We've had this disagreement before and I won't let that go. I believe it's important that this panel and this process take the best expert clinicians and use state of the art technology up front to identify explicitly the clinical subgroups that people are interested in, with explicit definitions, and the vague kind of questions that are given to the panel don't do that, and the methodology for doing this has advanced for the last 15 years. It's just like you used decision analysis methodology here; that methodology ought to be used in this process up front. I think until we

.00139

do that, good clinicians will not necessarily believe the kinds of results that we produce. They can be combined.

DR. SOX: John.

DR. FERGUSON: Well, I suppose until we have trials of diagnostic studies that have outcomes and are randomized and so on, that we will fall back on modeling, and I think that this model that the Duke people came up was quite novel, at least in my view. I was impressed by it. What hasn't been discussed here so much but is implied in their Table 10 of the Matchar report was, the specificity in the PET scan studies comes out to be around 75 percent. I realize that that's supposed to be better than the usual clinical thing, but the patients that I see are terrified of having a diagnosis of AD most of the time, and their families are, so I think that a false

positivity of 25 percent is not insignificant, and I wonder how the people who use PET scans a lot handle that, but that's sort of an aside.
I think the panel did a good job in using modeling to come up with a reasonable answer.
DR. SOX: Thank you, John. Barbara.
DR. McNEIL: It would be hard for me to say the committee didn't do its job, but I would say
.00140
actually that we did from several perspectives, and just integrate some of the comments that have been made so far.
First of all, I think we responded to a level of detail in the request that was made by those individuals requesting coverage so that we didn't get down to really tiny subgroups, we were responding in general to a general request. And had the formulation of the questions discussed at a previous meeting before a subcontractor was even hired for this process is my understanding, and then we asked that subcontractor to answer questions that had been publicly vetted using criteria that the MCAC had developed over the past year and a half. Right or wrong, those are the criteria and is what we had to go by.
And the subcontractor in my view, you did a terrific analysis that was really state of the art in all possible ways of decision analysis, and integrated the literature in a very sophisticated fashion and therefore, came up with a conclusion that I think was robust, given the givens. I think what the issue so far that has been raised is, wouldn't it be nice if we were able to get some more specific subgroups or some more specific questions. And maybe
.00141
the thing that the audience can do is say ah hah, there are some very specific issues that were raised and those issue are the ones that we should go research and get answers to, and then perhaps feed back in the form of another kind of request. The one that comes to mind was the predictability of PET in assessing the response to the cholinesterase inhibitors. So it would be that kind of thing that I

think an analysis of this may be helpful for, despite it's nonenthusiastic acceptance I suspect, by members of the audience.
DR. SOX: Daisy.
DR. ALFORD-SMITH: I too would be in support of the committee's recommendations and would also like to commend them for a job well done, I think it was great.
My decision was based upon not only the analysis that was provided by the subcontractors and how the panel took that into consideration, but I also found it interesting, particularly in reference to what Dr. Albert said, and she had the ability to differentiate what she was doing in a clinical setting but yet stepped forward in terms of making a recommendation on behalf of the Alzheimer's Association. And then lastly, my conclusion was
.00142
based on how the subcommittee made the decision to be inclusive in terms of various definitions without leaving that perceived gap within that whole process and so based upon that, I support it.
DR. SOX: Alan.
DR. GARBER: I agree with everything that's been said, but I especially want to echo one of Barbara's statements. We have a number of very distinguished researchers in this area in the audience today, and I was struck by how useful it would be to have a study that directly addressed the differences between outcomes when care is managed by the standard AAN approach and adding PET, and I think that if we had good studies of that kind, and you could even look at treatment of individuals in whom the standard workup and PET were discordant, it doesn't have to be a huge study, but I think it would be enormously helpful in trying to determine the role of PET, and we just don't have that yet.
But I think the committee clearly did their job, and I didn't take the panel's determinations or any other work to say that this test does not have a future for Alzheimer's disease, it's just given what we know now about the tests and about the treatments, it doesn't meet the criteria

.00143
that we've set up.
DR. SOX: Tom.
DR. HOLOHAN: I generally agree with Alan.
I would simply add one comment about the issue raised of subcategories of patients who may differentially benefit. We have seen this in this committee and in the subpanels in issues where it's easier to make that distinction, and that's in treatment. My own panel looked at litologous stem cell transplants in high dose chemotherapy for multiple myeloma, and then later at levo-carnitine use in end-stage renal disease patients. And when you have varying and usually more or less small response rates, it's easy post hoc to speculate that there may be some subcategory of patient that may benefit but the fact is that there is no evidence available to allow you to take that beyond the level of speculation, and I think that's kind of what Alan was addressing about a prospective study.
To force CMS to consider the possibility, unproven, that there are subcategories that may or may not respond when in fact the evidence is not available that would allow you to make that distinction, I think is totalogy.
MS. BERGTHOLD: I just want to make a
.00144
comment about the Alzheimer's Association's thoughtful response to the panel and to note that I believe this is the first time that any advocacy organization has actually, has actually not advocated for the coverage of the device or the technology, and I thought that was an interesting and balanced response on their part. I myself have encouraged consumer organizations to look at, you know, to really look at the whole spectrum of possibilities when they are coming before us, to not just advocate for coverage of something, but to look at sort of who it would harm, who it would help, and really on balance what the outcomes would be. So, I would just like to note that for the record.
DR. SOX: If I get to vote, I will vote to ratify. I think the committee followed the process

very nicely and interpreted the evidence that they got in an appropriate fashion. One of the advantages of decision analysis is that it tells you what the variables are that are likely to really make a difference in making their decision, and I think we have tended to focus the discussion on the key variables in this particular problem, and hopefully as we use decision analysis in the future, we can do so in the same way.

.00145

I guess the other comment is that we are also sticking pretty much to the rule of evidence here and we have been pretty careful about ruling out of bounds questions such as testing in patients who have a clinical condition for which there is no proven therapy, and in Alan's calling for a randomized controlled trial of management with or without the test, I think we are trying to push the envelope on the evidence and we will probably be back in a couple years with this problem.

So with that -- do I turn it over to you at this point?

MS. ANDERSON: I have a brief comment.

DR. SOX: Before we vote, I would like to give anybody a chance in the audience, a chance to step up to the microphone and for a maximum of two minutes say anything that they would like to influence the outcome, but two minutes maximum. Please identify yourself, state any conflicts.

DR. PHELPS: Mike Phelps, UCLA. I have some ownership in a company, CPI. I do not receive money from anybody other than that so I have no other conflict.

There are two comments I would like to make. One is that dementia is not a disease, it is a

.00146

clinical syndrome of symptoms or signs. The issue here is that both for good clinical management and use of treatment and by the FDA label, in the mild to moderate stages you have to make the differential diagnosis, and PET is the most accurate way to do that. We keep ignoring the fact that clinical diagnosis goes on for years, and it has a lower

accuracy than PET at the start of that, when they are actually at that stage. The clinical diagnosis accuracy comes after they have progressed to later stages.

The other comment I would like to make is just a general comment that if you treat everyone and in a broader category at the earlier times, aren't you telling all of them that they have AD? What is the implication of that? We said at the early stage, about 30 to as high as maybe 50 percent of the people have things other than AD. The good molecular therapy, pharmacologic therapy, is to identify patients that have cholinesterase deficits when you give them a cholinesterase inhibitor, and that's the liability and responsibility that one has to step forward with, and PET is a great aid to the physician, not to replace them but to assist them in making that at a very early mild to moderate stage as

.00147

opposed to just saying well, we're going to give you the drug and then the implication of that, that I don't know yet that you have a cholin deficit, and by giving you the drug that I think you have Alzheimer's.

MS. ANDERSON: For today's panel meeting, voting members present are Daisy Alford-Smith, Wade Aubry, Robert Brook, John Ferguson, Leslie Francis, Alan Garber, Barbara McNeil, Frank Papatheofanis, and Tom Holohan. Chairperson Hal Sox will vote in the event of a tie. A quorum is present, no one has been recused because of conflicts of interest. At this time the chairperson, Dr. Hal Sox will call for a motion and will ask the voting members to vote.

DR. SOX: Would anybody like to make a motion that we can act on? Tom.

DR. HOLOHAN: I move that the executive committee accept the recommendations and conclusions of the Diagnostic Imaging Panel.

DR. ALFORD-SMITH: Second.

DR. SOX: We have a motion and a second. Any further discussion before we vote?

MS. ANDERSON: I'm going to read the entire motion. The motion is to accept the

recommendations of the Diagnostic Imaging Panel
.00148
regarding the use of positron emission tomography (PET) for the diagnosis and patient management of Alzheimer's disease and other dementias. Those voting members voting to agree, or for.
DR. FRANCIS: I think I need to state for the record since I came late, that I have no conflicts.
MS. ANDERSON: Those against? And no one has been recused and no one is abstaining, so it is unanimous for.
DR. SOX: At this point I would like to declare a recess. We'll reconvene at five minutes to one.
(Recess from 11:55 a.m. to 1:05 p.m.)
DR. SOX: I would like to call the committee to order please. This afternoon we're going to have I guess what's considered to be an educational program on the role of decision analysis in coverage, our advice about coverage decisions. Some of you might think that we're going about this a little backwards in having a relatively entry level discussion of a topic that we have explored in such a sophisticated way this morning with David Matchar's talk. The purpose of this session, at least the purpose or my talk, which is
.00149
going to be an introduction, is really for people to kind of understand at a gut level the rationale behind using decision analysis in coverage decisions. And if you're like me, sometimes you'll here a presentation that sounds extremely logical but you really don't understand what's going on in the engine room, and you sort of nod and it sounds logical, and you accept it. So what I'm going to try to do is talk a little bit about what's going on in the engine room, and then Dave is going to take the discussion up several levels, and hopefully when we're done with this, we will have a much better understanding of how decision analysis can help us in the many instances in which we don't have high quality data.

So, if you look at our interim recommendations for operating policy, you'll see that we're told to pay attention to the validity of the evidence, and we all recognize that the ideal is to have several randomized trials that all point in the same direction, but frequently we are dealing either with discordant trials or with observational studies in which the assembly of the cohorts that you're comparing may be made more difficult by selection bias or by confounders that are influencing both the

.00150

tendency to get the intervention as well as the tendency to have the desired outcome. We're also told to pay attention to the applicability of the evidence to Medicare patients and finally, we're told to pay attention to the magnitude of the effect of the candidate for coverage relative to already covered options.

Now, I posit that the role of decision analysis in coverage decisions boils down to something we heard implicitly in discussion this morning. If high quality trials aren't available, decision decision analysis is the next best option for measuring effect size, for evaluating the role of bias and confounders, and extrapolating the evidence to other populations. So that's sort of what I'm asserting. Just a reminder. Confounder is an influence or factor that both influences the tendency to get the intervention on the left, and also independently influences the outcome so that the intervention and the outcome are moving in the same direction, and you are tempted to make the inference that it's the intervention that's affecting the outcome, when in fact it's the confounder that's moving both in a coordinated fashion.

Now, the basic principle of effective

.00151

expected value decision making is a sort of principle of utilitarian philosophy, which is always choose the decision option that has the highest expected outcome, and by expected outcome we really mean the highest outcome on average if you were faced with that decision many many times, you would pick the

outcome that would turn out the best in the long run. A decision tree is a method for assessing the balance across arms and benefits, basically a way of modeling a management strategy.

Now I have been fond of using a metaphor for expected value decision making that makes an analogy with gambling. One play at the slot machine, an individual gambler has no way of knowing how things are going to turn out. But a year of play at the slot machine, which is what the casino owner has to deal with, you know very precisely what's going to happen, because there will be tens of thousands of repetitions and under those circumstances the play of chance is actually very predictable.

Now applying this metaphor to the doctor who either has one patient for which he has to make a decision or on the other hand, the other perspective, namely that of casino owner, is a lifetime of practice, and what you should be doing if you're an

.00152

expected value decision maker as a physician is to choose or encourage the option that leads to the greatest gain in the long run, that is to say, the option that over a lifetime of practice would provide more benefits to the patient than the alternative.

Now calculating expected value is a pretty straightforward process. Here we illustrate a game of chance in which you have the option, if you choose to play the game, there's a 10 percent chance of winning $100, a 30 percent chance at $10, and a 60 percent chance of having to pay $60, and the question is should you play. The symbolism here, the circle represents a chance node at which the outcome is driven by chance. If there were a square, it would indicate a decision node, at which the decision is either yours to make or the decision would be to choose the option with the highest expected outcome.

So the way that you calculate the expected value at this chance node is to multiply the probability of each outcome times the magnitude of the outcome, so here we have .1 times 100, plus .3 times 10, plus .6 times -60, and it turns out narrowly if you play this game many many times you

will come out about $4 ahead. So, when David Matchar pushes the button on the computer to do a
.00153
calculation, it's doing many many calculations of this type.
DR. BROOK: I think you did your math incorrectly.
DR. SOX: I could go backwards to dispute this but if I did, we'd be in big trouble. Okay. Did I make a mistake?
DR. BROOK: Yeah, you set it up so .3 would make it worth playing the game, but that's all right.
DR. SOX: Well, if that's the worst outcome that happens in this talk, I will be okay. I knew I'd do something like that, and I knew that you would say something about it, Bob.
(Laughter.)
DR. SOX: So now let's take kind of a hoked up clinical example and walk through this hoked up clinical example. This is a man who has a slow growing malignancy surrounding his aorta, which is a bad place to have a malignancy. He's been told that he has two years to live if nothing else happens, and the question is should he go for curative surgery or palliative surgery, or optionally, simply accept the two-year prognosis. Now the way I've set this up, radical surgery looks pretty bad. There's only a 50
.00154
percent chance of surviving this operation, which is one that's being done in a very delicate part of the body, and there's only about a 50 percent chance of cure if he survives the operation. Whereas, palliative surgery has no chance of a cure, but improves somewhat, but has a very good chance of surviving the operation. So the question is, which option should he prefer?
Now here we represent a model of the decision and the square represents the decision, which is try for cure or palliate. And if you try for cure, there is a substantial chance of operative death represented by the first chance node, and then as you go to the right, there is the long-term

outcome of the operation which is whether there is a cure or no cure. And the tree has the same structure on the bottom, but as you will see, the numbers will be different.

Now the factors in a decision tree are pretty straightforward. There is the probability of the chance outcomes and then there is the outcomes themselves, and the key thing of course is you want to use the same measure or same unit throughout the tree. One outcome you could use is life expectancy, which is the average length of life. Alternatively

.00155

you could use utility, which is a number, which is a measure of preference for an outcome, usually expressed on a zero to one scale, or you could use quality adjusted life expectancy, in which you multiply life expectancy times utility, which in effect converts years in an unhealthy state if it was life expectancy only, into years in a healthy state. And finally, you could use the word that we don't use much around here, which is cost.

So here I put the probabilities on the tree, and here I put the outcomes on the tree. The life expectancy of 20 years if there is a cure, life expectancy of two years if effectively the operation doesn't alter the prognosis.

Now this slide shows how you would actually average out at a chance node and then fold back the tree to calculate the expected value of the two decision options. So in the upper right chance node, you would multiply the probability of cure, 0.5, times the life expectancy, 20 years, and add that to the probability of no cure times the life expectancy of two years; that gives you 11 years, which you then multiply times 0.5, and then add the 0.5 times life expectancy with operative death, which is zero, and that gives you 5.5 years on the average

.00156

that you would experience if you tried for the cure. In the bottom part of the tree you can see that the average length of life would be about two years.

So, although this doesn't look like a particularly good gamble for the patient to take to

try for a cure, in fact it's the best of a bad situation, according to this analysis.
Now, one of the most important parts of decision analysis is sensitivity analysis, which is a method you can use to decide if one of these factors, either the probabilities of the outcome measures could alter a decision. And the basic approach in sensitivity analysis is to first establish a range of reasonable values for the factor such as probability or outcome measured, and then you substitute the lowest value in the range of reasonable values into your decision model, and calculate the expected value. Then you repeat the calculation but now substituting the highest value in the range and you ask yourself, does the preferred decision option change as you go from the lowest to the highest value in the range. If it doesn't, then you can conclude that the decision is not sensitive to that factor.
So that's the basic principle behind sensitivity analysis, and Dave showed us a very
.00157
elegant example of two-way sensitivity analysis in the presentation this morning.
Now sensitivity, I think can help coverage policy a lot. For example, if you think that there might be selection bias, then you could postulate an effect of selection bias that could change the preferred option. Or if you think there's an unmeasured confounder that could be making it difficult to have an unambiguous interpretation of an observational study, you can postulate an unmeasured confounder and its effect, and then ask for either one of these, how large of an effect of these two factors would be needed in order to change which option is preferred, and if that effect is outlandishly large, then you could conclude with at least a moderate degree of confidence that those biases and confounders would not be important.
It can also help in trying to apply findings in one population to an older Medicare population. For example, you could model the effect of older age on life expectancy as an outcome measure. An older person has a shorter time to live

and has other diseases that are likely to prove fatal while waiting for all the benefits of an intervention to play out in the form of greater life expectancy,
.00158
and you can model that very nicely with decision analysis. You could also postulate an effect of age on the effectiveness of the candidate technology and again, ask yourself, how do these changes in the factors in the model alter the effect of the candidate technology.
So, I'll finish up with just a couple slides on cost effectiveness analysis, which conceivably could be in our future sometime later on. Alan Garber has described cost effectiveness analysis as a method designed to assess the comparative impact of expenditures on different health outcomes, so it's a way of comparing several different potential interventions and trying to decide basically, are you going to get your money's worth from an intervention that's more expensive.
The basic approach is comparative. You compare the cost of the proposed intervention with the cost of the currently accepted technology, and you divide that by the difference in quality adjusted life years with the candidate technology as compared with the existing technology. Now cost effectiveness analysis, therefore, measures the impact of a candidate technology on cost and outcomes. It's always comparative and it measures an effect relative
.00159
to the status quo. The CEA yard stick that's in common use is that less than $50,000 per additional quality adjusted life year is considered to be reasonable, 50 to 100,000 is considered to be a gray area, and more than $100,000 is expected to be an expensive policy to apply over a large population. Just where those numbers come from is something that Alan is going to explain if anybody is interested.
Finally, you can do decision analysis of testing, and that's really what David did in his analysis, and here's an example of a choice between testing, treating everybody, or observing and neither testing nor treating. And I have just modeled to

show that if you choose the test option, the test could be positive or it could be negative; clearly, that's a matter of chance. And if the test is negative, the patient could either have the disease or not have the disease, and we could have that same disease no disease chance option in any of those branches that don't have anything attached to them, and then model the effect of changing the probability of disease and see which of the three options has the highest expected value for different probabilities of disease, and define the zones of probability in which each are preferred, as David inferred earlier.

.00160

So, the role of decision analysis in coverage decisions is mainly when the only evidence is from observational studies and you need to try to stitch together evidence from different observational studies and then try to model the effect of confounders, model the effect of applicability to Medicare patients, and of selection biases, and when you want to assess the potential impact of unmeasured effects.

So, that concludes my brief presentation, hopefully a look inside the engine room for those of you who don't know this field and for those of you who do know the field, an opportunity to skewer your chair with a great question. Nobody is going to take the opportunity to skewer me. Sean.

DR. TUNIS: I will try to skewer the chair. The question is, in the example that you gave of the malignancy around the aorta, it obviously raises the issue that I think kind of underlines some of the discomfort with the decision model applied to the PET Alzheimer's case too, which is the lack of -- well, in that case, the lack of patient preferences as part of the calculation. So for example, I may be interested as an individual in a 98 percent chance of living two years rather than taking a 50-50 chance of

.00161

being dead right after the operation or during the operation, I guess.

And so, which raises I guess some issues around coverage policy, because the preferred

strategy in aggregate there was the five years to do the surgery versus the 2.99, but for an individual, it might be appropriate based on their preferences actually to not do the surgery, and I'm just wondering if you could comment about that in terms of addressing it through modeling and policy.

DR. SOX: Well, one way to -- you know, a policy, a decision to cover on the basis of the effects over a large number of patients kind of opens the doors for patients and clinicians to do it if it's clinically important. But for an individual patient, you would like to model their personal preferences which might, and one person might have very different preferences than another. Now some of the time the preferences won't make any difference, and that's one of the nifty things about sensitivity analysis, you can assume a utility of 1 for an outcome or a utility of .5, and if the preferred option is the same over that range of utilities, then you ought to be really coming on pretty strong in recommending that to the patient, because it's

.00162

unlikely that their preferences should affect their decision making.

On the other hand, if it's a close call and the preferred option changes as you go from a utility of .9 to 1, then you ought to pay a lot of attention to what the patient is telling you in making your recommendation. But I think the difference between a policy which opens the door and individual decision making is really very nicely encompassed by a decision model of this type where to develop the policy you might use utilities that have been obtained by doing a study of 500 patients and getting the utilities, but when you apply it in patient care, you use the patient's utilities.

Any other questions for me? John.

DR. FERGUSON: Did you mean to suggest that soon we will be taking up cost effectiveness analysis for our debates?

DR. SOX: I don't know of any plans for that.

DR. TUNIS: I don't think he was

suggesting that.
DR. SOX: I certainly wasn't suggesting that, but my guess is that some day we're going to get around to using cost in our decisions, in which
.00163
case we will need to go over this again.
DR. FERGUSON: Thank you.
DR. SOX: Well, if there are no further questions, then I will turn it over to Dave Matchar.
DR. BROOK: Hal, can I ask you one question? Is anyone ever going to put together the science of the reliability of the decision analysis process? If you give the same problem to two different teams, how consistent they approach it, I would just like to point out for the panel, that kind of work has not been done. And just because it's quantitative and looks simple, it's not; it requires skill an delicacy to do these things and I don't know of any literature in which the same problem has been subjected starting from scratch, by two decision analysis teams, and the way they've gone through the literature, come up with the sensitivity, what they have done, how they have modeled it out and actually produced comparable results.
And the reason I bring this up is because all of the quantitative types on this panel bemoan the fact that expert clinical judgment and appropriateness and all that process which has a known liability to it at least, because those tests have been demanded, there are no tests that I know of
.00164
the other form of work that's been done, and I would urge that this field -- you know, we may want to adopt this technique and I have nothing against it, but I would urge that the reliability of this process done independently by different teams of different merit and skill should be assessed.
DR. SOX: I actually know of one example, but it's unpublished, under review, and actually in that one at least, there are pretty consistent results in that particular one. Of course, it's not independent, because one came before the other.
DR. BROOK: I would make one other case,

that it's very hard sometimes, just like you see in the material, it's really hard to go through all of the work that has been done, especially when this comes out as a computer program, and there are mistakes that have been made. When we have looked at some of the modeling, for instance, of the decisions about using coronary angiography in bypass surgery, there were things that didn't make any sense, like there was no immediate death rate from surgery following bypass surgery, and we couldn't figure it out. And we go back, and the database they used happened to have no deaths over the first hundred patients. So yes, there are sensitivity, there are

.00165

analyses, but when you start getting complex things here, it does require devotion to the kinds of stuff that you've done in your career, which is really meticulous paying attention to detail and reading through what the assumptions of the decision analyst is who's doing this.

DR. SOX: Yeah, reviewing a really complex decision model can be virtually an impossible task, and I have seen decision models, one by Alan and one by Steve Powker on a very complex issue where you basically had to say this person has good reputation, and that's it. One think that would be nice would be if we could get to the point where somebody who is submitting a decision analysis for publication would send the model, and it would give a few brave souls an opportunity as reviewers to play with the model to see how well it behaves at some of the extreme ends of the ranges of assumed value. Barbara.

DR. McNEIL: This is sort of an example, an answer to your question, Bob. When we had the Diagnostic Imaging Panel in January, Peter Neumann was one of the advisors to that panel and he told us, and Frank, correct me if I'm wrong, that he and his colleagues were actually developing a similar model and when he looked at this one closely, he found out

.00166

that both his model and the Duke model had essentially the same structure, and differed ever so slightly on that relative risk. Duke was .72 and he

was .70, or something, and so at least for that situation --

DR. BROOK: And I didn't point out the calculation problem just to be a pain in the ass. We did raise -- we always run out of money on these things and when we did the Rand -- you know, we spent $100 million doing the Rand health insurance experiment, which was the most major health services randomized trial ever done, and there was no money ever to double independently code tens of thousands of line items that we coded when we reversed all these variables, did all the regressions, selected the variables and did all this stuff. They were never double coded; they were double key punched, they were double checked, but never double independently coded, and Joe Nuast to this day has nightmares about if somebody would say, I want to go back, take exactly the raw data set you had, and independently, using exactly the same logic, translate that into computer code and then see if you get the same results.

We tend to run out of money at that time,

.00167

and one of the things that we will need to discuss as we do this is, you know, after you specify the model and do all this stuff and you turn it over to a computer analyst or a software package to do it, I mean, simple mistakes produce errors, and unless you go through them, you can't find them. And as you say, going through some of these models, nobody really ever does in detail. I just think it's stuff that we have to be aware of.

DR. SOX: We will turn things over to Dave Matchar, who is going to take things up several notches.

DR. MATCHAR: Well, my job is to assuage all your concerns, obviously. This is a very daunting task. The discussion here is going to be about your bringing what I call quantitative modeling, so I'm using a term more broad than decision modeling. Quantitative modeling to the processes that you are engaged in in developing recommendations for CMS.

Again, I'm using a more general term, clinical policy formation. I feel that I'm bringing coals to New Castle. This is a group where many of you have been working in this area for years and have been concerned about the question of when and whether

.00168

to use quantitative modeling, but you asked me to talk and so I will.

So I will just give you some of my general observations based on not only this work, and I will try to allude to the work on PET scanning where I think that it relates to some of the questions that you just raised, and other questions as well. But also, I'm going to be basing these comments on experience I've had and interactions I've had with other people who have worked in the area of using quantitative assessment, quantitative modeling in policy formation both in and outside of the clinical or medical realm.

Steve Powker being one person who I've spent some time talking with some years ago. He was involved in working with the NIH consensus panels, trying to consider ways in which consensus development might be improved through the use of decision modeling. I don't believe that it ultimately got incorporated in the long term, but there were some exercises that were published out of that. And also, I have some experience working with someone in a business context, namely Larry Phillips at the London School of Economics, and also Dr. Bill Asher, who is in public policy, and much of the work

.00169

that he did is in the area of general public policy, and the issues he's been concerned with has been how, what have been the successes and failures in the application of quantitative analysis in policy formation, and I think those kinds of studies have a lot of relevance here.

Now, the overall goal I have here is to convince you indeed that there really is a role of quantitative analysis in your deliberations, and it may or may not be in the area of using expected utility decision making.

A couple of slides just to kind of warm you up a little bit is definitional. First of all, what's clinical health policy, both because the center that I direct is called the Center for Clinical Health Policy Research, and people wonder why did you add the word clinical instead of just health policy. But also really because we're talking about clinical health policy development in this area, as opposed to more general public policy. Who are the policy makers and what do they want? Well, that should be self evident; it's you, but there are other health policy makers and they do want different things. But then fairly quickly I will try to move on to the question at hand, which is

.00170

quantitative health policy models, what are they, how can they help make better health policy decisions and more specifically, how should models be incorporated in the policy making in the MCAC.

Now, clinical health policy, the definition that I apply here is it that it's decisions that relate to the clinical enterprise. It can be somewhat indirect, as in the case of your group, your concern with decisions that have to do with how physicians practice medicine, how patients receive care. They can be the decisions themselves you're concerned with, should a woman between 40 and 49 seek mammograms, should physicians recommend carotid noninvasive tests for asymptomatic individuals? Or it can also be issues that relate to the health system, structures which support those decisions, such as electronic medical records in anticoagulation clinics.

And the distinction that I'm making here is, and it's maybe idiosyncratic my using the term, but I think the public policy is often really not specifically intending to deal with clinical causal relationships, but may often deal with clinical issues as a black box. So why this is relevant here in the area of your deliberations is that I think

.00171

that much of what you're concerned with require the careful input of clinicians in the entire enterprise.

While the health policy makers are everybody, of course, and I only list this taxotomy of micro, meso and macro just to make that point, that everybody does want to make decisions, everybody does make decisions, but I think there is a more important issue to make, which is that when decisions are being made at these various levels, there is an opportunity depending on how the analyses are done, for them not to be concordant with one another, or not to be consistent with one another, and I think that fundamentally, one of the concerns that certainly I suspect you all are interested in is as people at the macro level, that what you produce is something that is also consistent with the other levels as well. You don't want to be making determinations that are inconsistent in that sense because as I'm sure you have discovered on multiple occasions, that is the formula for disaster.
What do health policy makers want? Now, you notice I haven't yet talked about quantitative policy modeling, because to me the issue really isn't about the models. I mean, we can talk about the models and I will talk about the models, but really

.00172

we're talking about the policy making and the policy makers, that's really what the subject is here. And then secondarily, can we use quantitative techniques or any other techniques for that matter, to facilitate the decision making process.
And so in answering this question, we're talking now about having a client or a customer if you want to use those terms, they're the health policy maker, what do they want, that's the first question you should ask. And whatever anyone might say about you all, I believe that you as health policy makers want to do the right thing. I hope that makes you feel good. What does that mean? Specifically to evaluate uncertain and controversial issues in a calm and rational environment. The kinds of decisions that you are faced with often are going to be very stressful for everybody involved. Stopping, taking a breath, being able to think through the various issues is an essential element to

doing the right thing. And most importantly, you want to choose an action that represents the best choice under the circumstances. It seems self evident but again, at the end of the day you want to know that you have done that. And that I guess really gets to the

.00173

second issue, which is you want to feel good about what you have done. When you finish making a decision, you want to feel as though you have had the opportunity to make trade-offs between competing considerations, complicated decisions, you got to think them through, you have been able to think of them, again, is this calm rational environment. But ultimately one of the things, and research in the area of general policy development has demonstrated that people feel good, people don't feel good when their decision making prerogative is taken away. I mean, your policy makers, I don't know that you necessarily want me as the decision analyst or model developer to take away your prerogative to make the policy. You would rather, and perhaps there is some sort of optimization analysis, but rather, I think that you might prefer to be able to make your own decisions based on these various considerations. Just from a very practical perspective, again, in the general policy field, there have been multiple examples in which optimization approaches or prescriptive modeling, which I'll talk about a little bit more in a second, have basically been the examples of failure of the application of quantitative methods for policy formation.

.00174

And of course this is terribly important to feel good; you want to have decisions that are defensible to outside scrutiny. Now let's move on to the issue of what are the quantitative health policy models, and as Hal mentioned, you're talking about decision models, and I'm using the notion that quantitative health policy models as being a fairly broad class of models. We can talk about models in terms of what they are, but we can also talk about them in terms of what they

do.
In terms of what they are, they have three basic elements no matter what technique or approach they use. The three elements are structure, inputs and outputs, the structure being the critical components of the decision. What are the elements of the decision? They are the choices that need to be made, the points of uncertainty and the valued outcomes is the simplest classification. It incorporates inputs, and ideally health policy models are evidence based inputs, as opposed to randomly selected numbers drawn from the air. And finally, they produce quantitative outputs. For example, survival, quality of life years, incremental cost effectiveness ratios, and any other parameter that
.00175
the decision maker potentially might find useful.
So fundamentally, what health policy models are as opposed to what they do, are quantitative tools which are intended to influence the agents of health policy, namely health policy makers. So here is the link between the policy formation and the quantitative modeling. Health policy models are good if they in fact influence the agents of health policy in a way that the health policy makers feel good.
So how can health policy models help make better health policy decisions? Now by the way, I could have gone through some of the taxotomy of decision models and simulation models and discrete event models, and all kinds of models that can be used, which I didn't think was the issue at hand. So we can talk about that, if you like, but these again, I think all of these apply to whatever models one might choose in an effort to promote informed and satisfactory decision making.
So how can health policy models make better health policy decisions? And I think that Dr. Sox mentioned one of them, one of the commonly used approaches to using models for improving health policy and that is to develop prescriptive models
.00176
based on say expected utility. Now the notion here

is based in a theoretical foundation of Norman Morganstern's axioms of utility theory, and again, without going into those issues, you should read Alan Garber's textbook, or chapter, or various chapters and textbooks, and publications. But basically the idea is that to the extent one subscribes to the notion that the real world does or should conform to the axioms of utility theory, prescriptive models can lead policy makers to an optimum decision. Okay? Now as I implied before, the problem with prescriptive models is that there is a tendency to see them as driving the decision as opposed to guiding the decision, and so I raise the possibility, I think an important possibility for you to consider, and this is basically the theme of the whole presentation, that policy models can be, quantitative policy models can be extremely helpful even when they are nonprescriptive, and I sometimes use the term facilitative, because as mentioned earlier, models can provide an explicit framework for supporting but not necessarily prescribing policy. And as I mentioned, you know, if you develop a nonprescriptive model, ideally that should be developed in such a way that it is consistent with

.00177

the prescriptive model and that can happen if you develop the model with a utility structure as one of the possible outcomes you can produce. So you can use all of the insights that one might gain from maximum expected utility decision theory but at the same time for those of you who may not find that to be a compelling way of making decisions, you also have the opportunity to look at the other outcomes that you care about like survival, five-year survival, disability free survival, and so on. So these nonprescriptive models are not different models, they are basically a super set in some sense, where the prescriptive models can be seen as one kind of models that also are nonprescriptive models. I don't know if I made that very clear, but hopefully I did. So what can models do? They can provide a structure for visualizing complex issues. You can

only keep so many things in your head at one time, so finally putting it down on paper is a nice touch. In the case of the PET scanning issue, we were talking about a variety of concerns, including the natural history of disease, the efficacy of treatment, the potential side effects of treatment, and so on. So there were a number of issues to consider
.00178
simultaneously.
The models can provide a tool for expressing how indirect evidence relates to meaningful outcomes, and this again speaks to the issue that Hal mentioned earlier. If you do not have a clinical trial that says that, shows, a randomized trial of PET scan, for example, in which you randomize patients to receive PET scans or not receive PET scans, and then follow them forward and see how they do. I mean, that would be one potentially ideal approach to obtaining the information that you all want. And that hasn't been done, and in the absence of that kind of evidence, you need to create some indirect links that allow you to make these inferences, and models are one approach to do that.
In the case of a PET scanning model, we found them an extremely useful way to guide our literature search, and again, because through sensitivity analysis you can identify those factors that are most crucial in determining which would be a preferred strategy, and you could make most of your efforts, put most of your energy into coming up with the best estimates of those factors.
And finally, and I think this is relevant
.00179
here, it becomes a dynamic document, in general terms a document that serves as a repository of best evidence and can continue to be relevant as new data is collected. I suppose we could call this the full long-term employment for our center slide, because basically the notion here is that data will hopefully improve. There's going to be new tests, there's going to be new information about the epidemiology of disease, there are going to be new treatments. And

as all of these unfold, to the extent that the model that we developed was developed with these general issues in mind, you can come back and use it again. And of course, that was one of the reasons, as I mentioned, that we even included the possibility, for example, that a new treatment might actually make people, if people had a side effect from the treatment, they might actually have their dementia accelerate. That wasn't anything that we had evidence for, but we did that with the notion that somewhere down the line, there might be such a treatment.

And again, you can get all of these wonderful benefits without having the model prescribing the policy. Now, before I get off the issue about policy models per se, I'm going to give

.00180

you a taxonomy that it's based on how policy models fit into the policy making process, which is analysis, formation and implementation.

I think what most of what we are all familiar in looking at quantitative models in clinical medicine are what we would call in this taxonomy the freestanding model. That's the kind of thing that we all do as academics, and we say okay, we're going to do the reference analysis based on what we're told to do by experts in the field, we're going to publish that, you know, we're going to use the societal perspective, we're going to use quality, all that sort of stuff. We publish it, we hope that it's going to gain purchase in the community by virtue of its wonderful writing and its wisdom.

So hopefully that influences policy but generally indirectly, maybe not at all.

Now the model that we're talking about, that we've talked about today, which is the PET model, was I would like to think of as a member of this class of facilitative models, something where you had an opportunity to use the model, to question it. We could go back. We had a lot of interactions with Dr. Zarin and others at CMS that allowed us to make modifications to the model. And the idea of a

.00181

facilitative model then is it's developed specifically with a policy making role in mind. And then finally, sort of the holy grail for anyone who wants to have quantitative modeling applied to real world applications, it would be the embedded model, and I sometimes like to think that the work we've been doing in stroke is ultimately leading to this holy grail model in which you develop the model not only to better understand the evidence but also as a mechanism for promoting practice improvement, and we can go into that subject, but that's not what we're here for. I just wanted to mention that because that's I think the ideal. But again, all without prescribing a policy. One can use maximum expected utility notions in order to help understand whether you've gained insights or not for a certain model, but that's not the only thing you can gain. Some caveats, and some of these were mentioned earlier. In turning to the lessons of general public policy modeling such as economic modeling, weather modeling and so on, these concerns that you have raised here are concerns that are raised in all of these areas. For example, assumptions can drive conclusions. I could think of

.00182

examples, but I won't go into, time won't allow me to go too much more into the PET scan issue and perhaps we shouldn't here, but assumptions certainly can drive conclusions and people worry about that. Competing models may conflict. In fact, Bob, the notion that competing models may conflict is a concern, but it's also seen in some ways as being positive, that one should not disregard when models compete. And when I teach a course in which my students evaluate decision and cost effectiveness models, one of the potential jobs that they have is to find competing conflicting models and try to explain why they are different. And you can learn a tremendous amount from identifying the differences and the reason for the differences. Complex models are difficult to understand, again, a point just made, and very

legitimate, and that speaks to the issue that models that are more complicated are not necessarily better. Sometimes complicated models are just complicated. And there is a notion here that I like to refer to which again, I learned from Larry Phillips, is this notion of the requisite model, and the requisite model is a model which is sufficient in form and function to address the problem at hand without being

.00183

more complex than necessary, and you know that you have a requisite model because the people you're working with, your clients, your policy makers, they do not have a sense of disquiet that the model is failing to address some fundamental concern that they have, so that I guess is the technical definition of a requisite model.

Insights may not be insights, they may be bugs. And I think that all of these together basically speak to the issue that when you work with models, you're never going to be able to get away from haunting concerns about validity, which again, reinforces the notion that you should not be using models. I would not recommend that you use models in a very automated way in an optimization mode primarily, but rather in a facilitative mode, because you should always be constantly concerned about validity and the degree to which insights may actually represent errors.

Okay. So finally, getting to where the rubber meets the road for you all, my recommendations for MCAC, this really is daunting, but I will go on. No when to commission a quantitative model, that's an important concern. As mentioned earlier, when there is no direct evidence, that's certainly one of the

.00184

areas where modeling is something you want to do. When there are complex competing issues and it's important to visualize them graphically, models provide you a graphical way of doing that. And third, an issue of ongoing interest. Again, as in the area of PET scanning, this is an area you are going to be interested in for a long time, you have this model, you own it now, go use it.

You paid for it.
You have to have basic expectations for a model. I know that you're all aware of this, but some of the basic expectations are involvement of impartial experts to assure that the model reflects the fundamentals of the problem at hand. Again, at an early point in the PET scanning model, we engaged clinicians and said are we capturing all of the outcomes that people care about, which spoke to the issue of whether all the health states were being correctly addressed in the model. When we were concerned about issues having to do with, for example, what was the progression rate of dementia among individuals without AD and we had a lack of evidence in that area, we were able to sit down and work with these experts and they helped us at a very early point in structuring the analysis, and also

.00185

structuring the model, and also structuring the analysis that went into developing the inputs into the model, which is an equally important problem to developing the physical structure of the model.
Attention to extant models. In this case, there were at least two models out there that were published that we specifically sought out and spoke with the individuals who were responsible for those models to see whether we understood them entirely and whether our model either was consistent with that model or whether we may need to make some modifications, or if indeed the other model might actually have some problems with it.
Avoidance of unnecessary complexity, I spoke to that issue.
Ability to generate outputs that decisions makers care about, again, qualities are great. I like qualities in a lot of ways, they can be useful, but not everybody loves qualities and we just have to accept that.
Explicit accounting of assumptions. You will see there's a table in the analysis in the technical assessment, and most good analysis I think will do that, will provide an explicit accounting of assumptions, the assumptions that go into the

.00186
structure of the model and the assumptions that go into the estimation of the model, because in order to make an analysis that you can do in a reasonable amount of time and explain requires assumptions, and you have to list the assumptions and you have to describe what you understand those assumptions, what the impact those assumptions might have on your conclusions.
And then finally, and this could be a talk in and of itself, the issue of formal validation is in the case of the PET model, I will just tell you what we did, is that we had an analyst who was experienced in developing models sit down with the code, with the model, and go through it line by line, and see whether it made any sense and see whether there were any bugs that they could identify.
We went through a process also of varying each of the inputs over not only its plausible range but its implausible range to see if we had some counterintuitive results and with the notion that nine times out of ten counterintuitive conclusions are bugs. And there are other forms of validation that can be done.
But ultimately, I think establishing standards for what you expect of a model in these
.00187
various areas is essential.
And then understanding how to use a model is important. I think this is a good start, and I think having a basic education program that's not designed to teach everybody here how to do these models, because if we taught you how to do them then we wouldn't have anything to do, but rather to make you educated consumers so you can feel comfortable that you ask the right questions, that you understood what the sensitivity analysis really was about. What the heck was that two-way sensitivity analysis there, you know.
And making sure that the contractor understands the committee's decisional needs. We worked through an intermediary, that is Dr. Zarin at AHRQ, and that was a very useful experience, and we

also had several conversations with committee members in to do this as well, so we don't go off onto some path of developing a model that was inconsistent with your need. Also, consulting an external model expert. We were happy about it because it turned out well for us, but getting an external modeler to inform consumers, you're going to have additional questions and you should have a direct advocate by your side to help, and that was something that was
.00188
done during the subcommittee meeting and again, since the comments all worked out nicely, I have nothing negative to say about that component of the process. So to summarize, decision makers want to do the right thing and want to feel good about their decisions. Health policy models are quantitative tools which are intended to influence the agents of health policy, namely policy makers, and this is what it's all about, making good policy. Policy models need not be prescriptive, but certainly a prescriptive model can be contained within a general model that you use in a facilitative way. And this kind of nonprescriptive model can provide a framework for explicit credible defensible definitions, all without usurping your prerogative as decision makers, and after all, what do you guys get paid for? If you give me the problem and you ask me to do a model, and I optimize for you and I come back and say well, the maximum expected utility solution is, what the heck are you going to do? Not to mention the fact that I might be wrong, by the way. Only commission a model when it has the potential to help, and the three issues that I raised were sort of guides to that. And I would say that in general, you know, if it's a controversial or
.00189
interesting issue, even a simple model can be very useful if there are three or four components. Just laying out that model in and of itself can be a useful exercise in addition to being informative for everybody on the committee to understand how to use this process most effectively. So I'd say basically, use modeling often

but commission a model and the expense of commissioning a model only when it's a big issue like this and there is an ongoing interest. Establish standards and be informed consumers, and that's it. Thank you.

DR. SOX: Does anybody have any questions for Dave before we open up the general discussion? Thank you.

Well, we now have some time to discuss the role of decision models in our deliberations. Our interim recommendations started out without really explicitly incorporating decision models into them. Then we developed a sort of simple approach, a quasi decision analysis approach for evaluating diagnostic tests which got us through the first round of PET scan discussions, and now we've gone to expected value decision making to guide us. I guess the real question before us is, are we going in the right

.00190

direction? Has this been a good experience or are we digging ourselves a hole, either in respect to making poor policy recommendations or losing our credibility because we're using a technique either inappropriately or a technique that people simply don't understand or want to be guided by.

So Alan, maybe if I could ask you to begin the discussion, what do you have to say on the specific question of the direction we're going and also are there some issues that we should be paying attention to that neither David nor I covered.

DR. GARBER: I think what we have seen thus far, at least from my perspective, and I guess I'm not purely objective or impartial on this subject, is that there are many questions we could not have answered if we had avoided the use of models. And I think there could be no doubt, for example in looking at the multiple indications for PET scanning, that if we had said can you directly demonstrate an improvement in health outcomes, we would have just had to say across the board that's just not possible with the existing evidence base. So I think if we were to say you can't use models, you can only use direct evidence, we would be

25 painting ourselves into the corner because people
.00191
1 need to make clinical decisions and we would not be
2 able to reach affirmative decisions very often if
3 that had been our criterion, we would have to wait a
4 long time for trials to be done, and they might never
5 come.
6 So I hope that it's not controversial that
7 the general idea of using modeling is the right one.
8 But I actually also have to say that I think in
9 practice it has worked out pretty well thus far in
10 terms of highlighting what the key issues are where
11 the data are uncertain, where we could use more data,
12 where we can draw fairly firm conclusions, and I
13 think it's very hard to generalize and we will have
14 to deal with use of modeling on a case-by-case basis
15 going forward.
16 But I personally am very encouraged. I
17 think that David Matchar and his group did a very
18 nice job in this study on PET for Alzheimer's, but
19 the other contractors have also done very nice jobs
20 and I found these very informative analyses. And I
21 have to also say that even if I were to disagree with
22 some aspects of the analyses, maybe some of the
23 assumptions made, they were extremely useful
24 frameworks for focusing the discussions. So in that
25 sense, I think it has been very successful.
.00192
1 DR. SOX: Perhaps you can address a
2 technical question that comes up a lot in my line of
3 work, and that's sensitivity analyses in which you
4 basically do a simulation in which instead of
5 assuming -- and you assume some sort of distribution
6 for the values of the various factors and then you
7 basically look at all of them at once by doing some
8 sorts of simulation process that gives you some idea
9 of the distribution of expected qualities or expected
10 value instead of point estimates, and also have a
11 chance to look to some degree at the interaction
12 between different variables.
13 How should we be approaching this problem
14 of sensitivity analysis so we don't over simplify it?
15 DR. GARBER: Well, Hal, I guess you're

asking should we routinely do probabilistic sensitivity analyses where eery parameter varies simultaneously, and you know, this is something that I'm not sure we could do justice to in a brief time today and I won't really attempt it except to make the general observation that it entirely depends on the individual study or the individual subject, what kinds of data we have.

I think the guiding principle for every study that's used to help us, I like David's use of

.00193

the term nonprescriptive and I think that all of the time we do nonprescriptive or we look for nonprescriptive analyses, ones that will help us reach a decision but that don't tell us what the decision will be. But in any case, in each situation, sensitivity analysis has to be suitable, and what's a suitable sensitivity analysis depends on the data and how many things are really uncertain. Hal is alluding to a technique that's very popular and has many advantages, but it's only useful, it is only as good as the assumptions built into that type of sensitivity analysis, and I think again, we would have to look at that on a case-by-case basis. It can be very helpful in some circumstances.

DR. SOX: Let's see. Go ahead, Bob.

DR. BROOK: I would just add two comments. One, it's foolish to think that the sensitivity analysis in these decision models or any models are not subjective and not always explicit. All of them require a great deal of judgment by the people who model them and how you fill in missing data even if you do this with wide sensitivity bands, because even that requires subjective judgment, and virtually all of these, they do not turn out to be robust to all of

.00194

the decisions that people are making. That doesn't argue against using them. I believe we should use them, but I believe the major purpose of this should not be for this committee. I believe that the major purpose of this should be to help provide guidance to the industry of what are the

critical pieces of information that are missing in order to make better decisions about what things ought to be done or not done, and I would urge the government to try to explore the use of these techniques way up front in the process when these technologies and devices are actually developed, so that people would really understand what are going to be the critical pieces of information that are going to drive the modeling and the evidence, and they are not always obvious without setting up these explicit models.

So in response to David's question about when to do these things, almost any technology that is going to consume millions of dollars of money, which I think would be the hope of almost any one of the manufacturers in the room, if not tens of millions of dollars or in the case of PET scan, billions of dollars, there is a need up front to determine what evidence is needed, and I think by

.00195

setting up these models up front, they ought to do it and we ought to encourage through our modeling just what the -- one of the products out of this ought to be what are the things, gee whiz, why did you miss doing this ten years ago and then come to us and be upset with us when we say that a specific technology is not worth it. There needs to be some thinking up front using these models, this kind of approach to figure out what kind of information is really needed in this new age to produce a positive response from this committee. That's where this ought to be put. For us to do it at the end and get into this usual contentious process, yeah, we're going to have to do it, but it's not very satisfactory.

DR. SOX: So based on David's model, the folks who were interested in building PET scanners might know the target sensitivity that they need to reach in order to have it be useful.

DR. BROOK: Absolutely, or the subgroups, or whatever, but I really don't believe that's what David wanted to do when he talked about the user. I don't know if there is something out of this work that you have written a report for the companies that

make PET scanners, to say here is the critical pieces of information you need to know to really prove that
.00196
your stuff is useful in improving the health of people with AD, or what are the critical questions, the pieces of information you need to know, if any, if the outcome of this modeling is until there is a better drug or a drug that is both better and no more side effects, which is what I think I heard, then that's another story. But what I'm really asking is, we ought to try to prepare useful constructive reports for the industry to help them make better decisions and collect better data, and that ought to be one of the options for these models.
DR. MATCHAR: In this specific circumstance, this was, this group or the subcommittee was the recipient of this, but I think historically manufacturers have not been interested in that, and only recently are they starting to, but I think that they are looking to it, and this is why I find this whole enterprise very exciting, because I think manufacturers are starting to look at groups like this and say well, okay, if this is how you're going to do it, we'll play that game, because they know that if they can understand someone more explicitly how you are going to be making your decisions, they will engage in this activity and they will start to develop models early on, and some of
.00197
the many manufacturers have approached various modelers and the Society for Medical Decision Making has even toyed with the notion of having some sort of consulting service to guide the manufacturers to increase their use of these technologies. And there are plenty of private consultants out there doing this. But it is certainly something that needs to be done very very early on rather than at the end, and I agree 100 percent.
DR. SOX: Perhaps I could interject a concern here. If decision models which operate on imperfect data become the norm, are we going to find ourselves dealing with more and more imperfect data because the motivation to create higher quality data

won't be there?
DR. BROOK: Well, Hal, I can't take it to devices, but flying over here I read a series of controversial papers on depression drugs that are newly developed now being tested against placebo and not other drugs. Now again, to test against other drugs, you have to do some sort of modeling, because you have to relate it back, and what they show is depending on the year and the study that things are done, the cure rate for major depression vice placebo varied from 3 percent to 29 percent.

.00198

So in the purest sense of the word, you know, everything has to have a health outcome and everything has to be tested in a randomized trial. We know that that's not going to happen, and I would think that we don't run the risk of encouraging both decision modeling and real data. And one of the things that ought to come out of this is the balance between the two.

Now, I'll add a third piece. I believe that it's absolutely incumbent upon people to change the willingness to give information so that we're going to have to go through large observational databases as well. For instance, I think it's absolutely wrong for somebody to offer PET scanning to anybody without first putting together, let's say I have an entry point for PET scanning and I make a decision to do it or not do it. If I'm going to do that, I think the commitment for people who get that, even though it may be paid out of their own pocket, ought to be that you develop an absolute complete database of the covariants up front, the decision that was made about whether to give it or not, and outcomes, and that that be pooled and not just be the purview of one single group of investigators. And that because I think in the modeling in general,

.00199

somebody is going to have to have that resource. And so I think these are the types of changes that we can stimulate, because I don't think we us observational data very well, we don't collect it very well, we don't use modeling very well, and we certainly don't

have a balanced decision between randomized trials and decision analysis very well. And with the new IRB approvals, I am just wondering what we will be able -- I mean, I am seriously wondering what kind of randomized trials we will be able to do given where the IRB process is now going in many of the academic institutions.

DR. SOX: That's the other side of discouraging randomized trials is maybe they are not going to be feasible. Leslie?

DR. FRANCIS: I just want to say that we're sort of roughly on the recommendatory, which is a kind of soft prescription side of things in what we do here, and as somebody who you know, sort of sits here expected to play something like that role, when I read a model, I have the meat hamburger theory of models, you stick stuff in it, and the most important thing is what you stick in it, and therefore -- and then grind it up, and therefore, it's most helpful to me when I read a model to get as much information as

.00200

I can about the choices and the controversies that were involved in the choices about what to stick in. I mean, I thought I had a pretty good sense of that from this model, but actually the one that was hardest for me to swallow just reading it, you know, before I came, was the no treat, I mean the treat all assumption. So that's just a comment about, as a consumer of a model who's going to have to make certain judgments about what it can be used for, the more information there is about those up-front choices and what they are based on and how hard or soft they are, the happier I am.

DR. SOX: Unfortunately, journal editors in their desire to publish more with fewer extensive pages often end up truncating the type of discussion you're calling for.

DR. FRANCIS: Exactly, but when somebody prepares a model that we buy, they don't have to meet the journal editors editorial strictures.

DR. BROOK: You may need to talk to David about this. I would think it would be useful for CMS if we're going to use these things, to develop a web

based library of these models. I don't know who maintains that and whether it's the Society for Decision Making, but I really don't believe that the peer reviewed 12-page format of models is appropriate at all, and I would believe that we probably eventually will need to go on record that any one such model that comes out of that, that the whole explicit activity ought to be in some web based preserve archive and that somebody needs to review that to say that if David Matchar no longer was interested in this subject, somebody else could pick up what he did, understand it, and do exactly what David said, which was to build on it with new data and new evidence, without starting over again. So how we catalog the programs, how we make this user friendly, we just have never developed that. So even if it's explicit on a web, it's probably unlikely that any smart computer person will be able to figure it out enough to actually redo it, without spending hours with David's team to do it.

DR. SOX: But one thing we could do along the same lines is to have a certain user friendly interface with the computer so that any person could go up and mess with the variables. Wade, I think you're next.

DR. AUBRY: I just wanted to add a comment that I thought we are on the right track and I would like to see CMS and this committee further develop the models, but I would also like to share the concern about potentially inhibiting the development of direct evidence. I know in my experience with the Blue Cross/Blue Shield TEC program, there were numerous instances of reviewing an intervention for a relatively common condition in which it would be relatively easy in terms of generating patients to study the intervention but it was not done, and the quality of the evidence was low. So I think the decision to proceed with a model for evaluating a given topic should be on a case-by-case basis and it should include not only the issues that David referred to such as there's little

if any direct evidence and that a model could potentially help, but also, there should be a determination that it would be difficult to develop the type of direct evidence that would be needed. And there are examples in my own career where I have looked at a topic, particularly in diagnostic testing, in which a study could have been done, and then I looked at it again ten years later, and there still is little direct evidence. So I think the committee and CMS should have a reasonable high bar for direct evidence when that evidence could be developed. If it's a very rare condition, that

.00203

may be a different consideration, and we touched on that a little bit at the last meeting with some, I think there was some discussion about pediatric cancer interventions for example. But I just wanted to underscore that point.

DR. SOX: Sean, do you want to comment on this issue of discouraging better forms of evidence by using models that incorporate relatively imperfect data. Maybe you'd like to say something after you have a chance to reflect, or maybe you'd rather not say anything at all.

(Laughter.)

DR. TUNIS: Maybe I'll do both of those things, say something without saying anything.

(Laughter.)

No, I will sort of turn it into a question, which is, you know, my assumption is that built into the techniques of modeling is consideration of the underlying evidence that goes into the model and therefore, the output of the model will naturally reflect, or the sensitivity analysis in some way, the quality of the underlying evidence. So I would presume that the mere fact of using modeling as a component of decision making wouldn't by itself, you know, undo the incentive to gather

.00204

high quality direct evidence as an inherently more reliable form of evidence. But, that's virtually a question, but that reflects my own knowledge about modeling but also the use of evidence in policy

making, so that we would continue to consider the quality of the underlying evidence even if it were incorporated into a model as part of the thinking process.

DR. SOX: Of course the decision model can tell you the parameters of the model for which you require high quality evidence and on those, you may not be able to make a coverage decision because, it was not the case here, but there are some situations where a critical parameter makes you unable to make the call because the range of reasonable values for a probability or utility, you know, the preferred option changes so it becomes a close call without more precise knowledge of what's good data.

DR. BROOK: One of the questions, Hal, that I thought would be interesting to look at would be, you know, the Duke cardiovascular database produced similar results to the randomized trials about when to do bypass surgery, very similar. And it would be very interesting if we could develop a body of information about when this kind of work, and

.00205

when good modeling and this kind of work, you know, the interface, the balance between that and randomized trials. Nobody has done that, but when those studies have been published, you know, how do we use them, because we're going to have to make some decision. And if we only had that kind of data on a controversial topic like that, would we approve it? I mean, if all we had was a good observational database about left mane disease and medical or whatever therapy, would we approve it?

DR. SOX: Barbara.

DR. MCNEIL: I have a slightly more simple minded suggestion than Bob's, but first a comment. Wade, I actually looked up the Blue Cross series of evaluations for the year 2000, and tallied the percent of them that we couldn't make a decision on because of inadequate database, and it was 40 percent, so it was quite large for that year. But my suggestion, going on what everyone said, is the following, just a summary. It seems to me that Bob Brook has said that these models are

useful for identifying up-front variables on which the decision is key. And then we said earlier that the result of this model was also useful for identifying prospective research studies, and there
.00206
is some relationship between the two of those. And the third thing that we said was that a lot hangs on the Y axis on the utility scale whether it's a simple expected value, or even if we get to cost effectiveness, that utility scale means a lot. And I have been impressed listening to the comments around the table, and I wish Randel were here. Would it ever be possible, and maybe you've thought about this, Sean, for CMS to have a conference in which we, a very small conference in which we invite leading device manufacturers and pull out some of the examples that we have had in terms of raw data and modeling data. And say for example, let's look at the breast cancer one on PET, and before we ever did a study, we would have known that that wasn't going to make it, just because of the implications of false negatives. And now let's look at the model of PET for Alzheimer's and what might have been predicted beforehand and what implications those models show for future studies. And basically have a little bit of a tutorial about how industry should be thinking prospectively about how to maximize the chance that the good things will get properly evaluated and move through the system. I think there really is an educational
.00207
gap, because this does get pretty technical and it has never really been put together in an integrated user friendly fashion, and I think maybe we could do that.

DR. SOX: If we could agree, if industry developed decision models to guide their evaluation process, and we did this in a more cooperative fashion with the clinical community, we could make an advance, I think.

DR. McNEIL: I was suggesting to help them come to their conclusion, we would give them some examples from our portfolio.

DR. SOX: Yours would be a first step, and not in the context of a specific issue where it's all charged up and nobody wants to concede an inch.
DR. MCNEIL: Right.
DR. SOX: Alan.
DR. GARBER: Well, Hal, I want to get back to your question about whether the use of models would dissuade various parties from carrying out trials basically, and when it comes to diagnostic tests, I am not terribly concerned about that issue, since right now there aren't that many randomized trials done, it's a fairly rare event to have randomized trials that look at final health outcomes.

.00208

So there is not really anything to discourage at this point. And I wish I had a view of the future that was so rosy that I thought there would be a lot of these trials coming forth that we might somehow discourage, but I don't think that's realistic.
I do think that it's very important for the panels, though, to continue to do what they have done so far in looking at the data and being pretty clear about looking at the assumptions and whether good data support these assumptions. In the context of the PET for AD, they look carefully at whether there is enough data to suggest it was sensitive and specific, what the basis for the accuracy result was and so on. And we depend very heavily on our panels to be consistent in application of evidence criteria for the components of the models, and that's a lot of what we heard about.
And as long as they continue to apply those criteria, the people who produce these technologies will indeed have an incentive to do appropriate studies to look at these aspects like sensitivity, specificity and whether treatment works given the diagnosis.
DR. SOX: John.
DR. FERGUSON: Just a question. Does the

.00209

FDA accept models from the industry or from anybody?
DR. HOLOHAN: Maybe I can answer that in the absence of an FDA official here, that the FDA

rules are in the Code of Federal Regulations, and the CFR states on the basis of well controlled trials, and that's been held by courts to be more than one trial. Controlled doesn't necessarily mean randomized blinded trials, but no, you have to provide substantive evidence of safety and effectiveness.

DR. TUNIS: The only other comment on that is that the FDA rules for adding a second clinical indication to an already approved labeled indication, in other words, changing an off label use to a label claim, can be supported in some cases by a meta-analysis of small trials without doing what would qualify as a fully powered pivotal trial, but that has only occurred in a few cases and I am not aware of any example where a decision model has supported even a second indication.

DR. HOLOHAN: To further complicate it, there's the issue of surrogate also. The FDA does accept certain surrogate influence as evidence of effectiveness of antihypertensive drugs, lipid lowering agents.

.00210

DR. SOX: Bob.

DR. BROOK: But we're in a whole different game here, and that game is the same game. I mean, if somebody asks us to evaluate the effectiveness of a neurological workup for AD, should, is the technology so simple if it was applied by an internist or family practitioner that you don't need a neurologist at all, then we shouldn't pay for it. Or what's the frequency of visits that a doctor ought to be paid for for hypertension, or do you need any visits with e-mail?

Now, all we have done is throw this back to recapitate physicians and then let them figure out what to do and worry about them being sued, but we have no evidence on any of this kind of stuff. So we're now in this whole issue of devices, visits, professional time. We have no evidence, nor do we randomize before we decide that an HMO can cut the time with a patient from 15 minutes to 12 minutes, we don't randomize that and look at outcomes.

So, we have a mess, and I just wonder whether or not the stuff that Barbara was talking about, if we change the way we operated and said that down the road this committee is likely to require randomized trial data here, modeling here, this was

.00211

based up front on getting together with the manufacturers of a product very early on in the developmental phase so that one could actually try to be smart enough to model out this, and say look, here's what the key data are really going to be, and then this committee could up front, or one of the subcommittees say, it's likely that we will probably want randomized trial data here, or model data here, that would be extraordinarily helpful.
I mean, they are in an industry -- I mean, this is not the tobacco companies trying to kill people. This is a group of people that are trying to make people better. And you know, the faster that happens, the better off we will all be. So the question is, can we change this process, are they willing to do it, but I think the government ought to explore that.
DR. SOX: Further comments or questions?
Yes, Leslie.
DR. FRANCIS: These models always get constructed as against existing practice in certain ways, so that suppose for example, PET had been on the horizon, I mean had been adopted 30 years ago, and the question were clinical evaluation. So the question was, is clinical evaluation, does it change

.00212

the decision tree, you know. In a way I would have wanted to know that, but I didn't know that.
DR. SOX: Compared with what?
DR. FRANCIS: With PET. I mean one of the things is you always do a model against a background, and the more we know about, one of the things that's just a mess here is that there's all this stuff that everybody just does that nobody has ever looked at, but the more we know about what the background is, the better we can decide how to use the model.
DR. SOX: Well, I think I will bring this

discussion to a close by suggesting to Sean that we have uncovered a lot of opportunities and potential problems as a result of having had a very successful discussion about a technology based almost entirely on decision modeling and imperfect data, and that we do need to continue to talk about this off-line, and we have had a number of good suggestions about developing our ideas, starting to talk to industry about it, developing databases of decision models, and trying to encourage a lot of very early stage planning of both evaluation and in fact what you ought to try to build, and we need to continue that process. And I guess on behalf of the committee, I will ask you to figure out a way to get some of us at

.00213

least together to start thinking about that.

DR. TUNIS: We will certainly do that and follow up on suggestions that Barbara, you raised, and I think Bob also, to get more fleshed out some of the ideas you had, including potentially working together with AHRQ to have a small conference to follow up and explore this in a little more depth. I'm volunteering AHRQ because they have more money than we do.

(Laughter.)

So we will certainly follow, take that as a directive from the executive committee and follow up on that.

DR. SOX: Now would be an opportunity for anybody who would like to step up to the microphone and comment on this particular topic. We would like advice from some of you who are in the audience. Yes, if you would identify yourself and where you're from.

DR. HERNANDEZ: I'm John Hernandez. I'm a health economist and health outcomes researcher at Boston Scientific, a medical device manufacturer, and I appreciate the opportunity to follow the committee's discussion on decision analytic modeling and the opportunity to share my thoughts with you on

.00214

this issue.

In general the medical technology industry

is strongly supportive of using appropriate analytical and empirical modeling to support decision making in health policy, and it's important that these tools be used in a manner that insures transparency and accountability in the Medicare coverage process. And if used appropriately, we do believe that decision analytic modeling offers the potential to support policy makers in understanding, analyzing and making decisions regarding these healthcare policy problems.

There are a number of examples where decision analytic models have been used successfully to support complex decision making. In World War II at military decision making, at Rand, and in the United States Government at the Pentagon, and in other situations. It is also be being broadly used in the business community to make decisions as well.

More recently, a growing body of research has been published demonstrating application of these techniques to analyze healthcare decisions. And that said, there has been limited study to date of the applicability and usefulness that decision analytic modeling techniques directly and formally to assist

.00215

healthcare policy decision making. The increased adoption of structured methods and processes for decision making in Medicare has been underway for quite some time and we applaud this trend as it provides a level of transparency and accountability that is rightfully demanded by the program's recipients and participants.

As you begin contemplating the appropriate role for decision analytic modeling in the Medicare coverage process, we ask that you keep in mind several issues, some of which will echo the points already made by Bob Brook, David Matchar and in the recent discussion, and we actually note that several recent reviews have concluded that decision analytic modeling is likely to be used and most usefully integrated into the policy process, if it is used simply as one form of input into the decision making process, and that there are dangers to utilizing such models if they are adopted in a manner that's overly

rigid or mechanistic, as they will be viewed rightfully and as mentioned by David Matchar as usurping the role of decision makers.
One researcher, and this was mentioned before, who has investigated the utility of decision analytic modeling in the consensus panel development
.00216
process, David Powker, has identified a number of problems that can limit the utility of formal modeling in conjunction with these panels. In fact, a major conclusion was that panels almost uniformly have avoided specific mention of decision analytic concepts in their recommendations.
And the role of decision analytic modeling in the panel recommendations discussed today is relatively unusual. Specifically, the medical, I as a representative of the medical technology industry recommend that if decision analytic models are going to be used formally in the MCAC process that you pay some attention to the processes and to the methods that is going to be employed by the MCAC and that specifically formal recommendations should be developed in the MCAC recommendations for evaluating effectiveness. And specifically, I believe this should be discussed in the suggestions for panel operations and in a separate section under issues on methodology.
We also would very much like to engage with CMS and perhaps AHRQ to discuss this informally as well.
I would like to point out that results, and this has been discussed I believe already, that
.00217
decision analytic modeling can be highly dependent on a large variety of methodological and technical factors. I think that there is a large literature showing that cost effectiveness analyses can have up to 20-fold differences in results due to technical decisions made by model constructors and similarly in the utility analysis world, Eric Nord and others have shown the impact of methodological differences in results on measuring utilities.
Other issues that have been mentioned that

can impact outcomes have to do with the subgroups chosen, time horizons used in these models, decision options, consideration of costs, and many other factors, and these should be specifically considered before being broadly used in the MCAC process. Thank you very much for your time.

DR. SOX: Thank you very much. Anybody else that would like to step forward? Yes, please identify yourself and where you're from.

MS. MARX: I'm Sandy Marx, with the American Medical Association. I thought this was a very interesting discussion and I was going to get up just to repeat something that the previous speaker mentioned about integrating the use of decision models into your recommendations for evaluating

.00218

effectiveness, so that there is some comprehensive tool that people have when they're looking to pursue a coverage decision that identifies all the possible ways you're going to look at the available evidence. And I think that document also gets into a sort of hierarchy of evidence, which sort of addresses the problem you raised about are we going to discourage doing randomized controlled trials. Clearly anyone who wants to see their new product covered is going to realize that if they have good evidence from multiple randomized controlled trials they are going to have a greater likelihood of getting a positive coverage decision than if they just have expert opinion. And somewhere in between those two extremes maybe, or two ends, is going to be this decision model process that can pull together the observational study. So to integrate that into the rest of the work that you've done I think would be helpful, because there are a lot more people interested in what you are doing and how you do it than just the people in this room who have heard this discussion.

DR. SOX: Thank you. We're now going to hear from Sean about changes to, in the voting process, and I guess our responsibilities. Is that

.00219

right, Sean?

DR. TUNIS: Well, actually just to get some input from you on some ideas about how we might reformulate some of the questions. And I promise, we will be done by five after three at the latest, somewhere around then, and actually, Jeff Kiang is going to come back and wants to talk to folks briefly, and he is finishing up a talk at three, so he will be back a few minutes after that.
So this is really again, to get some feedback from you all on some internal discussions we have been having about possibility of potentially improving the way we frame the questions in our panel meetings to get a full, the full range of views express that are relative to coverage decisions. Our observation is that our current voting questions may not give a full opportunity to address all the issues that are relevant to coverage decisions. We do sometimes find ourselves, you know, after MCAC meetings as we're trying to formulate policy, wishing we had a chance to go back to the panel and ask some additional questions, so what we are talking about here is whether we can modify the existing questions or add some new questions that would get some more information.

.00220

The issue is whether the current questions are too narrow, particularly a yes-no question on the adequacy of evidence, it's obviously a critical question, I'm not suggesting that we eliminate that question, but it does provide sometimes a framework that has in some cases been interpreted as too narrow. Obviously this is fundamental to the guidelines on evaluating effectiveness, so again, this is not intending to replace that, but trying to see if there is some way to open it up and maintain the value of that framework as well.
But there are other perspectives, opinions, observations about a technology, the condition that the technology is directed towards, alternative technologies that are important issues and may be relevant to have the panel discuss in this public forum.
One thing I just want to point out is that

the MCAC, as you know, is one of the only existing public forum dedicated to discussion of complex issues raised by new medical technology. There is the Blue Cross/Blue Shield has a technology forum and I'm not aware that that's actually a public forum. The NIH consensus forum is a forum in which technology issues are addressed, but really for

.00221

Medicare being the largest single payer, when it addresses a technology, it has ramifications that are much broader than just the Medicare program, and we may not be fully taking advantage of our opportunity to get into the more complex controversial issues raised by medical technology just focusing on adequacy of evidence.

So anyway, how can we make best use of this opportunity given these issues? So, the first notion is whether in addition to or instead of the yes-no on is the evidence sufficient, is it possible to, would it make sense to ask a question of how would you rate the overall quality of evidence, possibly dividing it into simple categories like good, fair or poor, rather than the simple yes-no of it is sufficient or it isn't sufficient. In other words, that the panel would come to some assessment, sort of a graded assessment, if you will, of the quality of the evidence. I think this issue has been raised and possibly discussed before but we're raising it again for discussion.

And then some additional questions, some of which by the way were reflected in those additional questions that were directed to the imaging panel related to PET for Alzheimer's was the

.00222

beginnings of our thinking on this. Magnitude of effect, this is already part of the interim guidelines, so this is just, this would not be a change. We're talking about the relative benefit, the size of the effect over existing alternatives, less effective, more effective, equally effective, I think there is seven categories in the existing interim guidelines.

But new questions would be what's the

clinical impact of the technology, in other words talking about what's the current state of practice absent the technology, what potential harms might the technology pose if it were broadly disseminated, are there specific patient selection criteria that need to be considered more narrow than should the thing be covered or not covered broadly, so to try to get some more into some of these issues.

The research consequences, a little bit of this has been raised today in the context of the modeling discussion, but would coverage and broad dissemination encourage or discourage research, or improve or not improve the existing medical evidence. In other words, looking at the potential impact of coverage on what sort of research might be conducted, identifying what additional research should be

.00223

performed to make the best conclusions even if we do offer coverage, you know, being more specific, trying to get the panel to be more specific about what the follow-on research should be, and also talking about what barriers might exist to performing the research, including burden on industry, the cost, ethical difficulties in potentially doing studies, and possibly methodologic challenges of doing follow-on studies.

And then, we don't often talk about criteria such as what sort of staffing and facility criteria should be considered in the use of the technology. We do this for transplant facilities routinely in the Medicare program of having certain minimum outcomes per year and certain numbers of transplants, but we don't talk about certain facility criteria for other technologies, and this issue doesn't often get raised.

And then questions about whether there are certain characteristics of a device that should be considered in terms of when we evaluate a technology. For example, the whole issue around PET scanners of gamma scanners versus full ring PET, we often don't get into details within categories of technology.

In terms of the process for formulating

.00224

these questions, these potential additional questions, the idea would be that we would try to come up with the questions prior to the meeting based on some standard formula, in other words, identifying the content area of the questions as we have done here, but they would be customized to the particular issue, and we develope the questions in consultation with the requester, advocates, industry, consumers; in other words, this would be part of the open public process of actually getting input on what questions should be posed to the panel in addition to the questions evaluating the evidence.

So, we would like to hear a little bit of discussion for the next five or ten minutes, this won't be the end of the discussion, but what are the pros and cons of modifying the existing key questions, just yes-no on the sufficiency of the evidence; if the answer is yes, what categories of questions should we be looking into. Any suggestions on the process for developing the questions, how we can engage MCAC in this discussion, and any other thoughts on this topic.

So I'll sit down and take your comments, but again, the idea is trying to figure out how, given that it's difficult to get all your brains in

.00225

the room at the same time, we want to take maximum advantage, plus, we pay you I think $400 a day and we want to make sure we get our money's worth for the public.

(Laughter.)

DR. SOX: Leslie.

DR. FRANCIS: I mean, you guys shape us in a way, so if you want more information, it's entirely reasonable to ask for it. The only thing that I must say that I found frustrating several times, and it was illustrated again today, is if the question gets changed in some way, that people don't feel that they had an adequate opportunity to answer it. What happens is that things get shoved over into process questions and rather than substance questions, and we want to try to avoid that as much as possible. So if questions get added or if people realize that

questions need to be changed, it should be done with adequate time for airing these discussions.
DR. SOX: Are you talking about the questions that sort of form the basis of the charge to the EPC that writes the evidence report, is that what you're talking about?
DR. TUNIS: No, this is questions to the MCAC, to you know, to guide at least part of the
.00226
discussion, and again, sort of with the intention of trying to systematically raise all of the critical issues that are going to be pertinent to making coverage decisions. So, I mean, just as an example, we don't often systematically talk about what's the standard of care and the adequacy of the standard when discussing a new technology. I mean, there are a whole set of issues that sort may sort of haphazardly come up in conversation, but we haven't traditionally made the effort of trying to identify those in advance and make sure that they get addressed in the conversation.
DR. SOX: Alan.
DR. GARBER: Well, Sean, I think the set of questions you raise is quite interesting and there are many directions you could go with this, and rather than address all the possibilities, let me just give you my two cents worth.
First of all, I think it is very important for MCAC as a process to be successful to be fairly consistent with at least a core of questions, and I think we have that start right now. So I would first suggest that anything we do should be considered add-ons rather than substitutes for what we have already developed.
.00227
And on the question of add-ons, some of the things that you are raising, first of all, will only be interesting in particular cases and would not be very interesting to do in general. And secondly, some of these may be far more difficult than we recognize today. Like an evaluation of standard practice. You can imagine how difficult it is just to define standard practice, and then to try to find

a body of literature that would enable you to draw fairly firm conclusions about the state of evidence backing up current practice, I can certainly imagine that that would be something worth pursuing in some instances but I would imagine that's going to be rare.

So, I think that if you're asking should we occasionally include some of these additional questions, the answer is probably yes. I mean, I think as a member of MCAC I would be happy to be involved in a discussion of these added topics, but it has to be your call about when it's important and relevant. And you have to recognize that you have to add on questions very judiciously because we could eat up all our time on things that may not be all that central to your coverage decision.

DR. SOX: Just to challenge you a little

.00228

bit on that, Alan, it strikes me that we don't do a very good job here of putting the candidate technology into the context of total care for a problem. To give you an example, the mercury sphygomanometer, we could evaluate that and find that it's accurate and that it adds lots of quality adjusted life years, but in so doing, if we didn't identify the problem with compliance with medication, failure to screen, failure to do cardiac risk reduction assessment in all patients that are hypertensive, those may be the things that are barriers that are getting in the way of successfully dealing with the problem of hypertension and its consequences.

And I have a feeling that when we respond to a coverage request that we're going kind of doing this, looking at it through a pea shooter, and that we really need to step back as part of our evaluation, put it into a larger context, not as the people that are advocating for a technology that may be fairly marginal, but how does it fit into the big picture of how good a job we're doing and what the rate limiting factors are in succeeding.

DR. GARBER: Well, the issue here in an ideal world where we had, or I should say imaginary

.00229
world anymore, where CMS had unlimited resources and time, Hal, I think the thing that you're proposing, to do a very comprehensive review of all the treatment conditions and the basis for all treatment and condition would be something you would want to do. But we're talking about tradeoffs here and what CMS really needs to know to make a coverage decision. And that's why I'm saying you have to be judicious. And what you're talking about doing for hypertension, we would be spending 10 percent of our time potentially on the topic of ambulatory blood pressure monitoring, and 90 percent on the treatment of hypertension. Now that's an exercise that has considerable value, but it also has considerable costs, and I think there are going to be times when that's going to make a difference to the ultimate coverage decisions, and times when it isn't, and I think Sean and the CMS staff need to be very careful when making those choices, deciding when to do the big big study. But we've seen how much discussion we generate over fairly narrow questions and when you throw in about 10 times as many considerations, and that's no exaggeration on the topic you're talking about, you have to figure, are we really going to be able to do an adequate job on any of it.

.00230
DR. SOX: Well, there is probably some middle ground between my quickly listing a few topics and doing a very expensive analysis that's already been done out there in several existing articles. But anyway, there's probably some middle ground. Other comments? Wade.

DR. AUBRY: I was just going to say that I think there are two categories of questions. One is being responsive to the coverage request. The coverage request comes in, there's a process that has been published in the Federal Register as to how the Medicare coverage process works, and so one question that needs to be addressed has to be responsive to the requester.

The second part of it, however, is to develop other questions that make sense to evaluate

in the context of looking at that subject, and it seems to me that that should include input from other groups, including the committee itself, perhaps the Medicare carrier, medical directors, perhaps interested specialty societies, and so a broader call for key questions may be appropriate. But I share Alan Garber's concern that that may yield sort of an unwieldy number of questions, and that would need to be sort of narrowed down to the key areas by CMS staff.

But I think that the committee in order to adequately address a topic should do more than just respond to the coverage request if they're going to look at a topic.

DR. SOX: Bob.

DR. BROOK: I'm a little confused and I think at some point we may need to revisit the entire process. Let me tell you what I'm confused about. Today we said we weren't going to use this procedure. In reality, if a PET scan was 50 bucks, given the supply constraint of anyone competent in geriatrics or AD to do a decent clinical evaluation, my guess is that I, living even in Los Angeles, would probably send any person that I identified a memory impairment knowing the modeling that David did, directly to a PET scanner. Just like we have determined, given the sensitivity and specificity, I think in your own work, Hal, of working up angina, that you probably should send them directly to coronary angiography, and do not stop, pass go, but certainly through these diagnostic tests and everything.

The question that I come up with is, you can't take these procedures in general except those that are really grossly not safe or efficacious, but we haven't seen any of that, we haven't seen any medicine where we're going to cure multiple sclerosis by opening the back of the neck and putting a forceps under the vertebral artery and tweaking it a few times, and closing it up. That's not what we get presented. We get presented material that's much better evaluated than that but not perfectly

evaluated, and how does it fit into this whole process of deciding what package of things are going to be done.

And literally, the fight that exists right now is between this limited budget that Medicare has, how much of it is going to being spent on going back to inefficient physician visits versus relatively ineffective technology or effective technology, how do you put all this together? And one of the things that bothers me on a policy level is that with all the bashing of managed care and the failure of the government to stand up for managed care, we're not left with anything other than going back to fee for service, higher deductibles and co-insurance, and fighting it out in the trenches. And we probably have something to say about this as we think about covering medical technology.

What if the neurologist had to make the

.00233

case in front of us that a neurology visit for AD should be covered? Could we document that anything they do is value added to a good general exam, and I would doubt it. And I would doubt that PET scan would look better than a neurology visit, if you could use cost, but I'm biased. So the question here becomes how are we putting all this together.

And I like your point, Hal, about trying to think about it in broader purposes. I don't know what it would exactly look like, but I suspect that even if we don't approve it, the price of PET scan will come down and some private entrepreneur is going to market it just like they're marketing heart scans for everybody, and Medicare is going to pick up the cost of evaluating the false positives and the true positives, and we won't have a real coverage policy, so we may spend more money on this than if we actually covered it and came up with rational subgroups on who it should be used. But we can't bite that bullet because we can't have federal guidelines or ways of using things other than sort of quote, coverage decisions. So we're really up a creek at the moment as far as I can tell.

DR. SOX: Daisy.

DR. ALFORD-SMITH: I guess the more I hear
.00234
of this discussion, it takes me back to where we started with the original mission statement, and I remember quite vividly in that particular discussion we had thoughts and opinions regarding the comprehensiveness of what we would do, which would be inclusive of various populations. And so when I see Sean putting something up like the clinical impact, for me I think we're beginning to speak to that. I really believe that it's going to address some of those things that I in particular have an interest in, and I, you know, highlight the one, are there specific patient selection criteria that must be considered? And so it really helps us.
And just using one example, we're all very much aware of the health disparities that exist in this country and based upon that we also know that there is a lack of access to a lot of the clinical trials for various populations of individuals. So that when you really begin to at least attempt to address perhaps whether there are specific patient selection criteria that needs to be considered, then it also affords us the opportunity to look at a broader picture of what needs to be done and addressed in various subgroups of individuals.
So, I am particularly impressed with the
.00235
additional questions, but then I also recognize the limitations just in terms of how extensive do we need to be, or how extensive can we really be. But if we go back to that mission, it speaks to that, and these questions address that particular area.
DR. SOX: Other comments? Barbara.
DR. McNEIL: Just one other comment.
Sean, I like your idea of grading the adequacy of the evidence rather than coming out with a black or white yes or no answer, I like the idea of some kind of gradation, much like the professional societies do in their guideline development, so I think that would be a step in the right direction.
DR. TUNIS: I'm curious to know, particularly I guess from the folks who were part of

the methods work group who, you know, developed the interim guidelines, Alan and Hal and others, whether we potentially would lose something by -- that would replace a yes no question, I guess, if we did a grading of evidence, and I just want to make sure that sort of the thoughts around the table of what's the potential risk for doing that.

DR. GARBER: In fact what we had said was, what we lack is a direct A, B, C rating of the evidence, but we rate the effect size, and I don't

.00236

recall all the details of the discussion, but this was an Executive Committee decision that what was important was if you could draw conclusions, and the reason that -- and then you would go to step two of grading the size of the effect, and there were these various categories that we used. As I recall, though, the main reason why we didn't want to rate the evidence in the first step is because we wanted to give the panels discretion in deciding whether it was adequate.

Now this decision of the executive committee wasn't written in stone and it is certainly something we could revisit, but the argument against doing it all with the grading of evidence was we had this algorithm that you go on to the second step if the panel viewed the evidence as adequate in the first step, and we can do the same grading, but it will still be the case that you have to say what threshold do you say the evidence is adequate. And I think one of the advantages or disadvantages, but one of the characteristics of doing it without an explicit rating is it gave the panels the flexibility to consider whether -- for example in some areas it may be almost impossible to do a randomized controlled clinical trial, and so they would have

.00237

some discretion to use the lower standard of evidence without making it as explicit as it would be with a grading scheme.

Again, you know, I don't think any of us felt super strongly that it had to be the way that we have done it. I think it has worked pretty well the

way we have done it, but certainly if people think it's not adequate for decision making purposes, we can change it.
DR. SOX: Here's our customer. What do you think, Sean?
DR. TUNIS: I think getting this feedback is useful, because not of all us were part of the entire history of this, and I think that in some cases we found that yes-no questions on the adequacy of evidence is thought to be so colinear or coequal with a decision about is something reasonable and necessary and therefore covered, that the MCAC saying yes on that question, even when the evidence has lots of significant weaknesses can in fact be something of a hurdle for CMS in that it looks like the MCAC has made a definitive, essentially said the equivalent of yes on the coverage decision just simply based on saying yes on the adequacy of evidence. And some of the panels have not wanted to say no because they
.00238
didn't feel like the evidence was convincingly inadequate, and so you ended up in a place where you had to go one place or the other. So in some of those cases, as you're well aware, where the coverage decision and the panel seemed to be at odds, it really I think centered around this issue of the difference between how good the evidence was and the yes-no decision on the part of the panel.
DR. GARBER: Well, in the reports I am familiar with, there has been a descriptive statement about the quality of the evidence immediately after that decision about whether the evidence is adequate, and so what we're really talking about is moving it out of the explanatory text into some bold faced statement that it's Grade B or something, and I really don't have a strong opinion about whether that's important to do or not, but I would certainly, I think that the panel discussions and the Executive Committee discussions have been nuanced enough that it would be unfortunate if the readers only looked at those highlighted points.
DR. SOX: You know, anything that we adopt, one of the reasons to do it is to improve our

external validity or credibility to the folks whose ox may be gored by the decisions we make, or who may benefit, and do you have an opinion about whether being a little more explicit about quality of evidence would be helpful in that respect?

DR. TUNIS: I think something along the lines of what Alan said, you know, of just bolding and highlighting the qualitative assessment of the evidence in some way so that it's clear that you know, the yes-no is matched with, and here's what these are, just on the scientific evidence front, what the quality is, and strengths and weaknesses, and I think that would be sufficient, and I think it would in many cases help us to maybe formally amplify it and maybe potentially formally have a vote on a simple grading system.

So, I think we don't have to decide it today, but I think this is useful input for us and we may want to, with this committee's approval, reconstitute the methods subgroup and try to come to some consensus to bring back to the committee about this stuff.

DR. SOX: Sounds good.

DR. TUNIS: Let me just do this. Jeff is here, he's going to come up in a minute. Just as a final bit of business, there is another panel meeting scheduled for I believe it's June 12th. It's the med-surg panel and the topic is bilateral deep brain stimulation for Parkinson's disease.

Because we are still under the old operating policy and we have announced a panel meeting on this topic, the executive committee will be given the opportunity yet again to convene to ratify that panel's recommendation. We are hoping that prior to that meeting and prior to announcing any future panel meetings that we will have the charter revised in such a way as to no longer require an executive committee ratification, and we are moving along well on that front, and I will be able to talk more details very soon. But their will be one more executive committee meeting, probably in the

early fall, to talk about deep brain stimulation about Parkinson's, so start reading up.
DR. BROOK: Is this to imply that there will be no other executive committee meetings after this, or do we not know, after we deal with deep brain stimulation?
DR. TUNIS: The function of the executive committee will be maintained in spirit if not in fact. We are moving to being able to convene a committee to deal with the broader issues of the MCAC but it will probably not have the title of the

.00241

executive committee, but all of your memberships are intended to be maintained, and roles. It's kind of in the final stages of clearance, so I just can't give a lot more details about it, but your name will be on a list somewhere.
DR. KIANG: I actually think that one good way of thinking through the very discussion that you had this afternoon about the role of decision analysis as kind of a policy issue, I think is a great thing for this kind of new committee to do and advise us on, and I really think it sets the tone for the panels' work. So I actually, kind of separate from how we organize this, I do think that there is needed desperately a role in kind of the larger policy area, and someone needs to do that, and somewhere we need to get kind of advice on, and I would see that for whatever we call this new committee.
But I'm sorry I missed some of the discussion you had, but the very discussion you had today is very important, I think, about decision analysis and the role in coverage decisions, and I think we need to continue to do that.
DR. TUNIS: That's all the business and Jeff, did you want to have a few words.

.00242

DR. KIANG: I wanted to actually say that some of you may have heard already, but I have taken a new job as the senior vice president and national medical director for Sigma, up in Hartford, Connecticut, it basically is their entire health

insurance portfolio, so I will be actually leaving and my leave date is May 3rd. But I did want to say that from kind of my perspective and also from the Administrator's perspective, I think that we as an agency have been really pleased with all of your hard work and I think you have set a very important tone and I don't want this issue of kind of reorganizing and everything, I do see a role for this group on a continued basis more in the policy making area, so just bear with us as we sort through some of the legal details, but the message that I wanted to send is you have done a fabulous job and I really think that, certainly I do, but also, I think the Administrator definitely appreciates what you have done here, and we are looking forward to more interactions, and maybe I will have some interactions in my new role at Sigma.

DR. SOX: Perhaps I could say speaking only for myself, Jeff, that one of the reasons that we have been for our success such as it has been is

.00243

that we have been working for an inspired leader, namely yourself.

DR. KIANG: Sean is inspired; I am more deranged.

(Laughter.)

DR. SOX: Your boldness, your sort of anti-bureaucratic approach really have been inspiring to all of us, and I think it has given us confidence in our government and CMS to know that people like you are important decision makers in the government, and speaking for myself, I'm delighted for you that you're having a new opportunity, and I'm sure it's the right thing to make a change for you, but we're disappointed not to be able to be seeing you regularly and we want you to know now highly we value your help, your esteem, your inspiration. Good luck.

(Applause.)

MS. ANDERSON: There are a few closing things. First of all, I do want to mention that I know that most of you have to go to your respective airplanes, trains, cars, whatnot, but we have been invited to the conference down the hall, and that is

extended to our gigantic audience, I don't think they would mind if you show up. You simply have to go to the registration desk, they have offered us

.00244

complimentary registration, they have lovely exhibits, and they are having a meet and greet at 4:30. I think at 3:30 they're actually having a discussion on who's going to pay for health care. Now, for continuing information, visit our web site at www.cms.hhs.gov\coverage, or simply go to the CMS web site and click on the coverage process. To conclude today's session, would someone please move that this meeting be adjourned?

DR. ALFORD-SMITH: So move.

DR. HOLOHAN: Second.

MS. ANDERSON: Thanks, you all. This meeting is adjourned.

(Whereupon, the meeting adjourned at 3:25 p.m.)