Transcript of September 25, 2002 Meeting

Note: The language on this website comes directly from the transcribed testimony taken at this panel meeting. The views and opinions are those of each of the experts and not those of the Centers for Medicare & Medicaid Services. CMS does not edit these transcripts and makes no assertion as to their accuracy.

Baltimore Convention Center
100 W. Pratt Street
Baltimore, Maryland

Attendees

Harold C. Sox, M.D.
Chairperson

Robert H. Brook, M.D., Sc.D.
Vice Chair (voting)

Janet Anderson
Executive Secretary

Voting Members
Leslie P. Francis, M.D., J.D.
John H. Ferguson, M.D.
Alan M. Garber, M.D., Ph.D.
Thomas Holohan, M.D.
Barbara J. McNeil, M.D.
Wade Aubry, M.D
. Ronald Davis, M.D.
Michael Maves, M.D., M.B.A.
Robert Murray, M.D.

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

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

Consumer Representative

Linda Bergthold, Ph.D.

P R O C E E D I N G S
(8:13 a.m.)

Ms. Anderson: Good morning, panel chairman, members of the Executive Committee, and guests. The committee is here this morning to discuss and act upon recommendations from the MCAC Medical and Surgical Procedures Panel regarding the use of deep brain stimulation for the treatment of Parkinsons's Disease.

The committee will also discuss methodologies for assessing evidence that could be incorporated into future MCAC reviews and referral criteria for issues that come to MCAC.

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 meeting and is made part of the record to preclude even the appearance of impropriety. The conflict of interests statutes prohibits 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 matter before the committee today.

With respect to other participants, we ask that 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 topic to be presented today. Sean?

Dr. Tunis: Thanks, Janet. I just wanted to welcome everybody and thank you again for taking the time and making the trip here for this meeting. I think we have today a number of critical issues to talk about, not only the deep brain stimulation in which there's a great amount of interest, but a couple of the issues of process for the MCAC and also for the coverage function in Medicare are quite crucial. This issue of delineating factors and considerations that go into obtaining external technology assessments for specific issues and also factors and criteria for deciding which issues to refer to this panel are quite fundamental to our functioning, and I'm looking forward to that discussion.

And then the afternoon, we hopefully can get a better grip on an issue that we've been struggling with, I

think, for a good bit of time, which is some of the situations under which we would find adequate less than direct scientific evidence on questions of effectiveness or clinical utility. The issue is brought up to us particularly by continuing considerations of diagnostic technologies in oncology, particularly PET scanning, and particularly in the context of rare cancers or infrequent indications within cancers that come to our attention.

Just a couple of quick announcements or sort of updates on some other issues that affect this committee. As some of you may know, all the Department of Health and Human Services advisory committees are under continual review, and there has been an effort over the last year to look at each of the numerous committees and kind of assess where they are and any modifications that might be needed in terms of numbers or charters or things. We're part of that pool around which there is a significant amount of discussion, so we're continuing to move a revised charter for this committee with some of the features I've talked to you about at previous meetings. Hopefully we will have some finality and a signature on that in the near future, and hopefully that would mean in the next couple of months.

We're also moving forward on some additional members for the committee and hopefully those will be made public in the near future as well.

So, I don't expect a dramatic earth shaking change in this committee beyond, again, what we've talked about at previous meetings, particularly in terms of eliminating the two stages of panel recommendation followed by EC recommendations, which I think would be a major improvement if we can get there.

So with that, let me turn it over to Dr. Sox, unless there's any particular questions. Otherwise, we can move on to the business of the day.

Dr. Sox: Well, I'd like to start out by asking the members of the panel to introduce themselves by giving their name and their job or affiliation, and then if you have any conflicts related to any of the topics that we're discussing today, I'd like you to disclose them now. Mike, will you start please?

Dr. Maves: Sure. Mike Maves. I'm the EVP-CEO of the American Medical Association, and no conflicts.

Dr. Brook: Robert Brook. I'm with Rand Health and UCLA, and have no conflicts.

Dr. Ferguson: John Ferguson. I'm a neurologist and consultant to the NIH Office of Rare Diseases, and I have no conflicts.

Dr. McNeil: Barbara McNeil, Harvard Medical School and the Brigham and Women's Hospital. No conflicts.

Dr. Holohan: Tom Holohan, Chief of Patient Care Services in the Veteran's Health Administration. No conflicts.

Dr. Francis: Leslie Francis. I'm professor of philosophy and law at the University of Utah. No conflicts.

Dr. Garber: Alan Garber, VA Palo Alto Healthcare System and Stanford. No conflicts.

Ms. Richner: Randel Richner, Boston Scientific. Conflicts?

Ms. Anderson: Actually as the industry rep, you're excluded from that, Randel, so you're okay.

Dr. Murray: Robert Murray, Advocate Healthcare in Chicago. No conflicts.

Dr. Aubry: Wade Aubry, University of California San Francisco, no conflicts.

Ms. Bergthold: Linda Bergthold, Center for Health Policy, Stanford University. No conflicts.

Dr. Sox: I'm Hal Sox. I'm the editor of the Annals of Internal Medicine, and no conflicts by definition of my job.

I will make my remarks fairly brief. I think that in the greater scheme of things, the biggest contribution of this panel has been to establish processes and techniques for evaluating the evidence so that we can carry out our statutorily mandated mission, which is to advise the coverage group. And I think we have actually come quite a long way in the several years that we have been functioning, and I think today's presentation about deep brain stimulation will be kind of a good example of process in action to try to give advice that is evidence based. This afternoon we will be dealing with a really tough problem, which is how to give advice when the evidence is inadequate for one reason or another. Don't wear yourself out in the morning because you really want to be at your best for the afternoon, because I think that's where we're serving our country best, is in trying to give Medicare good advice about how to find out what the best evidence is and make coverage decisions that are based on it.

So, with that brief introduction, I will ask Ron Davis to state his name, his job, and if he has any conflicts on the matters before us today.

Dr. Davis: Sorry I'm late. I'm Ron Davis, with the Henry Ford Health System in Detroit, and I don't have any conflicts.

Dr. Sox: Well, with that we're ready to proceed, and Alan, do you want to take it from here?

Dr. Garber: Yes, thank you, Hal. I'm reporting on the results of the Medical and Surgical Procedures Panel meeting that addressed deep brain stimulation for the treatment of Parkinson's disease. I think everybody received two evidence reports, one is thick and one is thin. The thick one is for bilateral stimulation, the thin one is for unilateral stimulation, and I'll just say a few words about the questions that we were asked to consider and how the panel discussion went.

I think everybody received this two-page handout with the voting questions, is that right? I think it was sent as an e-mail attachment. So the first question was, is the evidence adequate to determine the clinical effectiveness of bilateral subthalamic nucleus deep brain stimulation, and rather than attempt to say that again I will use these acronyms STN and DBS, for a well-defined set of Medicare patients with Parkinson's disease? The first question then under that is, if the evidence is adequate, what is the size, if any, of the overall health effect? Now let me just say a word about the study designs. There were actually, I believe, no large studies that had randomized double blinded control trials where people were randomized to either get the device implanted or medical management. But the study design that some of the studies used was one where they tracked symptoms when the stimulators were in place and they were turned off for periods of time and turned on, so you could track response that way.

So the panel felt that that was actually a fairly persuasive type of study design, one from which you could draw some conclusions. The totality of studies was very positive. This was done in people with fairly severe Parkinson's disease, and the belief then was that yes, they really did seem to be of adequate design to draw conclusions.

There is a second aspect to the discussion, and that was, there were two subquestions under deep brain stimulation. Should it be placed in the globus pallidus, or GPI, or should it be placed in the subthalamic nucleus or STN? And here we had some really terrific help from ad hoc members of the panel who were drawn from other panels, including a neurosurgeon, Kim Burcheil, who actually provided a lot of insight as a neurosurgeon. And he pointed out that this is a decision, the location to put the leads was a decision that a neurosurgeon often made on the spot and it would be very difficult to determine which location was better, and the studies weren't really set up to answer that question, which location was better. So he pretty much persuaded the panel that it was not going to be a useful exercise to try and establish indications separately for those two sites of implantation.

So that was the nature of the discussion.

So on the first voting question, is the evidence adequate, the panel voted a unanimous yes, and they designated the studies as of Level II quality. And even though that falls short of the ideal, they felt for the reasons that I have mentioned that that was adequate to draw conclusions. So that was a unanimous yes. Then if the evidence is adequate, what is the size, if any, of the overall health effect of bilateral STN DBS for those patients? And this is something I would really like to draw the Executive Committee's attention to because this may come up in other contexts, and it's something that may be, I think that the panel addressed this question very thoughtfully. They believed it was clear that this device for this indication was effective. And then when they struggled with which category to put it in, they were faced with the top category, which is described as a breakthrough, and some of the language in breakthrough is so effective it becomes the standard of care.

The second category was more effective, and I can't find the exact language of more effective - oh, here it is. Let me just read this: The new intervention improves health outcomes by a significant albeit small margin, as compared with established services or medical items. They thought this was bit too tepid to

describe the magnitude of effectiveness of this procedure.

So we did have a great deal of discussion about creating another category, and so this would be intermediate between Categories I and II, and the term that the panel used was substantially more effective, and the language under substantially more effective is as follows: The new intervention improves health outcomes by a substantial margin as compared with established services or medical items. So the panel felt very comfortable with this category in general, and for this particular technology and indication, so they unanimously approved assigning deep brain stimulation to this category.

So that was actually the biggest discussion. Then the issue was clinical effectiveness of bilateral internal globus palladus GPI DBS, and here the discussion was pretty much the same as what I just described for STN, and they voted unanimously that the evidence was adequate for that, and again they assigned it to the category of substantially more effective.

But the way you should interpret the panel's discussion is they were really lumping GTI and STN as sites together. Okay.

And then they modified the third major category of questions, is the evidence adequate to determine the clinical effectiveness of ventral intermediate nucleus, VIN DBS for central tremor and/or Parkinsonian tremor for a well-defined set of Medicare patients with Parkinson's disease? So that differs somewhat from the questions that CMS put to us.

For this one, it basically, there was a discussion, and this should have been for unilateral stimulation basically I believe, if I recall the discussion correctly. And here we relied on this thin document, which is mainly a set of tables. And the tables, most of the panel felt were adequate to summarize the evidence to their satisfaction, and concluded that not only was the evidence adequate but they had enough material to judge it.

There were two panelists who felt they really needed to see a more detailed report. And let me be clear. They weren't questioning really whether there was enough evidence. They were just questioning whether they had enough information in this format of the tables alone to really be able to draw conclusions about how good the studies were, and so on.

So two people abstained, and it was really without prejudice as to the issue of whether the evidence was adequate. It was more a comment on did they have information that they were able to assimilate and use effectively. So apart from those two abstentions, all of the panelists felt that the evidence was adequate here, and again, when they voted on the level of effectiveness, they said substantially more effective, and the same two individuals abstained, obviously because if they couldn't judge whether the evidence was adequate, they didn't feel comfortable assigning it to a category.

So in summary, they were all yes with the addition of this new category, substantially more effective. I personally thought that the addition of this category was very useful. I could imagine that we will be

faced with other types of technologies where this would be the right category to use. The breakthrough technology, at least as it's defined now, is a fairly high bar to reach and I think that something that really works well might not really be the standard of care in all cases, and they would fit well into this category.

Dr. Sox: I would like to suggest that we discuss this new category not now, because we really need to focus on the technology before us, but come back and discuss it later on and perhaps make an addition.

My suggestion now is that we, if anybody has any sort of questions of fact that they would like to ask Alan, this would be a good time to do it, and distinguishing questions from sort of assertions and so forth, which we could deal with when we have the scheduled discussion time. So, any questions for Alan?

Actually, I would like to ask one, Alan. I didn't see much about patient preferences for the various outcome states that they were in, either before the treatment or after the treatment, or any particular effort to try to balance harms and benefits. I wonder if you could say, tell us what you can about that.

Dr. Garber: Yeah, that's a very good point, and I should say that one of the things the panel did discuss is who should be performing this procedure, and there is some point on the learning curve. The panel wasn't interested in getting very precise about recommendations like how many operations per year a neurosurgeon should be doing to really be considered qualified to perform this.

But clearly there was a belief based upon some data that the side effects of the procedure were dependent upon the skill of the surgeon, and the population chose - the study samples basically were people who had refractory disease. And the studies did focus not on preferences in the sense of utilities and types of quality of life instruments that we might use in other settings, but they did extensive analyses of symptom severity, frequency, on-off phenomena and so one. And the evidence is really quite consistent that it was an improvement, and it seemed that the side effect profile was fairly low if done by a skilled center. So these are people who were pretty desperate, it sounded like, in many of these trials.

So there really wasn't really much doubt that they felt better, even though it might not have been measured with the precision we would like. So it turned out that this balance of harms and benefits, we felt the studies had really addressed to our satisfaction. That is, for a person who had severe refractory Parkinson's disease and some of the other indications, they were often immobilized by their disease and so it seemed that it would be reasonable for a person to draw the conclusion, I'd rather undergo the procedure than continue with medical management.

Dr. Sox: Thank you. Other questions? Yes, Leslie?

Dr. Francis: I didn't see your presentation or the minutes addressing specifically the question of early versus later, and which patient population, and basically what's meant by well selected.

Dr. Garber: Well, thank you, Leslie. That was - I'm sorry that the record didn't reflect that, and maybe

that's because the discussion was not entirely conclusive. There was a question about at what point you might initiate this therapy, and it came up in the context of Medicare beneficiaries, because these would oftentimes be people who had onset of their disease maybe in middle age or 40s and 50s, and then would it be Medicare beneficiaries in the studies. And the treatment, the age of treatment for many of these people was people in their 60s and so on, so we felt that it extended to the Medicare population.

I don't think we satisfactorily resolved the question of what is the best set of symptoms in time and age and so on at which to consider this therapy, but it was pretty clear if you look at the studies, there is a wide range of ages represented. The criteria to be entered into a study were reasonably well laid out but maybe not with the precision we would like, but generally it's people with severe Parkinson's disease. And if the question is, is there a point at which they became too burned out to benefit, that's something that we could not really resolve based on the information at hand.

Sean or some of the other people here who were here at the meeting might be able to comment more.

Dr. Tunis: Actually I wasn't at the meeting.

Dr. Garber: Oh, that's rght.

Dr. Sox: Tom.

Dr. Brook: Most of the published studies that the panel reviewed were composed of patients who at that point in there disease were more or less refractory to drug therapy. These were patients where there didn't appear, certainly in the eyes of the people doing the study, to be anything further gained from drug manipulation. I don't want to get into the technical details of on versus off periods but for the most part these were patients who would have unexpected periods of akinesia, and were not able to be treated successfully to any significant extent by drug therapy. So they were kind of in the final common path end stage of most Parkinson's disease patients. None of them were subjected to surgery, none of the patients reported in the literature were subjected to surgery early in the course of their disease when they were drug responsive.

Dr. Sox: Ron.

Dr. Davis: I think Tom might have just provided the answer to the question I had. The evidence report, the larger one on page 4 notes that, "There are no large prospective randomized studies with long-term follow-up of bilateral DBS for treatment of advanced Parkinson's disease." And I was going to ask, well, why is that the case? And maybe Tom just answered that; is it because this is seen as a therapy for people who are already refractory to benefit from medication, so that there wouldn't be a medical management comparison group that would be appropriate for the deep brain stimulation?

Dr. Garber: Yes. And that's why the panel was persuaded by this alternative study design of turning the device on and off. So they have the leads implanted but the machine isn't firing. So for placebo effect,

you would expect the benefits to continue during periods when the machine was turned off.

Dr. Davis: And presumably, the patients were blinded to whether it was turned on or off?

Dr. Garber: Right. And I think Tom may remember this better than I do, but people seemed to believe that it's likely the patients didn't really know when it was on, but it wasn't totally clear.

Dr. Davis: Thank you.

Dr. Sox: Although some patients apparently reported kind of a little bit of a funny feeling when it was first turned on and it was sort of transient, so they may not have been -

Dr. Garber: Yeah, that was part of the issue. But apparently it was subtle enough that patients might not have known.

Dr. Sox: And were you pretty convinced that the observers, the people who were doing the grading, were blinded adequately?

Dr. Garber: Well, I would not say that I felt they had a detailed enough knowledge to go beyond what was in the study, what was reported. And so you know, there's always caveats but I didn't have any reason to think that they were not being objective and really knew how they were being treated, but I wouldn't hang my hat on that.

Dr. Sox: Okay. Bob.

Dr. Brook: Yeah. I just want to ask a question of fact. In your first major recommendation you have the words "for a well-defined set of Medicare patients." Nowhere in the technical report that I read by whoever did it, did I understand what well defined meant. And after I read your minutes and recommendations, is it the intent of the committee that a well-defined set should be stricken from that and it just means anyone with Parkinson's disease done by anybody? Even though there were caveats in the report, nobody - this doesn't say that a group of patients that had failed drug therapy, had the disease for a number of years, have more off periods than on periods. Nor does it say in this recommendation that it needs to be done by somebody in a center of excellence that has a set of side effects that are below a certain level. All of those issues were raised in the technical report, and I really don't understand. I understand the vote but I have no idea what you were voting on, by reading this thing. So the way I would interpret it right now is that anybody, that the committee believed that anybody that any surgeon or doctor believed ought to have this thing done, as long as they had the diagnosis of Parkinson's disease, it would be covered. That's what I understand the committee did and I just want to make sure that that's what you did.

Dr. Garber: No, that's not what the committee did. And the committee - let me explain where the questions came from. The questions in their original phrasing came from CMS, which in turn was

following the Executive Committee's guidance about how the questions should be phrased. So this is the question the way it would be crafted if you started with - and I think that CMS did this absolutely correctly - if you start with the Executive Committee's recommendations, and the terminology "well-defined set of Medicare patients" comes from us. It does not come from the panel or from CMS staff.

So the issue here was can you find a set of Medicare beneficiaries for which there is adequate evidence to draw a determination, and that was what a lot of the discussion was about age and when you start treating, and so on. And the panel felt it was quite conclusive that there was a subset of Medicare beneficiaries who were also a subset of the randomized trials and that you could draw conclusions about that subset.

Now Bob is raising a very valid point that well, how well defined is this really for clinical purposes. That is, who should get it, who should perform the operation, and so on. We in consultation with CMS did not get into great great detail about all the parameters that should guide who should be doing it and who should be getting it, and we had an extensive discussion about what makes a surgeon qualified. My understanding was that CMS thought we provided adequate general guidance for them to work with about level of skill and so on.

So I think that we could have a discussion about whether we the Executive Committee want to direct the questions to be different, but the panel and CMS were faithfully adhering to recommendation that we had made in setting up the questions that way. Tom, you might want to add to that.

Dr. Brook: Can I just respond? Alan, I don't think you answered the question. I agree that you did all that in the committee discussion. I mean, I read all the minutes, I read all the materials. But the way that this is phrased, it's a matter of fact that there is nothing in this summary document, nor is there any reference to any other document that indicates what well defined means.

And then in the second recommendation when you talk about a breakthrough series, you say for those patients, and there's no definition - this is just a matter of fact - I have not been able to find anywhere a definition of what those means. And in the absence of doing that, I don't know how to interpret - I know what the discussion was, I commend the committee for the discussion, but I can't relate these recommendations to any group of patients other than Parkinson's disease, because there is nothing beyond what you have here, at least that I've seen, unless there is something that goes with this that would indicate what those mean and what well defined means. And I was just asking as a matter of fact, is there a definition of well defined, and is there a definition of those? And if there is, I'd like to know what it is.

Dr. Sox: So, can somebody please answer Bob's question?

Dr. Garber: You know, there's two aspects to Bob's question, and I will have to come back to the question, is there a subset of Medicare patients? The second is, is there a set of clinical indications that's very precise about who should receive this? And I would have to differ with Bob about the second

aspect, which is what I think he is really pushing. I think he is looking for a definition of, say, people who have had symptoms for X number of years, refractory treatment and so on. The evidence report makes it quite clear, as Tom has stated that this is people who had refractory disease unresponsive to medical therapy. It's very hard to go beyond that. I think the transcript will reflect that's what the panel had discussed and concluded that these patients would benefit. There were some aspects that they could not get more precise about, like duration of treatment. If you press very had, what you mean by refractory in very precise terms, that you have to be guided by studies on.

So I get back to the issue. If the question is, should we have very very well-defined clinical criteria to decide for whom the evidence is adequate, beyond the issue of can you define a subset, and basically in most of these cases we talk about people like the people enrolled in the studies and a general description of who is enrolled in the studies, and the discussion reflects the fact that these were people with refractory disease.

So the panel didn't feel that this was a great point of contention, although they recognized that there were some details, especially about who's qualified to perform the operation, which is a very important issue, that they couldn't give precise guidance about it, and CMS didn't seem to think they wanted it.

Dr. Brook: Alan, then can I reach - it's the intent, and I'm asking you a fact question. Is the intent of the first recommendation and the second, and the second part of the first recommendation in the meanings, that by well defined, you mean the panel considered this to be that the evidence was available for a set of Medicare patients who have been shown to be refractory to the best medical therapy for Parkinson's disease. Is that what you mean?

And then for people who are refractory to medical therapy, that's what you mean by those, that this procedure produces substantial effectiveness. That's all I'm asking.

Dr. Garber: Yes. I would strike the word best, but yes, that's correct.

Dr. Brook: Well, that was done in the trials, so trials usually produce the best therapy, so should we or can we replace, as a matter of fact, these vague terms with something that we now understand, or are you in terms of what's going on? And the reason I bring this up is because we have been working on epilepsy surgery and have exactly the same set of questions about what's going on, and I just didn't know what the panel's intent was. I understand what was the evidence in there and I understand what you reviewed, but there seemed to be a disconnect between these questions and the panel. And that's not discussions, it's a matter of fact. I just want to know how I ought to interpret this.

Dr. Sox: Our job, of course, is to advise the coverage group and to provide them with as much precision as they need in order to do their job. So Sean, I guess I'd like to ask you whether the coverage group would be served by having a more precise definition along the lines of what Bob is asking for and Alan was in fact saying.

Dr. Tunis: Well I do think as -

Dr. Sox: Because if you did, then we could change the wording of the motion that we eventually vote on.

Dr. Tunis: Right. So I do think, you know, as far as you can go along these lines of being precise about the language of the patient selection criteria would be helpful. Whether that is, you know, the patients that are refractory to best medical therapy, if that's the language that is consistent with the spirit of the panel's recommendations. Or I guess another alternative is patient populations essentially meeting the selection criteria for the major trials, I suppose those may be similar, but that may be similar to refractory patients but not identical. But I think, you know, as far as you could push on this would be helpful, because obviously this is clearly going to be addressed in the coverage decision and you know, we take from the panel's recommendation that we ought to expressly address this in the coverage decision, so I think the greater the detail that can be provided, the better.

And I think that applies also to Bob's second point, which is you know, if this is to be limited to certain centers or practitioners with specific characteristics, we would also want you to go as far as you can go on that. I'm not sure how far the panel went, since I didn't hear that discussion, but I think both of those issues would be important to press on.

Dr. Garber: Well, I think these are excellent points and let me say that I personally, and I believe the rest of the Medical and Surgical Procedures Panel would like to provide as much precision as possible. And let me just point out some of the challenges.

First of all, to say that it corresponds to the people involved in the trials is something that's implicit every time we make a recommendation. I'm not sure, if you think it's helpful to make it explicit, fine, but everybody who looks at it will know that our statements are made to imply to people like those in the studies we looked at.

So the issue becomes, can you define subsets, recognizing the messiness that the trials don't use identical patient populations and they typically don't stratify their results. So when you think about does this work, is there evidence in a specific age group, is there evidence in a specific indication, what do we mean by using the term best medical management? We usually in many of the studies don't have that much information about what the medical management was. And as Bob said, we would expect it to be the best if there are people involved in the trials. We don't know that as a fact. That's a statement of faith.

So you talk about what are they refractory to. Well, we usually aren't very sure and even when studies specify it, they may have used, in fact almost always will use varying criteria for what constitutes both prior treatment and treatment failure. What's a good enough attempt, how many different medications have people been tried on? And the list goes on and on. So there's a balancing act here and we have to be guided by CMS staff about what type of response from the panel is going to be most useful to you.

And I think that one of the most important issues for this particular technology is skill of the surgeon and

the center in which it's performed, and virtually all the useful information we had about that came from our expert Kim Burcheil. So, do you want to enshrine that in a report, recognizing that's the nature of the information, which I think was quite valid and helpful, very useful, but it's not quite the same as saying we know that centers that did 20 of these operations a year had this complication rate, centers that did 50 or 100 had this complication rate, and so on and so forth. We simply didn't have that kind of information.

So I think that's very important information to have, but panels can't produce that out of the air though, and unless that's part of the evidence reports, it's not very useful to have the panels try to address that in any great detail. And I think the reason that it's not in the evidence reports is because there isn't much basis for anyone studying this question for this particular technology to draw firm conclusions.

Dr. Sox: If I could just comment first, this has been really a useful discussion because it's in effect raised another potential voting question, which is, is the evidence adequate to specify any contingencies or conditions to be placed on the recommendation, such as eligible patient population or skill of operator and so forth. So it sounds to me in this case, you would probably resist being pinned down and rightly so from what you said, but it sounds to me as if it's an issue that we need to address more explicitly in the future than we have in the past. And for that reason, the discussion has been very useful.

So, I guess John is next and then Randel.

Dr. Ferguson: The issue of where these things should be done has come up for centuries, decades anyway. I mean, it came up with heart transplants and so on, and CMS has always as far as I know been able to in their coverage decisions say that these things need to be restricted to places of excellence, defined as 12 procedures per year, 25 procedures a year or whatever. That's not the job of this committee because the evidence is saying that these things should only be done at Mass General and Johns Hopkins, or something like that, is minimal. And even if this group gets down to saying that, it's restriction of trade.

Ms. Richner: I have one question also about the definition of the Medicare patient population. In this particular disease, I notice from the studies that the mean age was less than 65, and Parkinson's certainly affects a broad range of patients. Now they can have Medicare disability and be a younger population but I would really caution as well restricting the analysis to those 65 and older if that was the intent.

So you have to be very careful as well in terms of how you define the Medicare population in this particular disease.

Dr. Garber: Randel also brought up, if I could just make a point of information, I'm very happy Randel brought that up because that was part of our discussion. It turns out that the panelists, the neurologists who were on the panel and the neurosurgeons believed that a very high percentage of the people under age 65 would be eligible for Medicare on the basis of disability. But we didn't have the information from the studies to be able to say that with certainty. So that entered into the discussion but we really did believe that a lot of the under 65s would be Medicare beneficiaries.

Dr. Sox: Well, I think we need to close this discussion so that we can stay on schedule, but Bob, do you have a discussion closing issue to raise?

Dr. Brook: I have one quick point to raise. I don't understand either Alan, your comment at all, because I read the evidence report, I didn't think there was anybody enrolled in any of this clinical series that had never been tried on medical therapy for Parkinson's disease. So I couldn't find any evidence in the evidence report as a matter of fact that basically applied to people that weren't at least tried and failed medical therapy. And if the second part of this labels this as sort of a substantial breakthrough, then there are two issues that the panel talked about but are not captured, as a matter of fact, in these minutes. One is it's a substantial breakthrough only if the side effects approach a certain level, which the committee discussed but didn't incorporate into its recommendation. And two, it only refers to patients, at this moment the evidence, that basically have been treated presumably in a competent way with medical care and with standard drug therapy. And the fact that we can't define it precisely doesn't mean that we go the other route, and I'm concerned because if this report comes out, I'm concerned that the first thing every patient with Parkinson's disease, maybe they should get it, a stimulator and no drug therapy. But the evidence here that you reviewed and what you just told us doesn't support that conclusion. So I'm concerned about why the panel didn't say that as part of its recommendations.

Dr. Sox: Do you want to respond?

Dr. Garber: Well, Bob, I lost you in a few places. I think you had a double or triple negative along the way. But first of all, if you were saying, it is correct that this was primarily about people with severe disease who were unresponsive to medical therapy, and at least virtually and maybe all the studies include only people who had been tried on treatment and by some definition had failed. So I'm not sure there is a disagreement there. This is not about initial treatment at all. And if your point is that we should say that this is people who have failed medical therapy, that that's who this applies to and for whom the benefits exceed the risk, that's true. It was in the discussion explicitly and perhaps the minutes should be expanded to include some of the material from the transcript to reflect that. I don't think that contradicts anything about how the panel voted. It's really explanation about what we mean by well-defined set of Medicare beneficiaries. So I'll take your comment of suggestion that the minutes should reflect more fully the discussion to define the patient groups to whom the vote applies.

Dr. Sox: We're going to close this discussion but give Sean the last word.

Dr. Tunis: I just wanted to address a couple of the points that John Ferguson made as well. One is that in terms of center selection criteria in Medicare coverage decisions, those are unique to transplant procedures and haven't been used in anything other than transplant procedures, and it has to do or is driven fundamentally by the limitation in organ supply. And so for other procedures such as this one, it wouldn't be an issue and we would not traditionally set criteria, numbers of procedures, etc.

That being said, I would in my own view think that it is within the charge of this committee that if there is evidence that outcomes of a particular procedure are quite sensitive to either volume of procedures or

the skill of the practitioner and that the harms potentially could exceed the benefits in hands other than skilled practitioners, and there's direct empirical evidence to suggest that that's true, or strong testimony or something else that the committee finds credible, that ought to be part of the advice and recommendations that you give to us. So, I would just make that observation.

Dr. Sox: I might also be part of the report, to say when the evidence was inadequate to address that question, since it's always an important question. I'm going to turn things briefly over to Janet to get us launched on the scheduled public comments.

Ms. Anderson: We've actually only got one comment. I have been asked by the initiator of the request for Medicare coverage for this issue to read the following statement for the record.

I was the requestor for this issue of Medicare funding. I fully appreciate your taking the final step for certification. Many Parkinsonians will now have the coverage they have been patiently waiting for these past few years. Their quality of life will unquestionably change for the better. I can testify to that fact, as I am a DBS graduate and my quality of life has dramatically improved.

Thus, I fully support your certifying this coverage at this time. Hopefully, it will allow those waiting an earlier operation date, for my only regret is that it has been too long a wait for some Parkinsonians.

It is signed by Dr. Barry Green, and dated September 13th.

That concludes the scheduled public comments.

Dr. Sox: Well, now we can move on to open public comments. Could I just ask anybody who wants to make a comment to raise their hand. Nobody raised their hand, so we won't have any open public comments and we will go right on to discussion and vote concerning the panel findings.

So I guess the first thing to do is put something on the table to discuss. Would somebody like to make a motion so that we can move ahead? Ron, would you make a motion?

Dr. Davis: Well, I guess just for the sake of getting the matter on the table for official consideration, I move that we approve the panel's recommendations.

Dr. Maves: Second.

Dr. Sox: Thanks, Mike. Now the panel's recommendations are open for discussion. Who would like to raise any questions? John.

Dr. Ferguson: To address Bob's concerns and this discussion that preceded this, would it - I'm going to go back to these questions again, Sean, excuse me, about - but using the consensus programs we didn't want the questions to be changed because they were crafted so cleverly. But for substituting well defined

some other verbiage that you can decide on, and then to answer that, the panelists believe the studies were for this group of patients, was a unanimous yes, or something like that. I mean, Bob, would that address your concerns, or partly?

Dr. Brook: Well, I'm troubled, especially since Alan tells me that we were responsible as the Executive Committee for these questions. I am deeply troubled by the disconnect between the minutes and what I think is the intent of the committee. I really believe that there is a process missing here that is, that could have given CMS a lot more guidance regarding where the evidence actually supports the use of this procedure at this moment in time, and it's not reflected in the document. My hope would be that somebody prepares, like you do with legislation, sort of an intent of the committee of what they really were talking about here, and what the evidence really looks like.

I also am shocked about the notion there are only a few - I mean, I though the evidence report, by the way, was very difficult to follow, and it was hard to relate the evidence report to what was going on with the committee's discussion, in writing at least. But the overwhelming message is, very few people have been included in these studies and some of them have been duplicably counted and all these other kinds of questions. I'm not questioning that it seems to work, and I'm not questioning that it helps many patients, and I'm not questioning that it should be covered for those patients. I am questioning that we ought to be precise, very precise about this, and that the committee could have gone many more steps to be precise.

I'm also convinced that just like the FDA sometimes does post-approval marketing, that the committee based on the testimony could say that because the evidence shows that there are very very few people that have actually undergone this, that for initiation of this technology, that at the very least there needs to be standards kept of post-coverage descriptions of side effects and complications from this procedure. Because, it's not unclear to me that there could be some catastrophic results from this. And it's also unclear from the data which the committee doesn't say, that the follow-up periods have been very short in the evidence report. And just like dopa, after five or six years winds up becoming less effective. This thing may actually do things in five or six years that we don't know anything about, and I would want those caveats to be included in these recommendations. But given the short-term benefits when done apparently by competent people in refractory patients, I agree with the committee's assessment that this ought to be covered. I just am very concerned that none of that intent is moved from this very difficult to read evidence-based report to a document that can be useful to CMS. That's my only problem, and I don't believe it's insolvable.

Dr. Sox: Ron.

Dr. Davis: Well, just for the sake of discussion and to allow us to consider Bob's point and make a decision as to whether we should act on it or not, I drafted some language that I could present to the group, if now is a good time to do that.

Dr. Sox: Now would be a good time to do that.

Dr. Davis: Okay. This is on Alan Garber's note dated June 6th, 2002, where there is a statement at the end of item number one that says, "The panel unanimously approved assigning DBS to this category." I don't know that we want to change the question at this point, but perhaps a sentence could be added after this sentence about the panel conclusion that might go like this: The Executive Committee believes that the well-defined set of Medicare patients for whom this conclusion applies is the group of patients with Parkinson's disease for whom medical management is less effective or no longer effective.

And maybe I will look to Alan or others who participated in the panel discussion as to whether this captures the discussion and the feelings of the panel, and the overall gist of the evidence report.

Dr. Sox: Are you making that as a motion?

Dr. Davis: Well, I could do that if people would like to consider it that way.

Dr. Garber: Could I just make a comment that might move us along a little quicker? I think Ron's suggestion is an excellent one, with a slight modification about which is the group, and I would say the patients for whom medical therapy has proven ineffective, or who no longer benefit from medical therapy. It has to be people on whom it's been tried and it's not working.

Dr. Holohan: You mean instead of the phrase less effective?

Dr. Garber: Well, yeah, less effective relative to -

Dr. Davis: What I was trying to capture was the patients for whom the medical management has waning effectiveness or no effectiveness, and my assumption was that some of these studies might have addressed a population that could have fallen into either category.

Dr. Garber: The difficulty is when you get very deep into what constitutes medical failure it's a really messy question. I think what you're saying, Ron, is probably right but if I could try to summarize what I believe the studies were recruiting, it's supposed to be patients who have terrible symptoms, worsening symptoms while they're on medical therapy. And so worse, it's ineffective relative to how effective it was in the past, and the symptoms are considered, we were led to believe anyway, unacceptable by these patients.

Dr. Davis: But presumably the medical treatment would still have some effectiveness; otherwise, it would have been discontinued.

Dr. Garber: On some of them, I think it might have been. So yeah - Tom, did you want to comment on it?

Dr. Sox: It's sort of no longer satisfactorily controlled, or something like that.

Dr. Garber: It's worse than no longer satisfactorily.

Dr. Ferguson: As somebody who has had a lot of Parkinson patients in that situation, it's not always easy to determine. You can manipulate the medicine up one side, down the other. You can actually stop the medication which we've done, we used to do for years, stop it and give them a dry-out period, in which case they're good for a few days to a few weeks sometimes without medication and then boom. So there is not an easy definition of refractory. Dementia enters in in a number of the older ones, and that's another thing to tackle which was mentioned a little bit there. But it's not easy to define refractory and there's not a clear set point. It's difficult, so I think a little waffling there is not unreasonable.

Dr. Phurrough: Just to -

Dr. Sox: Why don't you introduce yourself so we know who you are.

Dr. Phurrough: I'm Steve Phurrough, I'm a medical officer at CMS and I was on this particular panel as the CMS representative.

Just to clarify a bit on the issue around L-dopa and DBS, for DBS to be effective and in fact in all of the studies to be a candidate for DBS, you still had to be responsive to L-dopa. If you're no longer responsive to L-dopa, then DBS isn't performed. So it's not that the medication is no longer effective, it's that you have reached maximum effectiveness of L-dopa with increasing symptoms or increasing side effects from the medication. So no longer effective is not an accurate description of these patients in response to L-dopa.

Dr. Garber: Well Steve, do you want to suggest substitute language?

Dr. Sox: Yes, please do, so we can move on.

Dr. Phurrough: I think the issue is patients who have reached maximum medical therapy with increasing symptoms or side effects to the therapy, I think would perhaps be appropriate language.

Dr. Davis: Could you say that again?

Dr. Phurrough: Patients who have reached maximum medical therapy with increasing symptoms and/or side effects.

Dr. Brook: Can I just suggest some simple - let's just say that the medication, that current medical therapy has produced an unacceptable level of benefit or an unacceptable level of side effects in the opinion of the doctor and the patient, and for those patients this procedure has been proven to be effective. I think that's probably as good as we're going to say.

Dr. Sox: Alan, Bob's Solomon's choice. Which of these do you think would be best?

Dr. Garber: Well, I think this is an excellent illustration of why we left it to CMS staff.

Dr. Sox: Good.

Dr. Garber: You know, no matter what we do, if we start talking about what happened in the physician-patient relationship, I don't see how that's going to guide CMS in making an applicable coverage rule.

Dr. Sox: Well, I think we should go with Dr. Phurrough's suggestion since he has lived with the problem.

Dr. Garber: Well, let me just suggest that we don't vote on the specific language, because it has to be nuance to exactly reflect what was in the studies.

Dr. Sox: Okay, but we still need something to vote on.

Dr. Brook: I object. I think we -

Dr. Sox: Bob, please don't interrupt. If you want you can raise your hand and I will recognize you. So, where was I? How should we handle this in terms of a formal vote? I think what we probably ought to do is to vote for what - Ron ought to make a motion. We then ought to have any more discussion on the motion, then we can vote on the motion, and I think we should all have the understanding that CMS clearly is going to use our recommendations in the way they see fit in making the final coverage decision. So what I'd like to do is now to get this discussion on a little bit more formal basis, hearing a motion, getting a second. Bob, you will then have an opportunity to make the first comment if you wish. So, would somebody please - Ron, would you please make this as a motion.

The Reporter: The basic motion without any amendments is still on the floor.

Dr. Sox: Right, so we're talking about amendments.

(Discussion off the record.)

Dr. Davis: I'm happy to amend my original motion with language like we've been discussing, but I need another 60 seconds or so to craft something, so maybe in the meantime you can -

Dr. Sox: Well, in the meantime - well, we can't have a discussion without something to discuss.

Dr. Davis: The main motion is still on the table.

Dr. Sox: Leslie, do you have some help here?

Dr. Francis: I just wanted to suggest that one alternative is to craft language along the lines that Ron is trying to craft now, but another alternative would be for us to simply call CMS's attention to the textured nature of the discussion at the panel, and have the Executive Committee indicate how important we thought those issues were for CMS to look carefully at when they actually craft a coverage recommendation. That's, you know, saying we think the issue is a central one but we're not going to try to substitute language for yours.

Dr. Sox: Alan, does that make sense to you?

Dr. Garber: Yeah, it makes perfectly good sense in the spirit of both what Leslie and Ron have said.

Let me just point out an issue pertaining to the motion on the table. The bold faced questions with the answers, I think are completely consistent with everything that we've discussed. The issue is explaining what they mean in greater detail, and I think we've always felt that there should be - first of all, in a brief statement you can't possibly capture all the nuances for any of the technologies that we consider so we rely heavily on the explanatory material. I think that Bob and others have correctly pointed out that the explanatory text in my note in the minutes are not adequate to capture that. So what I hear the discussion is saying, fill that out to address some of the issues like which are the patients who benefit, and to the extent that we can, who should be performing the procedure. That was in the transcript but did not make it into the minutes in adequate detail.

So I'm not sure that that requires any major changes to the major questions on the table. It's saying amend the text to more accurately reflect and more completely reflect the discussion. So I hear that loud and clear. And I'm not sure whether that means that the panel has to decide anything, other than to approve what the - or whether the Executive Committee has to do anything other than approve what the panel did with a suitable amended document that is amended just for the explanatory text.

Dr. Sox: What I didn't here is whether we should have some language like the panel calls the coverage group's attention to the issue of the importance of defining the study population and things of this sort. Is that the sense of what you were saying, Leslie?

Dr. Francis: Yeah, not that we try to define it but that we indicate the importance of the problem.

Dr. Sox: Well, since we don't have a motion on the floor right now, Leslie, we're in a position to accept a motion, or I'd say an amendment to the motion.

Dr. Francis: Well, I don't know whether we want to do Ron's first, that approach first, and then do the other approach.

Dr. Sox: Well, we have a motion. We have been talking about amending what we're going to vote on, and I think the proper procedure here is to put a motion on the floor, discuss it, vote it up or down, and give opportunities for other people to have a substitute motion if this is voted down, and then to vote on

the main motion.

Dr. Francis: Then I would move that we amend the original motion to call CMS's attention to the importance of looking more closely at the panel's minutes concerning the definition of the patient population and perhaps also, the centers at which the procedure should be performed.

Dr. Sox: Okay. So we have that motion. Do we have a second to that motion so we can discuss it?

Dr. McNeil: Second.

Dr. Sox: Barbara has seconded it. Now we have a chance to discuss the motion and vote it up or down, and if we vote it down, then we can consider an alternative motion. If we approve it, then I think we can move on.

Dr. Garber: I think we have to hear from Ron about his own amendment first, and then discuss what Leslie -

Dr. Sox: Well, did you formally make a motion to amend, Ron?

Dr. Davis: Not to amend.

Dr. Garber: No, no. He has the primary motion, but he can amend his own motion.

Dr. Davis: But now there's a secondary motion on the floor that's been seconded.

Dr. Francis: I will withdraw that and we'll talk about Ron's first and then we will talk about that.

Dr. Sox: Okay. We have a motion on the floor to adopt the language that's in the committee's report, and the question is, do we want to amend that language. Would anybody like to propose an amendment to the language? Ron.

Dr. Davis: I will offer an amendment to my main motion, which I presume would amend Alan's note date June 16th, 2002. The amendment would be to add a sentence after the sentence that states, "The panel unanimously approved assigining DBS to this category." And that sentence would read, and let me see if I can get this straight: "The Executive Committee believes that the well-defined set of Medicare patients for whom this conclusion applies is the group of patients who have reached maximum benefit from medical treatment and continue to have unacceptable symptoms from disease and increasing side effects from medication."

Dr. Sox: Do I have a second?

Dr. Maves: I'll second.

Dr. Sox: Thanks, Mike. It's now time to discuss that motion. Wade?

Dr. Aubry: Just a point of clarification. Did you mean and or or, in regard to side effects?

Dr. Davis: Well, I raised that question too with the CMS staff, and I was advised to leave it as and.

Dr. Sox: Other discussion? Barbara.

Dr. McNeil: That was my question as well, Hal. That seems to make it much more

restrictive than I sense from some of the discussion.

Dr. Davis: And/or would be another possibility, but I think we need to hear whether this wording reflects what's in the research that we've reviewed.

Dr. Ferguson: Read it again, just where the and/or would go.

Dr. Davis: … the group of patients who have reached maximum benefit from medical treatment and continue to have unacceptable symptoms from disease and increasing side effects from medication.

Dr. Garber: Yeah, I think - Tom may want to weight in on this and Steve might want to, but I think or may be a more accurate description of the literature. They didn't necessarily have to have both to be included in the study populations, as I recall.

Dr. Sox: Dr. Phurrough?

Dr. Phurrough: Since the literature did not clearly define which patients had one or the other and as a group they had both, and I can't recall instances where they had exclusively only increasing symptoms or increasing side effects, and/or would be an acceptable -

Dr. Garber: I think the operative principle here is people for whom the net benefit of treatment was unacceptable, that is, the balance of side effects and effectiveness was unacceptable, either because side effects were so great that it was difficult to continue to take the treatment, or the treatment was providing minimal benefit.

Dr. Davis: I agree with that if you will allow me, Hal, to amend my secondary motion to change the "and" at the end to "and/or", I think that might better accomplish what the Executive Committee is trying to do.

Dr. Sox: Okay. So, other discussion on the amendment?

Dr. Ferguson: Is this under the first or number - you mentioned the -

Dr. Davis: The last sentence under number one. Do you want me to read it one more time?

Dr. Sox: Could you wait just a second? We're moving toward a vote. I have been informed that someone has arrived who would like to make a public statement, and since they arrived at a time when according to our agenda we would still be having public statements, I want to give that person a chance to make their statement. And I'd like to ask now if that person could identify themselves please, and I want to ask you, sir, whether your statement is pertinent to what we have just been discussing, because if it is, you should talk about it now. If not, I will give you a chance to say that after we have voted on this amendment.

Mr. Cohen: No, it's more a general statement.

Dr. Sox: Okay. So if you will sit down, I will recognize you as soon as we have voted on this amendment. I guess we are ready to call the question. Alan?

Dr. Garber: Could I suggest, logically I think, Leslie was unwilling to say this publicly but I agree with her so I'll say it. It should be or, not and/or. My preferred phraseology would be something along the lines of treatment is no longer providing effectiveness commensurate with the risks or side effects.

Dr. Sox: Yeah, or really covers the and/or is what you're saying?

Dr. Garber: Right.

Dr. Davis: I'm comfortable with that if the committee is comfortable with that.

Dr. Sox: I think logically it is correct, or. Okay. Ready to vote?

Ms. Anderson: We do have some administrative things to do before we vote, sorry. For today's panel meeting, voting members present are Wade Aubry, Robert Brook, Ron Davis, John Ferguson, Leslie Francis, Alan Garber, Tom Holohan, Michael Maves, Barbara McNeil, and Robert Murray. Chairperson Hal Sox will vote in the event of a tie. A quorum is present. No one has been recused because of conflict of interest. Now at this time Chairperson Hal Sox will call for the motion for the members to vote.

Dr. Sox: So, all in favor of the -

Dr. Davis: Hal, I'm sorry to interrupt, but I just - since we made all these wording changes -

Dr. Sox: You want to restate it?

Dr. Davis: Yeah, I think for the record it would be good to have this clear. So here's what I have. The Executive Committee believes that the well-defined set of Medicare patients for whom this conclusion applies is the group of patients who have reached maximum benefit from medical treatment and continue to have unacceptable symptoms from disease or increasing side effects from medication.

Dr. Sox: All in favor of the amendment, please signify by raising your hand.

(Show of hands.)

Dr. Sox: Any against?

(Dr. Francis raising hand.)

Dr. Sox: Any abstaining?

(Dr. Garber raising hand.)

Dr. Sox: The motion to amend carries.

Now we're in a position to discuss the amended motion, and what I would like to do just so we stay on target is give our late arrival an opportunity to make his statement. If you could identify yourself, sir, tell us your affiliation and explain any conflict of interest that you might have for us, and then go ahead and make your statement, and I guess I will give you a maximum of five minutes.

Mr. Cohen: I won't take five minutes, thank you. My name is Perry Cohen. I am a Parkinson's patient and an advocate, and I work with the Parkinson's Foundation, and I have no conflicts of interest.

I am speaking on behalf of patients. I was a patient representative on the FDA panel that reviewed this technology about two and a half years ago, and my one comment is why does it take so long? Patients are waiting. The technology appears to be very dramatic in terms of its improvement of patients for whom other therapies are not available, as you just voted. And I am glad to see the committee taking this issue up and considering it carefully, but I'm wondering why it takes so long. The other thing I wanted to mention is, the deep brain stimulation is, the quality of the treatment is the most important thing. Most of the adverse events are results of the treatment, not of the technology, so that Medicare has an important role to play in promoting quality of treatment and it's good that this technology is here, because it provides more incentive into the healthcare system to provide optimal treatment to the patients, which they don't always get right now. We know from studies that we've done at the Parkinson's Foundation that no more than 50 percent of the patients even see a neurologist, much less a movement disorder specialist, and so we think a lot of people are getting inadequate treatment. So, I would like to encourage the panel to consider working with the patient groups to educate doctors and educate the public about

this technology so that it could be used optimally. That's all my statement, thank you.

Dr. Sox: Thank you very much, Mr. Cohen. We appreciate your coming here today. Well, we now have an amended motion on the floor and I'd like to encourage anybody else who wants to discuss the motion to do so now, and if you don't, then we will vote. Bob.

Dr. Brook: I would like to have the committee consider an amendment to have another piece of language, which basically says that the intent of the Executive Committee in approving this motion is that CMS should be instructed to maintain a standardized database on the indications for the use of this procedure and the outcomes of it so that, as Mr. Cohen said, we can make sure that the benefit that looks like in the literature is actually achieved.

Dr. Sox: Is there a second to that motion? It looks like the motion dies for lack of a second, Bob, but we have that statement on the record. Are there any other comments anybody would like to make? In that case, it's time to vote.

All those in favor of the amended motion please signify by raising your hand.

(Show of hands.)

Dr. Sox: Any opposed?

(Dr. Brook raising hand.)

Dr. Sox: Any abstaining?

Ms. Anderson: The motion carries, with one against.

Dr. Sox: Bob, I will give you the opportunity to explain your no vote, if you wish.

Dr. Brook: I believe this is a complex procedure and we were as an executive committee asked to deal with difficult issues and not refer things back to the staff, and that's within our purview, that we should make sure that this procedure is used well. I am concerned that if we don't, we will wind up having a lot of people that are both inadequately evaluated and treated poorly, and the success of these clinical trial data or clinical series data will be mitigated in the real world practice and implementation of this procedure, and it's the purview of this committee to comment about it and to push the government to actually cover this in what I would consider a sensible way, as opposed to what we're doing now.

Dr. Sox: Thank you. Sean, do you wish to respond?

Dr. Tunis: Well, had there been a second on your motion, I was going to see if you were interested in

rephrasing it as, instead of instructing CMS, encouraging -

Dr. Brook: I would be happy to.

Dr. Tunis: -- CMS to undertake the sort of review and analysis that you suggested, which is very much an interest of the program and I think if there were other folks on the Executive Committee who would support the notion of CMS making its best effort to collect data to determine whether or not the procedure is being used in an effective fashion, that would be a welcome commentary.

Dr. Sox: Let's see. I think, John, you were first.

Dr. Ferguson: Well, first I think that the FDA in their approval said that they wanted Medtronic to get data for several more years. And second, I believe that the NIH and the VA are carrying out a large trial. Is that not correct? So, there is an ongoing, or starting ongoing studies. So there will be lots of data coming forward, I believe.

Dr. Sox: Thank you. Alan, I think you were second.

Dr. Garber: Yeah, but I think Tom was going to comment, were you?

Dr. Holohan: On the trial, and to some extent this may answer however imperfectly, Mr. Cohen's comment about why things take so long. The VA has funded a number of centers to do a prospective randomized control trial between best medical therapy and deep brain stimulation in a number of centers nationally. All of the patients enrolled eventually will be eligible, whether they're randomized to the surgery or medical arm initially, will all be eligible at some point in time for the surgical procedure.

I would argue that the existence of that trial demonstrates that at least VA is one federal agency who is in a state of equipoise as to whether this is in fact a breakthrough technology with significant benefit. And if in fact it was clear that this should be widely applied to every case, that it would be unethical to carry out that trial. So I think the existence of the VA trial is further evidence that the available data that the panel reviewed is probably from a logical viewpoint, not as adequate as we would like. And that's one of the reasons why I think the panel was not able to answer these questions as specifically as some of the Executive Committee members would like.

Dr. Sox: Question, Tom. Is this study population for the trial, would they be eligible for coverage according to the statement that we just made? Is it patients that are longer have a satisfactory balance of harms and side effects?

Dr. Holohan: I would have to read the details of the protocol but two of the people on the panel, Kim Burcheil and somebody else whose name escapes me are in fact principal investigators in the VA in this trial, one in Portland or Seattle and the other in Iowa.

Dr. Sox: Alan, do you want to comment?

Dr. Garber: Well, I was getting back to the earlier question of what Bob had raised as a motion, which is what I thought we were talking about.

Dr. Sox: Other comments on the subject of the trial, and then we can - yes, Ron?

Dr. Davis: I also wanted to respond to what Bob had suggested, but if Alan wants to do that, he should go first.

Dr. Garber: I just want to make a very brief statement, which is that the question for me, what Bob proposed I think is something that CMS should always do, not only with this technology or things that, even established technologies, knowing how they're being used and how well they seem to be working in broad dissemination. The question in my mind and the reason I didn't second his motion is should this be singled out especially for that. I think that the answer is no, both because there is a major ongoing trial and because it would seem to minimize the importance of doing this for everything else, and I think we really should encourage CMS to have the system in place, and in fact they do to some extent.

So you know, I would just endorse the broad idea of doing this for all technologies.

Dr. Sox: Ron.

Dr. Davis: I would like to make a similar comment as Alan did. Bob is basically saying it's important to consider what happens when you go from efficacy as shown in clinical trials to real world application where effectiveness is the issue at hand. And what happens when you have coverage and dissemination in the real world, that's a big issue, it's an important issue. I'm not sure it's in the remit of this particular committee. There are other agencies that would certainly have an interest in this, including FDA and AHRQ and others. I didn't second the motion just because I think this is a big complex important issue which I don't feel prepared to discuss and vote on today. I think it would be better if we do believe that this is in the remit of this committee, to come back and make this a separate agenda item for a future meeting, and have some background materials sent to us in advance, and perhaps have a presentation or two that somebody would provide to the committee, and then we could discuss it. Thank you.

Dr. Sox: Sean, do you want to respond to Ron's statement?

Dr. Tunis: Yeah, let me just say, and I think that's a great suggestion, we would be very happy to do that, particularly I think prompted by some previous comments by this committee along these lines of CMS really needing to develop the capacity to look at the impact of coverage decisions in dissemination.

Dr. Sheingold, who is going to speak to you later, has taken the lead in actually developing what we're calling our post-coverage analysis capacity, and we're in the process in the next couple of week of launching four major studies of recent coverage decisions, using Medicare claims data primarily, to try

to look at some of these questions and perhaps in our next EC meeting we can have Dr. Sheingold actually describe what we're doing in those four projects, and get your input and feedback on how we can improve and institutionalize that activity.

Ms. Richner: Sean, has that been, those four projects, have those been defined and certainly have they started, and what, is that in the public domain in terms of what you're evaluating?

Dr. Tunis: You know, that's a good point. We actually are, I think in the next couple of weeks, we are having our kickoff meetings with our contractor to actually try to identify what the questions are. We've identified the topics for these four things, the general topics. I've not actually thought much about the question of, you know, how much to do this kind of with any public notice, not that we're particularly trying to do it in secret, we just haven't gotten that far. So my guess is that it probably makes sense for us to maybe on our web site to mention what we're doing and the topics, and I guess people can contact us like they do on everything else to give us some suggestions on those.

So again, I'm mentioning it now because of, we're not trying to be secretive about it. I think it is something that, again, we have been encouraged to do and are getting a start on it, and we will see how it goes.

Ms. Richner: And also on the point about essentially follow-up clinical trial evaluation or further prospective studies, there's a panel that you were recently on that is evaluating that particular topic. That was the Clinical Trial Research Council, I believe, and that would be very interesting to bring in them as well and I think this panel would benefit from that.

Dr. Sox: Is there somebody along here? I think you've been waiting longer than anybody else, Linda.

Dr. Bergthold: I just wanted to address just the issue that Mr. Cohen brought up about why does it take so long, in sort of more plain English. Because I think for the general public sitting and listening to this, this is technical stuff, this is research based discussion of really fine points, but it really doesn't kind of touch I think one of the major issues of why this takes so long and why it should take so long, so if I could just make one or two comments about that.

I think the reason why we're here, my understanding and the role that I feel I play here is once Medicare decides to make a coverage determination, it flows out into the system very rapidly and gets implemented in a number of ways at a local level which the public probably doesn't quite understand. In other words, once it's covered, almost anybody can do it and almost anybody does, and they do these procedures sometimes well and sometimes not well. So the reason it takes so long and ought to take so long is that once it's out, it's very hard to pull it back. And I actually applaud you for looking at other coverage determinations that have been made, and the balance that we all try to strike is between trying to get things out to people who need it without getting it out to everybody in a way that can be harmful. And I just, I don't know that we talk about that enough, and I don't think any of us tries to sort of slow the process down for some arcane reason of research evidence, but only because we know that once

these things are out there, they can be done in this American system in ways that could be harmful.

Ms. Richner: But the FDA essentially monitors, the FDA also has that remit in terms of protecting the public safety.

Dr. Bergthold: Right, for some things, not surgical interventions.

Dr. Sox: Bob's next, and then Mike.

Dr. Maves: I actually wanted to comment, and actually Dr. Bergthold brings up a good point. Bob actually asked two questions at the beginning of this discussion. One was, can we define the patient population that we're trying to treat more accurately, and I think we've made some steps along that way.

The second thing that he asked was, can we better define those who should do this, and maybe Sean can give me some insight. Because if you look at the Blue Cross/Blue Shield TEC assessment, one of the sort of criteria is the improvement must be attainable outside the investigational setting. And really the only thing they said here was this should be done in experienced centers demonstrating comparably low procedure related morbidity and mortality, I would presume to just general neurosurgery.

On the other hand, Sean, I would appreciate some feedback as to where you go on that because as sort of a recovering act omission, typically we will have early case reports of something going on, a few promising short studies. We then get a clinical trial, and then using the next cycle once this is out in the community, it's the complications of articles start coming out. And so there is almost a reproducible cycle. So one does worry, and I think Mr. Cohen's statement that, you know, side effects can be a significant part of this, how do we go about doing that? I mean, that's a tough job to wrestle with and I think as Dr. Bergthold pointed out, once a coverage decision has been made, then really qualifying, you know, or trying to better describe who should do this process even beyond clinical trials or VA study, becomes kind of a moot point at that point. So how do we do that, Sean, on things that we know can have some really fairly serious risks, and what are we doing at CMS to look at that?

Dr. Tunis: Is the question how do we in the coverage policy itself address the issue of limiting this to specific centers? Well, it's not done easily and there's, you know, depending I think on the situation there might be some tools. But as an example in this case, one at least could consider requiring as part of the coverage policy that only institutions with a dedicated neurosurgical intensive care unit would be eligible to do the procedure. I'm not saying that's what we would do but that's the sort of thing you can look at doing as a way of trying to limit this to the more experienced centers.

So the challenge here is operationally defining more experienced centers in a way also that's consistent with the scientific studies and the facilities and capabilities and training of the individuals that are involved in those studies. So we have to try, we can't make it up out of whole cloth obviously, so there has to be a way to do something from which the data was actually derived.

I don't know if Dr. Phurrough or Perry Bridger has any other thoughts on this. You've clearly been thinking about this topic, but what other sorts of things might you be thinking about?

Dr. Phurrough: We spent a lot of time thinking about this topic, how do we clearly define the criteria and in fact we have some things such as should patients, should we require evaluations by movement disorder specialists and not just the neurosurgeon. Should we look at the issue of whether the facility has on site the appropriate equipment to do the procedure, versus in some cases where stereotactic surgery is done, you're sent across town in an ambulance to have this CT performed that says here's where the electrodes go. Things such as should you have electrical mapping as a requirement on site.

So there are specific criteria that we think we can require that would limit where it can be done. The question is, is the evidence available and adequate to say that yes, we should make those specific criteria. So we're struggling with it, it isn't a simple process, and in fact would be one of the newer, one of the precedent setting times we've done that outside of transplants.

Dr. Maves: And Hal, if I could just follow up, that might well be a topic for future discussions from the Executive Committee, to really look at that and perhaps have some examples of how that's been done very restrictively. Someone mentioned transplant obviously, and sort of what's been our experience on that. I mean, that's the only reason in bringing it up today is it clearly is a concern and maybe that would be something the Executive Committee should discuss and perhaps make some

recommendations, or at least get a better feel for how that works in CMS.

Dr. Sox: If I could just step down for a second and comment, it sounds as though CMS, the coverage group is trying to do a lot of the things that we have been discussing this morning, and the question is who should be doing it and where do we sort of draw the line between this committee's responsibility and the responsibility of CMS staff. I think that's an issue that we've raised very effectively and forcefully, and that we're going to have to wrestle with at a future meeting. Okay. Now, Bob.

Dr. Brook: I raised these issues because I did read the evidence report, and I want to read you some sentences from the evidence report. The improvement - the last page of the evidence report says, "The improvement must be attainable outside the investigational setting." And it concludes that, "Bilateral DBS meets this criteria when performed in centers that can demonstrate" - I'm just quoting from the evidence report - "comparatively low procedure related morbidity and mortality." Then back on another, page 18 it says, "The current level of technical variation would seem that until appropriate clinical trials have been completed, choice of target and method of target location will depend on the center in which it is performed."

Now what I'm concerned about and why I voted no on the motion is I think we have a mess at the moment. We have not stated precisely where this place ought to - the responsibility of those centers that actually are doing, or are going to do these procedures, and I think it's within our purview. I would be happy to wait another year to have that discussion, but I really do believe we could have incorporated it

under this discussion. That's the first thing.

The second thing is that based on this other evidence about, that you don't know which kind of procedure to do, that it would have been within the purview of this committee to suggest that anyone getting this procedure when covered by Medicare be enrolled in a clinical trial that compares the different techniques. That's easy to do, it's been the policy of many places in treating childhood leukemia for a long time period. Nobody is going to do this other than CMS if they want to do this. And third, I'm really concerned about if we label this a substantial breakthrough that, is this the preferred first treatment of choice. And there's nothing in what we've said that suggests that this ought to be tested aggressively under - for people who have early Parkinson's disease, before you wait until they're unemployed, disability, you know, everyone's on MediCal because they've run out of money in terms of what's going on, is this something that we ought to have pushed the science, because there is no science of using this early in what's going on. And finally, I am concerned with Tom's statement, really concerned now. I was trying to find in here the description of the VA patients, but if we have decided this is really eligible for coverage, then his trial is unethical if it has the same patients that are in it right now. He answered his own question. We would, as a human subjects committee could not approve that trial and that trial ought to be stopped. I will go on record to say that, if it includes the same kind of patients that we have just approved for routine coverage.

Dr. Holohan: I said that both arms get the surgery.

Dr. Brook: But it's a delay.

Dr. Holohan: Yes.

Dr. Brook: It's a delay, and you haven't done it in earlier patients. I mean, if it really includes then at least people ought to be offered, the people ought to be offered the coverage for this procedure, and they ought to be told right now that they are being offered it. I don't know if they are still enrolling, or they finished enrolling or where they are in this process, but we have some loose ends that haven't been put together.

I'm sorry for being negative about it but I think we have advanced the state of this field, that we can do more than what we have said to do, and I think the only justification for the Executive Committee is to do that more. Otherwise, I really do believe it ought to be disbanded and we just ought to, you know, use the process from the subcommittees to give HCFA or CMS as a body.

Dr. Sox: Do you want to respond to Bob's question or raise something new, because otherwise, Leslie is ahead of you.

Dr. Francis: No, you should go.

Dr. Garber: No, I want to respond to Mike Maves' earlier comment. Is that in order now?

Dr. Sox: Go ahead.

Dr. Garber: I just want to underscore something about what Mike said. I think it would be extremely helpful for us to get some feedback from CMS, especially the reimbursement people, about what's really feasible in terms of restricting a procedure at certain centers. After having spent years, and Wade having spent even more years I think on Blue Cross/Blue Shield's medical advisory panel wondering about what constituted attainable outside investigational setting, it's a tremendously tricky issue. But it's got to be driven as much by what's your goal. Your goal is to make sure that people get good results and on the one hand, if there's a volume outcomes relationship you could say that everything should be done only in specialized center that achieve certain volumes. That's not really the intent here because you think that it might, I believe the intent is to make sure that a procedure which is potentially harmful or falls well short of the potential benefits if done in the wrong setting should be done only in the appropriate setting. And in the case of deep brain stimulation we did hear these comments that it should be done, they should have a full movement center evaluation and so on and so forth, but that wasn't really addressed directly by the studies so it was hard to come up with anything concrete.

So what we need, first of all I think, is some guidance from CMS about what's feasible in terms of making sure that patients get channeled to centers that do it well, and then we I think can provide more useful advice when we evaluate the technologies of what are some of the characteristics that CMS could use that would help channel patients the right way. And I feel that right now in that respect, we don't know quite enough about how CMS operates to be as helpful as we could be.

Dr. Sox: Leslie?

Dr. Francis: Well, I was going to make a more general comment that went back to Sean's, so if you're -

Dr. Sox: On this point, Wade, or something new?

Dr. Aubry: Yeah. I would just second what Alan says. As a former Medicare carrier medical director, there are a lot of operational issues. One in particular is that there is no preauthorization in Medicare and a lot of these controls and direction of patients to appropriate evaluation and appropriate centers would best be done on a preauthorization basis. Otherwise, you're talking about potentially denying an entire hospital stay or professional services after it's been done because some aspect of the coverage determination has not been met. So it's complicated and there aren't mechanisms in place outside of the transplants to do that, so it's a very complicated issue. And also, I think it's probably not feasible to do it for very many topics, so if it is done it should probably be selected carefully in terms of which procedure.

But I think it is a worthy subject for discussion at a future meeting, to see what could be done, particularly for high risk procedures where there's a lot of variation in quality, and to more precisely define the role of the Executive Committee in these decisions.

Dr. Sox: Thank you. Leslie.

Dr. Francis: Well, I took us to be doing a little scoping out of the territory, and at the risk of getting us or the coverage and analysis group into an awful swamp. What you're proposing to do is look at what's happened when coverage decisions have been made, right, some of the recent coverage decisions? But what about all the rest of the stuff out there which is being covered without any kind of a recent coverage decision having been made? And I don't know whether it's ever going to be - whether anything is ever going to be referred to any of these groups on the question of whether there's enough evidence to support continued coverage.

Dr. Tunis: So, I guess stay tuned is a good answer to that. I mean, as most of you probably know, there is virtually no historical precedent in Medicare for withdrawing coverage for things that have been either covered at the national level or routinely paid for. It is my sense that to the extent that the procedures being developed for coverage and some of the framework, the evidence based framework, to the extent that becomes more acceptable and routine and fluid, the opportunity to look at new evidence, particularly if things are ineffective or potentially harmful, might actually lead to a serious consideration of revising coverage, and there probably will be an example of that in the near future.

I think the issue, Alan, that you raised on what are the tools, potential reimbursement or what tools does CMS have potentially to limit coverage and payment to particular centers is another very difficult area. To the extent that there are potential tools, again, they would most likely exist in coverage policy, not in reimbursement policy. I'm not aware of any payment policy mechanism for limiting to centers of excellence outside of demonstration programs or outside of transplants. But obviously a case like this one, DBS, where there seems to be some reasonably compelling evidence that it ought not to be done by unskilled or you know, or where the evidence comes from the most proficient centers and there is evidence to suggest that if it is badly done it would be worse than nothing, this is the kind of issue that would prompt us to pursue that.

So I'm sorry we have not been able to give you clear direction on this. It's not an area where the policy has been very clear or the pathways very well tread.

Dr. Ferguson: What about carotid endoarterectomy? Didn't you restrict that when you covered it?

Dr. Tunis: Not that I'm aware of, no.

Dr. Holohan: Actually the technology recommendation did include limitations to centers that had less than a 3.1 percent operative mortality rate, but that was not implemented.

Dr. Tunis: So that came from the OHTA evaluation, but it's not part of the Medicare policy.

Dr. Ferguson: Oh. You didn't follow OHTA?

Dr. Sox: If there's any more discussion on this particular point, let's have it, and then we're going to stop and take our well deserved break. If there's no further discussion, we will break for 15 minutes.

(Recess from 10:03 to 10:23 a.m.)

Dr. Sox: I talked to Sean about this new category that Alan used and just ask you to recall for a minute that the interim guidelines for procedure for MCAC contains a system for ranking the sort of impact of a new technology that's under consideration, and Alan's group has pointed out that there seems to be too big a gap between the top of the scale and the next level down. So the question I'd like to raise before we move on to the next scheduled item is whether we should insert another level of impact along the lines that Alan's group found useful in their discussion. So Sean, if you could just tell us that this is potentially useful, so that a discussion of this would be in order.

Dr. Tunis: I would just endorse that since it is part of our methodology to assess the magnitude of impact, and to the extent that there are things of magnitude that aren't expressible in the current framework, that it's worth looking at that and potentially modifying the framework as we go along. So, I think that would be useful to consider, to have this discussion and consider adding a new category that reflects that.

Dr. Sox: Just to remind you that every time that we have modified the interim guidelines, which is two or three times now, we have taken a formal vote, so if we were to make a change now, it would require a formal vote. But my suggestion is that we have discussion informally to see whether we in fact do want to make a motion.

Alan, maybe you could just tell us what you did and briefly why you thought it was important to put something intermediate between breakthrough and small but definite advantage.

Dr. Garber: I will be very brief. I mentioned this when I presented the panel's findings but essentially we thought that in general we will come across technologies frequently where there is a substantial benefit but it falls short of the language that we use for breakthrough, which was that it becomes so effective that it becomes the standard of care. I think this is a perfect example.

As Tom was saying about the trials that the VA is running, if you thought it was the standard of care, you couldn't possibly do a trial. Now I don't want to get into the details of why the trials might or might not be appropriate. There is certainly unanswered questions, even though we think that this technology is effective in certain patients. So we thought that this not only applied to the technology we evaluated but it had greater use and I think the panel believed that the Executive Committee should therefore consider this as a level of effectiveness that should be included in our general set of levels.

Dr. Sox: Leslie, we'll just sort of have a general discussion now.

Dr. Francis: Yeah. I had a question actually about whether the problem that you're pointing to is the need

for a single new category or whether the problem you're pointing to is that two is way too restrictive. I mean, it seems to me there's going to be a whole continuum of the extent to which health outcomes might be improved, like somewhat more effective, substantially more effective, a great deal more effective, and so on, all the way up to breakthrough. And you know, another possibility would be to take out the "by a significant albeit small" language, and say that the way we understand more effective is that the new intervention improves health outcomes as compared with the established services or medical items, and then invite the panel to comment on the extent to which it seems that health outcomes are improved. That allows for a whole variation of gradation rather than trying to - I mean, I can envision just extraordinary arguments about whether it's somewhat or substantially, and which category to put it into, which is probably not worth it.

Dr. Sox: Alan, do you want to respond to that?

Dr. Garber: Well, I think that is definitely another way to go and it's a very thoughtful way to go. The issue really comes down to how many discrete categories you want and how simple you want this to be. I think in the view of the panel and certainly my view, you can do it with basically the three top categories and you don't need a lot of fine details beyond that. That is, there's some stuff that's just clearly home run, incredibly effective, everyone should get it who fits within the right subgroups, that's a breakthrough.

And then there's another area where yeah, it works and it's better, but it's incremental. And I think there's some fairly broad agreement about what that means. You know, obviously there could be disagreements around the edges and then in between that, stuff is substantially more effective and not quite a home run. So I actually think that's a workable set of categories and we could do something that allows more continuous gradation but I think if you want discrete categories, those ones work and are likely to work.

The big issue for us is we'll learn as we go along. I mean, when we get experience that's how we came up with this category. We tried to apply the existing categories and maybe we need further modifications. But I think if you wanted to make a small change from what we have now, just adding this category would do it.

Dr. Sox: Linda.

Dr. Bergthold: Just for sake of symmetry, I think taking out the "albeit small" would make two much more parallel to the others, that you wouldn't have to create a new category, as Leslie suggested. That you just sort of look at the parallelity of these different categories.

But as I recall when we were doing this, it was done fairly quickly and there wasn't a lot of time to discuss all the nuances, so it's not like it's written in the Bible or anything, we could change it to make it a little bit simpler.

Dr. Sox: Sean.

Dr. Tunis: And just as a thought, really emerging from some of the issues we have been dealing with lately, I think a nuance here that we could propose that the committee think about whether they want to help us with is a category where there is, you know, clearly an improvement in outcomes or a benefit in outcomes, but the importance or clinical significance of that improvement is still somewhat in doubt. So it seems to me that that's a category we kind of need and maybe you meant by the albeit small, or maybe it's something lower than that. And then there's the things that are clearly improvement, it's clearly clinically important but it's not a breakthrough. So you know, there's sort of no doubt in your minds this is about a real effect and it's an important effect. But what I'm not sure you have - so that might be the new category you're proposing, but I'm not sure the existing category kind of gets at what I'm suggesting, which is it's a small benefit and it's not even completely clear, while it's a real improvement it's not completely clear that that's clinically significant or clinically important, or particularly superior to what's already out there.

Dr. Bergthold: Isn't that number three?

Dr. Tunis: I don't know, maybe it is number three.

Dr. Garber: Yeah, number three gets at some of that.

You know, I don't want to make more work for us but a useful exercise at some point to go through is to actually take a number of technologies that we're familiar with and try and assign some categories and see how workable it is. I mean, that's a little bit of pilot testing that might go a long way to giving us workable categories. But my impression is that something like this current scheme if amended would work.

Dr. Sox: Other thoughts? We're kind of, I guess as I see it, we're trying to decide between three options. One is no change, one is to change the more effective category to delete the "albeit small" and just simply say significant margin, perhaps with adjectives to be inserted at the committee's wish. And the third to be to create a new category. So I think those are the three options that we are discussing. Ron.

Dr. Davis: Well, I imagine that most of the matters that come before us are going to be in these upper levels and not in the lower levels, and therefore a little bit more precision at the upper tier would be helpful to us. So my preference would be to support Alan's recommendation and to add another category so that we have more options at the level of more effective. So when you get to the point, Hal, where you might be willing to entertain a motion, I can make a motion.

Dr. Sox: Yes, Bob.

Dr. Murray: Everything that we have been talking about has been clinically oriented but presumably these categories would have to apply to laboratory and diagnostics if anything does come before the laboratory and diagnostics panel again. Actually Alan's original suggestion or proposal would work

better for laboratory where you're dealing with quantitative, more easily measurable effects, so I would support Alan's original proposal.

Dr. Sox: Other discussion or questions? In that case, I guess we are ready for a motion, Ron.

Dr. Davis: I move that we add a category like the one that the panel brought before us earlier today.

Dr. Sox: And could you just, Alan, sort of -

Dr. Garber: Do you want me to read that?

Dr. Sox: Please.

Dr. Garber: It is: Substantially more effective: The new intervention improves health outcomes by a substantial margin as compared with established services or medical items.

Dr. Sox: Okay. Is there a second?

Dr. Murray: Second.

Dr. Sox: Any discussion of the motion?

Dr. Brook: I would just like to take Alan's suggestion as a positive one, that it would be useful to try to provide guidance to the committee and I think quantitative guidance, what these words mean, and I think that can be done. And that we might want to revise the scale beyond that after we've done that, but I really do believe that quantitative guidance can be given to what these qualitative words mean.

Dr. Sox: Okay. So the motion is that we insert a new item between one and two on our scale of size of health effects which would be substantially more effective. All in favor of this, please raise your hand.

(Show of hands.)

Dr. Sox: Any opposed? Any abstentions? So, the motion carries unanimously. We will make that change in the interim guidelines.

The next item on our regular agenda is a discussion of CMS referral guidelines for MCAC issues. And we're all eager to try to hear about this so we won't delay things any further. Steve, are you going to make the presentation?

Dr. Davis: Hal, I'm sorry to interrupt, but did we formally act on items two and three in Alan's note dated June 16th? Was that considered to be part of that original motion or was it item number one only?

Dr. Sox: I'm pretty sure the motion covered all three.

Dr. Davis: The reason why I'm asking is because when we added the sentence that I crafted with help from others, I recommended that we add that at the end of item number one, and I'm just wondering whether that needs to be added at the end of these items as well.

Dr. Garber: I think it was understood that you intended it to be parallel and to apply throughout. That's how I interpreted it anyway.

Dr. Sox: If there is no objection, then it is understood for the others and should be inserted.

Dr. Davis: Thank you.

Dr. Sox: Steve.

Dr. Sheingold: Thank you, Hal. For those on the committee that don't know me, I'm Steve Sheingold. I'm the director of the division of operations and committee management in the coverage and analysis group. That's the division that helps run this committee and make your arrangements. I'm fortunate to have the people you know well like Janet and Kim and Michelle outside on my staff.

In your packet today you should have a little two-page document that starts out with Guiding Principles for Development of Referral Criteria to external TA assistance and MCAC. I just wanted to make some brief remarks to explain sort of the origins of that two-page document, what our objectives were and why I bring it to the Committee at this time.

I think over the past few years while the committee has been in operation, the evidence based coverage process at CMS has really grown and matured. Sean has done a great job in building a staff that's increasingly sophisticated and able to do systematic reviews and internal analyses. And so we really do have a mixture for our national coverage determinations of internal analyses, those we seek assistance from AHRQ, both from their internal staff and through the evidence centers, and then those NCDs which whether or not there was external technology assessment assistance, come to this committee for further review.

Within our process for making coverage decisions, and there has been a considerable amount of interest in the public, how we make the decisions to take these paths, how do we decide when we ask for external TA assistance and when we ask for MCAC review. Actually the document got started after a visit from Randel in my office, who was asking these very questions.

And so what you see in front of you, although it's two pages, actually reflects a tremendous amount of internal discussion that we've all had over the past few months. Just in general, what you see describes our current thinking about the technology assessment process and how it fits into the coverage

determinations, and some basic guidelines for referral. And after just going through it very briefly, I'd like to open it up for a discussion. We thought it was time to get some expert and public inputs on these important areas.

The first thing we tried to do here is first of all set the record straight on the role of technology assessment in coverage decisions. There seemed to be thoughts both internally and externally that some national coverage determinations use technology assessments and some don't. What we're saying here is that every national coverage determination requires a full technology assessment process, and that technology assessment process will have the steps that you see outlined in bullets before you. That is development of the background information, development of an analytic framework, the specification of the assessment questions, a complete description of methods, the critical appraisal of evidence, and development of methods to summarize and describe the findings.

Internally we're piloting a few different types of quality of evidence tables that we hope to make a routine part of our decision memos in the near future. What we've also found is that starting in the middle of this process, just getting some literature and doing a review without the pertinent background information, analytical framework to guide us, can cause more problems than it's worth in skipping those steps. So we wanted to put that on the record and just following up on what Sean said about some post-coverage analyses, those bullets you see are the pre-decision steps to the technology assessment process. We are in full recognition that a complete technology assessment process is ongoing and continues after a decision is made.

So all NCDs, national coverage decisions, will have this complete process, and we're working with AHRQ and with our internal staff so that the technology assessment reports that you see, in addition to our internal presentation of results and our decision memos, will all follow very similar format so that there will be greater consistency not only in methods but in presentation from decision to decision. For selected coverage decisions, we will seek external assistance on one or more of these technology assessment activities. It may be just in the literature review or the analytic framework, it may be for the entire set of these activities, and we are working closely with the AHRQ staff. Some of these activities may be done by their internal staff and some through their evidence centers.

And again, for a different subset of these determinations, whether we've taken that external TA assistance step or not, they will come to MCAC for review.

And on number four on the first page I've outlined very briefly, sort of a summary statement we all agreed to, of what we look to MCAC for in terms of expert advice and discussion of other factors that should be important in helping us make a coverage decision.

The second page then looks at some very general statements, criteria or guidelines for when we should request external TA assistance and when we should refer a copy for MCAC review. For external technology assessment assistance it really comes down to two broad issues, scope and complexity, and can we do it in a timely manner with the competing uses of our resources, and of course the complexity.

Do we need a supplement to our staff on clinical or subclinical or methodologic issues.

Again, given what we've said about what we look to MCAC for, then again, there's two very general criteria there for when issues should come to this committee for review. And we've added that last one on the second page, which is discussion or clarification of general methodologic issues which may affect all or some future national coverage determinations, and is certainly a valid reason for convening a committee.

And so, with that brief description I will open it up for general discussion.

Dr. McNeil: Steve, thank you. I actually found this very helpful. I have one question and may it's outside the purview of the discussion today, but the question that I'm frequently asked is the criteria that go into a national coverage decision versus a local one, because that may be more important in terms of what gets covered than this finer gradation here. So could you say a little bit about when things get bumped up to you versus when they just stay locally or don't get acted upon at all?

Dr. Tunis: Well, that also is not a highly defined process. It doesn't rise to the level of mysterious but I mean, the most typical situation is obviously, the default situation is that coverage decisions are made at the local level, with the one exception of when there is already a national coverage policy, particularly a national non-coverage, it can only be amended through another national coverage decision. So if you take the default that things are local until some problem develops, and that problem would usually be there's a degree of variation or variability between the local policies, or there is a consistency of local policies that somebody objects to and thinks that needs to be addressed at the national level and brings it to attention at the national level.

So mostly, national coverage decisions have been taken on in a kind of reactive or responsive way when problems develop and are brought to our attention, mostly through variation.

There is a small movement afoot to try to be more proactive about taking on national coverage decisions, particularly for technologies that are emerging that are, you know, potential great immediate impact on public health. But that's the only other criteria that exists for taking them on.

Dr. McNeil: Could I just follow that up, Hal? As I hear what you just said, Sean, this would imply to me that individuals who were worried about coverage for various drugs or devices or procedures, whatever it is we're talking about, really should try to keep the coverage, or should try to prevent the coverage decision from becoming national, because a no nationally is really horrible for them, whereas a lot of variability in the local sites doesn't hurt them as much. Is that right?

Dr. Tunis: Maybe Randel has some thoughts on that, I don't know.

Ms. Richner: Well, I think we have to remember - I mean, this is the first time I've seen this, today, so I haven't digested what you've written here. But essentially I think one of the stimulus for going to speak

to Steve directly about this is MCAC was started because there was a closed process in terms of how decisions were decided on the national level, and it was closed, and we had the TAC system. So then, this was a public way to identify that.

So we still have a problem where if you have a technology and it is essentially, or a new procedure or whatever it is, it's sort of unknown as to how it's elevated to the national level. It can happen in a lot of different ways. It can be that internally CMS sort of decides, they can just essentially informally within their staff decide that they want to evaluate a technology. Or a manufacturer can come and say we want to have this covered. Or there are ways where, you know, a payer or Blue Cross, there's a lot of different ways in which this can go to Steve for a decision.

And so we're concerned that okay, then what happens once it goes there, in terms of how do they triage it, how do they decide that it comes to MCAC. How do they decide that it should stay with the local carriers. How do they decide that they are not going to do anything about it. I don't know if this document does that or not.

Dr. Sheingold: Yeah, this document starts down that path. It assumes that the national coverage determination request has been accepted, whether that was generated internally or as the topic you discussed this morning, that was a request from a member of the public. There's a well-defined set of criteria for when there's an acceptable request for a national coverage determination, so that any member of the public or another payer can make that.

Ms. Richner: A doctor or anyone can request a national coverage decision on a technology, and that's a concern.

Dr. Tunis: And it's important to know that if a national coverage decision is requested and it's got all the contents required, a letter that says it's a national coverage decision request and it's got the data, and for a beneficiary we will generally actually provide, you know, do the searches. We don't have a mechanism to say no, we're not going to do it as a national coverage decision. So that's actually the way our procedures are set up.

Dr. Sox: Bob?

Dr. Brook: If it's already done then maybe I'm just being redundant, but in the interest of transparency and public accountability, I believe that on a yearly basis we ought to ask from the Executive Committee to CMS, to release a comprehensive report that indicates what's happened to change in coverage at both the local and national level, and what was the criterion that was used by HCFA or CMS to make those decisions. It would be useful for us to understand the scope where what we do fits into all this other activity. If such a report is available, then that would be great and I would like to see a copy of it.

But the question is, is that a transparent public document that - do you know the variability of coverage of all these things at the local level? Do you know how - I mean, how are these things or requests

handled, all these letters that come in? And it would be useful if we sort of were able to see a report on that if that's available. I don't think I've seen that report.

Dr. Sox: So what you're asking for is a report of the coverage decisions that CMS made during the year, but kind of broken down according to whether it came to us or didn't come to us?

Dr. Brook: I mean, it would be interesting to see the technologies, what's been requested, what's no, what's yes, what's going on here, so that we could maybe give some advice about the whole process to CMS, since we're groping for a new role for this Executive Committee. But may increasing transparency of the process would make it a better place for everyone.

Ms. Richner: There's a few studies that are going on now, a project called the Med-Pac on this particular issue, on local coverage decisions. As well as, if you go to their web site, to the CAG web site, they have a list of all of the technology assessments and coverage decisions that have been made over the last year. And generally, Sean, I don't want to speak for you, but the industry has asked repeatedly for essentially what are the guidelines and criteria associated with those decisions, and the answer generally is that if you go to our web site you can get a good idea of how they've made their decisions, and get a general feel and flavor of the types of methodologies they've used for those coverage decisions. The ones we've done through MCAC have been very few a far between so you really, if you want to know what the coverage group is doing, you go to their web site and you essentially look at how they've written their decisions.

Dr. Sox: Mike, I think you're next.

Dr. Maves: I would applaud any kind of transparency we can put in the system. I get lots of questions from a variety of different constituents, physicians, companies, you know, government agencies and so on, so I think anything we can do. I think also it would be very helpful because I think one of the things, and Randel and I have talked about this, is to also put this in the context of parts of this such as the CPT codes, the RUC, you know, there are many many steps in getting this. You know, it's sort of fighting the ground war to eventually have something that's going to be out there which, this is one part of it but it's not the only part.

Ms. Richner: It's a small part.

Dr. Maves: Yeah. I think if there were a way to just simply put that together and let individuals kind of know that. I know when we first started this committee, I think there was a couple of presentations by folks from then HCFA talking about the process in its entirety and in the context of that, but I think that's one of the sort of misconceptions that I regularly have to kind of get on the stump and tell people, let me explain to you all the little steps that are in here. So at the end there's a CPT code with a RUC, and Sean's folks at CMS will pay for it. That's a long road to get to that point, and one where I think the more transparency, the more openness, the more information we can give people, I think it will help take care of a lot of problems for all of us.

Dr. Sheingold: It's a long road and it's not linear.

Dr. Maves: That's right.

Dr. Sheingold: But I will just mention one other thing Janet reminded me of. Of course as Randel said, anything we do as far as national coverage decisions is available on the Internet when the decision is made, as is the proceedings of this committee, but we're also required to do a report to Congress on an annual basis now to describe what's happened to national coverage in the past year. The first one has already been delivered to Congress, and when is the next one due?

Ms. Anderson: December 1st.

Dr. Brook: Does this include the issues that Barbara raised?

Dr. Sheingold: It does not address the local coverage policies.

Dr. Brook: For instance on this one right here that we reviewed today, there was no information. Is this covered anywhere under Medicare now? I mean, did the panel know that before they, and the justification for why individual carriers or fiscal intermediaries or whatever decided to cover this?

Dr. Tunis: I don't know. Perry, was that information provided to the panel?

Dr. Bridger: I'm Perry Bridger, I'm one of the analysts in the group and the lead analyst for the deep brain stimulation issue. We presented to the panel the status of local coverage for DBS, and we also on the panel had one of our carrier medical directors who's a neurologist and has been involved in this for quite a long time, to explain the evolution of coverage on the local level.

Essentially, currently most Medicare contractors do cover deep brain stimulation for Parkinson's, and they've got local medical review policies detailing their coverage.

Dr. Tunis: Bob, to sort of respond to you, there is no place where there's kind of one report that looks at local policies, which new ones have been done, all of the requests that have come in, why we've decided to take one on or not. Obviously it would be a lot of work to put such a report together, but it could be done. Certainly one thing that we could do for this committee is have a series of presentations from different staff on coding, you know, local coverage, national coverage, at least to sort of map out what the whole process looks like and where this piece of it fits in. Whether or not after that you all would encourage us to do some sort of report that you know, catalogs all these activities on a regular basis, you know, that's something that can be talked about again. It would be a significant amount of effort to do it.

Dr. Sox: I think Alan was next, then Leslie, John.

Dr. Garber: Well, one of the things that I think would be helpful to explain to the public and people like me who really don't know the details, is what the rationale is for having this split between national and local coverage determination. Because I would suspect that to an outsider, it looks kind of odd that there is this duplicative activity of evaluating the technologies at the local carrier level when presumably they are drawing on largely the same databases. And maybe there's a lot of local variation in outcomes and if we knew that there was such local variation, that is, if you get DBS in Georgia the outcomes would be different than if you get it in New York, one could understand why you might have local coverage. But given that we had so much trouble finding out who does well and who doesn't with this, it's a little hard to imagine that they have much better information to support that.

So I think if we had a better understanding of why the default seems to be local coverage rather than national when from an outsider's point of view it seems the default in most coverage decisions would logically be national with occasional local coverage determinations, I think that would be helpful. I'm sure there are very good reasons but they're just not ones that are known to those of us who are outside this system.

Ms. Richner: Think about the magnitude of the numbers, the sheer numbers, and you're talking about over 6,000 or so, that's a number that's floated around in terms of local coverage decisions, whereas there was 30 done last year on a national level. And you've seen the onerous process that we've undergone for these particularly vexing problems we've had, and consider what this would do in terms of workload and operationally to a national system. I'm just speaking very specifically and tactically that it would be an unbelievable burden in some sense.

Dr. Garber: But Randel, I think you've just given a very strong argument for why it should be national, because a lot of these are judging the same technologies multiple times at the local level, whereas there might be a single national determination.

Ms. Richner: Right. Generally they are fairly consistent. You will find that the variation is very minimal in the studies that have been done recently.

Dr. Garber: Exactly. So that would tend to support a national process so they don't reinvent the wheel multiple times.

Dr. Sox: There were several people waiting but let's see if there are responses to Alan's comments.

Dr. Bergthold: Yes. I just wanted to make a comment that Susan Foote at the University of Minnesota is doing a study right now that should be completed within what, is it a year, or parts of it soon and parts of it later, on just laying out sort of what happens at the local carrier level and historically why it was developed that way, what happens, what's the variation, what kinds of things are looked at, what kinds of issues. It's really going to be I think a very instructive study. So I would say wait for that study for the details.

The second thing is that I just want to make a comment. I would love the day if it comes when the public really participates in this process. Right now when we talk about transparency and the public, it's really about the industry having access to this process, it's not generally about consumer groups or advocacy groups. Basically they don't even know we exist, they don't know how to interact with this process at all. So you know, let's just qualify when we talk about, in grand terms about public access, that that day has not come yet and so far it is mainly industry access.

Dr. Sox: Leslie, you're next.

Dr. Francis: This is actually a related point. I mean, I took Barbara's initial question to be about what gets to the level of having a national coverage decision be made about it, not how once you decided to make a national coverage decision - this really speaks to once we've decided to make a national coverage decision, how does it work. And to me, and I notice in your list you didn't - you got to individual patients, and I didn't ever hear patient advocacy groups. But I think one of the things that we really should be concerned about is what about, you know, what gets something to the level that a national coverage decision gets made about it. Are there ever requests that don't get acted on? How much help does - I mean, industry has money and the interest in getting a technology out there - particularly where there are small groups of patients or where there isn't an effective patient advocacy group.

I mean, I know the Alzheimer's Association, for example, has been very thoughtful and careful about trying to bring to the attention of CMS some of these questions, but I don't know whether there are other patient groups or patients that just don't get represented in the decision to look at the local carrier policies.

Dr. Sheingold: I would just make the point that in the process the way it exists now, anyone can formally apply to have an issue looked at as a national coverage decision. If I recall, and Perry, if you're still here, the deep brain stimulation issue was brought to us by an individual beneficiary.

Dr. Bridger: Maybe since we've discussed deep brain stimulation, I'll just take a minute to explain how it came to us, because I think it's a good story and it's relevant to the conversation here.

Dr. Barry Green was the requester. Dr. Green is a Medicare beneficiary, is a Parkinson's patient who resides in Texas. He wanted to get the surgery, thought he was eligible, his physician thought he was a reasonable candidate, and he was told by them that the Medicare contractor in Texas at the time did not cover the surgery. Dr. Green then approached our regional office and had some correspondence with them, and they notified him that indeed, there was no national coverage determination about deep brain stimulation, it was at contractor's discretion, but a process existed for anybody to make a request for a national coverage determination. And they gave him our office's information.

And Dr. Green then contacted us by telephone and followed up with a formal letter, and we worked with him informally to assemble the package that's necessary for us to consider something to be a formal request, and we accepted it and the process moved from there. Since then obviously, Dr. Green being the

requester, we've worked with him throughout the last year throughout the MCAC process and so forth.

So I think that's a good example of an individual beneficiary and the system sort of working the way that it should work, and him having access to making a request and the process moving forward.

Dr. Sox: I think the next is John.

Dr. Ferguson: I didn't hear you mention it, Steve, but from what I recall of the old TAC, there was a cost factor that HCFA or CMS, which keeps excellent marvelous databases on how much money they're spending, very useful databases, and every once in a while something would come up because they found lo and behold, they were spending a lot of money at it. And this could come either from the central place or maybe Wade Aubry, who has been a local carrier as I recall, as a member of the TAC, that the carriers would say hey, this seems we're getting a lot of bills for this device or this procedure, or this drug, why don't we take a look at that and bring it up to the national group. So the cost played a role in what was going to be looked at.

Dr. Sheingold: I would say that a little bit differently.

Dr. Ferguson: I know that CMS doesn't like to talk about cost.

Dr. Sheingold: Yeah. I think the fact that an increasing number of bills were coming to the attention of carrier medical directors and they had direct access to the TAC certainly was a route by which we decided to take on national coverage determinations. And that same process would still work through the more formal request process we have now, although I think it's - I know it's the formal CMS line, but it's not so much the cost when it's a growing procedure as it is an indication that something is growing rapidly and there is really scientific issues about the risks and benefits that we ought to look at.

Dr. Ferguson: I'm not denying that is part of it.

Dr. Sox: Wade's next, and then Barbara.

Dr. Aubry: My recollection of that is that it was seen in the context of impact on the program. If they were receiving a lot of bills, particularly if there was controversy over the science, if there was variation in coverage across the country and it had the potential for a significant impact on the program, that that was a factor in determining whether to have it reviewed at TAC, the old committee. I wanted to point out that there is a local medical review policy, LMRP web site, that has local policies posted. I believe it's lmrp.org.

Ms. Anderson: It's .net.

Dr. Aubry: Is that right? Lmrp.net. And there are a number of deep brain stimulation policies up on the web. And it might be helpful for future topics if there was some sort of, as background information,

some sort of review of that as part of the committee materials, to have some sort of review of that as part of the committee materials, to have some idea of how many policies there are out there, and any pertinent information from that. So I think that would be useful background information.

And as part of Susan Foote's study at the University of Minnesota, the LMRPs will be evaluated as well as the survey of how medical directors use information and how the state carrier advisory committees work.

I did have a question. I'm interested in the national coverage determinations that are done outside the MCAC and whether there is any public input to that at all, or are those more administrative. Because certainly there's a lot of opportunity for public input in this process, but a very small number of technologies. What happens to that large number of NCDs that are done outside of the MCAC?

Dr. Sheingold: I'll say a few words and let Sean follow up. Actually there is a substantial amount of public input on every decision we make. We basically have an open door policy at CMS. As we say on the front page here, the MCAC, besides providing us the expert advice on clinical and methodologic issues, does provide the formal forum by which that public input can take place, but informally there's a substantial public input on every decision we make.

Dr. Tunis: The only thing I would add about that is there is a lot of input. In general, all the professional organizations related to, that have an interest in a particular NCD will either ask to meet with us or have a phone call, or submit a letter. Generally if there's an industry particularly interested we will get input from them. You know, all of this, our whole process is kind of kept track of on a tracking sheet that's on our web for each decision that tells you exactly where we are in the process.

And the only thing, if you have suggestions you can make them, but it's a process that counts on people paying attention and knowing we're doing something, and giving us input, as opposed to having some kind of ability to broadcast to the Medicare community that we're doing something. You know, the time frames are fairly short, it's not immediately obvious how one would announce that a particular thing is going on, but we do - at least it's a step forward that we have it on our web site, and we do get a substantial amount of interaction. Again, it's probably what you could call a biased sample of people who have a particular reason to be interested, and we do a little bit of outreach but not a huge amount.

Ms. Richner: That was the genesis of our original discussion, was essentially the informal process that occurs and how will - what questions are going to be answered and how do you triage those decisions. It was essentially how do you send it to MCAC, how do you decide that you're going to take it internally, you know, how do we make sure that we get the right physician groups engaged, the right patient groups engaged, how do we make sure that all the right clinical data is available to you at various stages of the manufacturing process or the academic process of whatever. So all those things, you know, we discussed those and it was essentially well, is there a process that perhaps that every time you have, you know, an informal or a formal decision is going to be taking place, is post it on the web, that there is - also the questions that you want answered are also posted on web so everyone can have input into how do you

define those questions.

I took that from Dr. Ferguson. You mentioned this several meetings ago. You said it all comes back once again to how you define the question of evaluation, and that is still a concern to us in terms of your informal decision making process, and how are you defining the question, what comparator, all this kind of thing that we're discussing today. And it's still not formalized yet, so I think this is your attempt to try to do that.

Dr. Sheingold: I think that's a critical issue and once we post that we're doing this national coverage review, there is that comment period, and we take that kind of input throughout. But I think there is also a very scientific basis to formulating those assessment questions which is why it goes back to what I said, that those first two steps you see there, that the pertinent background information and the analytic framework that should be developed really gives you the scientific guidance on how to formulate those questions.

Dr. Sox: Barbara.

Dr. McNeil: Well, this is slightly different but I think I read the report, the first report that you talked about on the web to Congress. And what I read talked about the time course for making for making a determination and then once a determination was made, when coverage was going to take place, and I'm actually not sure how I found that report. I guess just in surfing I found it, and it was very interesting to me because I personally would love a lot more information on it. I was actually interested to know that after you and whatever, or we made a coverage determination, the time frame from that decision to implementation, there was a range of something like 100 to 200 days. There was a rather long time period. And I didn't get all of the reasons for that from the report, but I gather one of them was that in some cases a CPT code had to be created for that technology, and then another reason was that I guess language about, your language about what actually, how it would be covered also had to be written. But it struck me when I read it that that was kind of long, that 100 to 200 days might be long, and that maybe this group should think about how to shorten that, given that it might not be possible to do too much shortening, or perhaps it would be, on the time frame of how to make a coverage decision, because that also was something like 100 to 300 days.

I would love to have a discussion about what's feasible to cut down the time frame on either the coverage decision or the implementation of the coverage decision.

Dr. Sox: But the 100 to 200 days that you mentioned, that was the post-MCAC.

Dr. McNeil: Or post-CMS, whichever the three parts of that. Do I have that right, Steve?

Dr. Sheingold: Post-CMS, no.

Dr. McNeil: Is it post-MCAC?

Dr. Tunis: It's once the decision memo is actually posted on the web.

Dr. McNeil: Correct.

Dr. Tunis: Once the decision memo is actually written, the implementation is close to 200 days.

Dr. McNeil: Okay, it was something like that. I didn't realize it was that long.

Dr. Sheingold: And add that to the list that Sean had of maybe talks that could be presented at this meeting to understand that process. And I'm not an operations person but just in general, the payment systems of Medicare are updated on a quarterly basis and not in between. So when a decision is made to pay for something new or pay for it differently, it has to be made at that quarterly time period, January 1st, April 1st, and then you have to have a substantial lead time I the previous quarter in order to prep them to make that change. So that's where the time gets eaten up, not so much in making the decision but implementing that decision into the standard payment system.

Dr. Sox: John.

Dr. Ferguson: Steve, do you get external reviews for the issues that do not come to MCAC? Do you get Blue Cross or ECRI or other, you know, technology assessment reviews of the literature?

Dr. Sheingold: That is one pathway that can be taken.

Dr. Ferguson: Do you do it very often?

Dr. Sheingold: Yes. We can request that AHRQ purchase a technology assessment for us, either off the shelf or an original. I notice that Deborah is back there and she could speak to this as well. There is even external review of that external technology assessment.

Dr. Ferguson: On issues that do not come before the MCAC?

Dr. Sheingold: Yes. As I said, it could be -

Ms. Richner: And that's one of our concerns as well, you see, because essentially they can request a technology assessment, CMS can request a technology assessment, and it wouldn't have this kind of public review in a sense, unless you have committed stakeholders once again, patient groups or whatever to know that this is happening and to actually have some input into the process and the design and the methodology and the question, and on and on and on. So that was one of our concerns, how can was make this a more public and also more efficient process in a sense.

Dr. Sox: Do you want to follow up, Barbara?

Dr. McNeil: A little bit, I think. I was at a meeting last week of a bunch of managed care executives and medical directors. And one of the comments that they made was the following one. When they are asked to consider a technology for reimbursement, or use, they can go to a number of different sites, and I guess ECRI is one, Blue Cross is one, and is it Winifred Hayes? I've forgotten what the other one is. But anyhow, what they said was that the Hayes one tends to be incredibly conservative, largely making approvals only if there is a randomized trial. And that ECRI tended to be a little bit more liberal. And Cochran was one of the other sources that they used.

But the bottom line was, several of them said that they had four or five possible sources of outside technology assessments that had been created in a formal way by somebody, one of these groups, and that frequently the came to conflicting recommendations, largely on the basis of their willingness to accept data less than Level I or AA, or whatever we want to call it. So it would be interesting to actually, I think maybe look at, and this may be well beyond our scope, but to look at a given technology and to see how several different sources have rated it. I didn't even realize there was so much variation in the sources for a given technology, let alone the simple science.

Dr. Sox: Okay. Ron Davis is next, and then Bob.

Dr. Davis: Well, I have a question and then a comment. My question is if I understand correctly, when anything moves forward in a matter that's subject to a national coverage decision, there's an update on the web site; is that right?

Dr. Sheingold: That's correct.

Dr. Davis: I think the difficulty that many of us face is that we don't have the time to be going back to the web site every few days to be looking for updates. And so I want to make a recommendation I made probably a year or two ago in some MCAC process that we were going through, and that is that there be a way to sign up on the web site to get e-mails any time a change is made to the web site. Any time there is an update, that a new issue is being considered by CMS, or an evidence report was commissioned, or an evidence report was received, or something was put on the agenda at MCAC, and of those updates on any issue, benefits issue that you're considering, there ought to be a way that is technologically easy for people to just put their e-mail address on a sign-up thing on the web, and get an automatic e-mail alert any time a change is made.

And ideally when your IT people make the update on the web, maybe they fill out a sentence that says such and such an issue was updated with a referral to a committee or whatever. So there might be a one sentence statement about what the update is, and that would become part of an automated e-mail to anybody who signed up on this thing. I mean, we may get 5,000 people signed up for this, and then they know when they need to go back to the web site to see what the update is. And then you don't have to keep thinking, do I need to go back and check.

Dr. Tunis: I think we have it, I think you're all on it, so I'm not sure why you're not getting those, because we actually established a list serve and put all the MCAC members on the list serve, and we have several thousand people on the list serve. It may be that we have an old e-mail address for you.

Dr. Ferguson: I do get some stuff on mostly the final ones.

Dr. Sox: Actually one came in yesterday, but it's the final new coverage decision. It doesn't talk about when there's any new information on the web site or the development of any technology, but final coverage decisions that are posted then go to the list serve. Go ahead, Janet.

Ms. Anderson: Actually there are several levels to our list serve, we try to make it as user friendly as possible. I was part of the development so I feel comfortable speaking on it. You can have new alerts, which is pretty much everything, that's the catchall. There's new decisions, which is what we signed the MCAC up; we didn't want to inundate you with information because I think we send you enough. And then there's also an area on there, I think, for just when we open a pending decision. So there's actually levels and if you'd like us to raise the level so you get everything, we can certainly do that. And you can choose. You can go in to the web site, it's on our home page, and change it on your own.

Dr. Sox: Maybe you should just send out an e-mail reminding everybody of their options.

Ms. Anderson: We can do that.

Dr. Sox: Great. Bob is next.

Dr. Brook: I was going to make a suggestion. The title of the memo you gave to us is referral criteria to us, and we've sort of turned this around and said can we help you in an advisory sense to streamline, make it more efficient, reduce the variability, reduce the bias and increase the fairness of the whole coverage decision. I can't figure out whether we're less than a half of one percent or one percent, but I think on that scale we don't make much substantial difference about anything that we've been arguing about.

So my question, or at least my suggestion to you guys is, is there a way that you want to use this committee to do this? The very least that I've heard from this conversation is that somebody needs to provide an evidence based report of what's going on here. You need an external group to do that. And whether that's the Institute of Medicine, whether that's somebody else to do that, or whether you want to use in a role to help play that, I for one would be willing to spend two days not being described what the process is, but if you really want advice on all of these things that we've given you gratuitous advice on today, if you really want advice that would have some likelihood of affecting your process, I think that that, it's obvious that there are a large number of people sitting around this table that are interested in doing that.

Dr. Sox: Well, if I could just comment on that, it's not just this group wanting to be useful, but it's the

whole process constituting more than a tiny fraction of all of the coverage activities. One advantage of having a list of coverage activities and which ones MCAC is involved with is to create some sort of way to communicate about whether we're in fact being asked to weigh in on issues that meet whatever criteria you establish, or whether there's only a random relationship between the characteristics of a particular coverage question, whether or not it gets referred to us.

So I think the next person is Bob.

Dr. Murray: Steve, I have a question. I recall about, I think it was in May of 1999 or May of 2000, CMS back when it was still HCFA published a description in the Federal Register of the criteria that it would us for accepting proposals for NCDs, the criteria for when and how it would accept them or not accept them, refer them to local contractors. My question is, is that still effective? Has that been changed, or can we rely on that guidance?

Dr. Sheingold: I think you might be referring to the April 1999 Federal Register notice on the process.

Dr. Murray: Probably.

Dr. Sheingold: Yes, that is still effective. It is in the process of being updated per some of the provisions that were in the Benefit Improvement Protection Act a couple years ago. But as you see it now, it's still effective.

Dr. Murray: Wouldn't that answer some of the questions that have been asked around the table, how does the process work, what are the deadlines and so on?

Dr. Sheingold: That does have that information, yes. What it may not answer as much is how outside of that formal process from the public, as Sean mentioned, how do we internally try and start a program to track technologies and bring them to the national coverage determination level.

Dr. Sox: Leslie, I think you're next.

Dr. Francis: I just wanted to follow up on Barbara's point. One of the things that would be tremendously helpful to me as a panel member is to know more about how the TEC assessment entity gets selected, because I worry sometimes when I see, well, the same group seems to have done the TEC assessment for everybody, and I worry about the level of independence. Or whether what we're in effect seeing is the same theoretical framework for doing TEC assessments, the bar is set high or the bar is set low. When we read repeat packets essentially, you know, I've seen the Blue Cross one, then I've seen the more recent one, and sometimes it's the same group doing it for both. And so I'd really like to have a little more discussion of that whole process.

Dr. Sheingold: We have a number of centers we can work with through AHRQ's contract with the evidence centers, and of course we work with them closely to make sure we're getting high quality

technology assessments. You do see a lot of assessments from a very few well known bodies. I think the critical factor is how rigorous and how consistent and transparent their methods are, so that when you come to the case that Barbara mentioned where you have the same technology and five different assessments that may say different things, you can track it back to their methods and look at the quality of the methods used to draw the conclusions rather than the conclusions. That's what we're working closely with AHRQ on now.

Dr. Sox: Linda, did you want to comment?

Dr. Bergthold: I think it would be helpful just as kind of a summary, talking with Ron here, to get something from you all sort of describing the list serve, how to get on it, a little bit about the background of the process, if you could just send that out to us.

But I will say that I stumbled onto it. I am on the list serve. An e-mail came to me and it sort of had a funny title and I almost deleted it and I thought it was just another whatever. Then I opened it up and it described all the pending decisions, and one of them was about something I knew something about and wanted to weigh in on, so I sent an e-mail to the staff person who was very responsive, who sent me an e-mail right back and told me exactly what to do, what was needed and when the deadline was, and I just thought that was an amazingly productive way for me to get involved. And it just, the problem is I sort of stumbled on it, I didn't know it was there. So maybe that's one of the issues.

Dr. Tunis: So we will go back and send out a note to everyone on the MCAC about the list serve, how to get on it, how to change your level of frequency and detail of the information that comes out. And you know, we did institute it at, I believe it was Dr. Davis's suggestion about a year ago.

Dr. Garber: That was your reward, Ron.

Dr. Tunis: Anyway, let me just make one comment on something Dr. Francis said, and then I want to sort of move to one of the particular questions we want to get some feedback on. In terms of the selection of who does the technology assessments, in large measure when we decide to get an external technology assessment we work through the Agency for Healthcare Research and Quality, and Dr. Zarin and her staff. They will try from amongst their network of evidence based practice centers to find the center that either has some content background in that area or some particular methodologic expertise. But we don't particularly weigh in on the selection of who's going to do the assessment, it's not a CMS decision but an AHRQ recommendation which we're not generally in a position where we would want to decline that, and I don't know that we've ever done that.

The other thing is to realize that the TEC assessment itself is one of the pieces of information that you all get to be part of your consideration when the issue comes to the MCAC. You also hear from members of the public, the interested parties, so presumably what part of this process is supposed to do is allow you to determine through other input whether or not the assessment has some bias or comes to some conclusion that's not legitimate. And whether or not the process is fully capable of doing that, you

know, that's another question. But it's not as if you all are sort of, you know, have to be driven by the conclusion that the technology assessment comes to. It's just one piece of that overall process. So you might have some suggestions about how to make sure that the discussion is full enough to give you a basis to determine if the TEC assessment is somehow biased or defective, but if you want to comment on that, that's great, and then I want to hopefully go back to Steve's memo.

Dr. Sox: Did you want to comment on a specific point?

Dr. Brook: Yeah.

Dr. Sox: Then it's Tom's turn and then we'll turn it back to you to redirect the discussion. Go ahead.

Dr. Brook: Through the practice centers the agencies, we're one of them, which means I've got to disclose that at Rand, the agency basically has a series of topics, about 60 or 70 of these reports, and there may be more of them. Most of them are just thrown out there. In other words, because they cover something that people want to summarize the evidence for. This has a specific use that affects millions of dollars or more, and that is it affects the coverage system. It sounds like to me that this needs to be interfaced a little bit better than the standard AHRQ process the way you're describing it. Because when I read the evidence report and your committee questions and report, it was unclear to me in places about whether the evidence center actually did what you needed them to do. And so I'm getting some sense of who controls, where does this go and all this kind of coordination issue but again, Sean, the question - I want to put this in a bigger question.

I threw out, you know, something to you. But the fact that nobody answered that question, I'm taking away the belief that you haven't decided yet in CMS whether you want us to answer all of these questions, or have all of these questions addressed that have been around the table or not. Am I taking that assumption to be correct?

Dr. Tunis: I think actually after you asked your question there were a whole lot of comments.

Dr. Brook: But that's from the panel, that's not from the Agency.

Dr. Tunis: No, no. I'm saying I just haven't had a chance yet to respond and you know, it's a complicated question. I think the answer to the question is yes, I think this committee ought to get engaged in those things and we ought to have a fuller discussion, and you certainly ought to understand better where this committee's function currently sits in the whole universe of things that are coverage policy, you know, all the things that you have discussed, and where it might most properly evolve to. So I think we need to have that broader discussion. We at one time were hoping to have such a discussion in sort of a retreat but I think the retreat concept raised lots of issues about FACA, so it's probably going to be that we're just going to have that conversation in a regularly scheduled EC fully open meeting, no sort of retreat context, and just kind of live with some of the limitations that come from having to do these things where every word is on the record, which isn't always the quickest way to get to a sensible end point. So

again, we've got a record of this meeting, we're going to go back and tick off these issues, and find the most effective way to move forward.

Ms. Richner: The general consensus is that MCAC sort of works in a parallel universe and that, you know, that CMS essentially works and they're taking our advice in terms of our evidence requirements and this kind of thing, they're taking the guidelines and using them internally, and that is very much appreciated in terms of what we provide you. But we're really seen as being almost irrelevant to the entire coverage process because of the inefficiency, that we're been PET 'r us for the last two years, that's all we've been doing is essentially evaluating PET every single time we meet.

So the thing is, it's like what do we do as a committee and how can we be most beneficial to the process and make ourselves relevant and important. I mean, we had an original intent to be an advisory committee for very difficult technology problems and that really sort of hasn't happened, other than PET.

Dr. Sox: Tom, you've been waiting.

Dr. Holohan: Changing the subject slightly, when you talk about external TA assistance might be considered, and in the context of Barbara's comments about the various groups that do usually for money, technology assessments, I know CMS is aware of this but virtually every other industrialized western nation has a central insurance program, central health coverage, and not all but most have very robust governmentally sponsored or governmentally implemented technology assessment programs. And there is general international agreement now. There is an international association that I know CMS recently joined that has attempted to set up international standards for evaluation of technologies.

And the specific question that I'm asking with that as background is, would CMS consider providing either for themselves or the decision processes that don't directly involve the MCAC, or to the MCAC, technology assessments performed by and in other nations that address the specific questions we might be interested in? In other words, would you accept the Swedish Institute of Technology assessment report?

Dr. Sheingold: I think I'll give an informal answer and let Sean follow up. We look at all of those. We are becoming more and more active in INATA and are interacting with them more and more on those type of collaborations. How we apply specifically, say to Swedish technology assessment, I would presume we would have to assess the assessment for its applicability to the U.S. practice of medicine. As you know, there can be differences, and I think that could be one thing that this committee could do in examining evidence. Sean, do you want to -

Dr. Sox: Deb, I want to call upon you please. Deborah Zarin, from AHRQ.

Dr. Zarin: When we get a technology assessment request from CMS or even before it's formal and we're talking about it, the first thing we do is send it out to the centers and say does anyone have a recent technology assessment on this topic. We also have our other favorite spots to look at, Cochran and other

places. So the first thing we do is see who else has evaluated this. We look at their technology assessment to see how relevant they are to the exact question that CMS is grappling with and we use them whenever it's at all relevant. So we try very hard not to reinvent the wheel.

And for many of the PET issues we used the Australian reports and some other countries have done reports on the same topic. We are trying hard to get arrangements where we can also share technology assessment capacity that's ongoing, so for some of the other technology assessments we've done, we know of other countries that are currently grappling with it and we try to share information there. And that's always summarized in the technology assessment so if you look back at some of them, you will see that.

So just to clarify that and to respond to something Dr. McNeil said, we are aware of some of the other groups that you've mentioned and that are at least purported to have various biases, and one thing relevant is that in AHRQ's role by statute and by tradition at least, we don't in any technology assessments that we do or that we commission, we don't draw the line. We never make a coverage recommendation and we don't say whether something is good enough or not good enough to be recommended for coverage. That's on purpose so that sometimes if you look at some of the Hayes reports, ECRI, Blue Cross/Blue Shield, the summary of the data is all the same. They're drawing the line at a different place and we don't get into that business, which is sometimes frustrating, but obviously there's a reason for that.

Then to just respond to what you were asking, Dr. Brook, is that the technology assessment program at AHRQ is managed within but separately from the general EPC program, so that any request from CMS to us goes through my office. We've been having a main relationship with the New England Medical Center's EPC over the last year just to handle the capacity. Moving forward now, we have three EPCs that will be doing technology assessments mainly. We might use other EPCs when there's a special reason to, or overflow capacity. We're going to be meeting soon with those three EPCs, that's evidence based practice centers, and CMS and us to try to clarify exactly what is wanted to be in the technology assessment, the format, procedures and all that. So we're moving towards having it be more uniform but by having three EPCs we hope to avoid the problem you raised also, in terms of at least a perception of bias or some normal way of doing things.

Dr. Sox: Barbara.

Dr. McNeil: I think this is on the topic. One of the things that I think would be really useful is the extent to which we can cut down the time it takes to get to a coverage recommendation. And what I felt happened on several of the PET studies, and this gets to your PET 'r us comment, Randel, is that some of the questions that you, CMS was asked to address by the group that made the request, in my view didn't stand a chance of coming up with a positive answer. Because just from a clinical perspective, they just weren't going to make it on the basis of the very quick but really critical review of the available literature.

And for you to make a new decision on those required an enormous analysis to prove what was really

the obvious, and an example of that turned out to be I think the axillary node question about breast cancer. Just on the basis of understanding the clinical component, looking at the data there was no way that you could make a yes to that. On the other hand, to make a no you had to really spend a lot of time, and the time that was taken on that took away from anything else that might have come before this committee.

Now maybe three practice centers will get around that problem because you can have multitasking but even with that, I wish there were a way that you didn't have to deal with frivolous, and maybe frivolous is too strong a word, but questions that really weren't burning or were likely to be a no, and could get right to the more controversial ones or the ones that benefited from input from a group like this. Or even in your second level where you had just the external technology assessment.

Dr. Zarin: I'll just make a comment and then flip it over to you. One of the ways that I think about that issue is that sometimes very quick literature search and analysis, say within a week, makes the answer clear. And the next six months, or four to six months has to do with what I call bulletproofing the reports and there is a question of does every single - does the answer for every single question have to withstand that level of scrutiny.

Dr. McNeil: That's what I'm asking, exactly.

Dr. Sox: Well, just to sort of review where we are, we're at the end of the scheduled time for this topic. We actually haven't got around to critiquing the criteria and I apologize for having not kept us a little bit more on task. What I think we ought to do, though, is take a break now, resume at the scheduled time of 12:45, and we'll start off by basically asking Sean to help us help him to get these criteria right.

We will have a little bit of time for public comment and then we will go on to the afternoon topics. So we will resume in an hour.

(Lunch break from 11:47 a.m. to 1:00 p.m.)

Dr. Sox:: This afternoon I'm going to ask Ron Davis, in just a minute, to make a conflict of interest statement that we agreed that he should have made a little earlier. Then Sean is going to lead the discussion of a critique of the criteria for requesting external TA assistance, and specifically the criteria for getting help from MCAC.

And I'm going to leave the chair for a few minutes to get set up for the second part of the afternoon session, which will be discussing a rather more specific question than we had originally planned. Instead of a general discussion about evaluating evidence concerning diseases that are quite rare, we are going to focus on the following question: Should MCAC change its approach to evaluating diagnostic tests to deal with the specific instance of rare diseases? And I'm going to help us to get through that discussion by going through some slides about how we actually do it, and then we can kind of focus on how if any or how if in any way we should alter the guidelines to deal with tests for rare diseases.

So Ron, could you just make your brief statement?

Dr. Davis: Thanks, Hal. When we went through conflict of interest disclosures this morning I neglected to mention that my mother has advanced Parkinson's disease. She's never had DBS, she's never sought DBS to my knowledge, and I have no idea whether she would even be a candidate for DBS, but I just thought I ought to get that on the record.

Dr. Sox: Thank you very much. Now we will turn to a critique of the criteria for referring a topic to MCAC, and I'll let Sean take this while I get set up.

Dr. Tunis: All right. Hopefully now that people have had lunch they will be a little more tired and won't be quite so feisty. But what I thought we could do is try to finish a little bit of a conversation on the specific issues raised here, which is given that we have heard lots of issues raised about which things actually come to the attention of a national coverage process, and we talked about the local process and other issues. But given that certain things now do actually come to our attention for national coverage decisions through one mechanism or another, we still have to make the decision of whether we're going to do the literature review internally, whether we're going to ask AHRQ to work with the EPC to do a literature for us, and whether we're going to review that issue to the coverage advisory committee.

And as you know, a decision to refer something for an external TEC assessment or to the advisory committee adds a significant chunk of time to the turnaround time for making the decision. It can be, you know, four months or so for coverage, it can be four to six months in the best case scenario for a TEC assessment. It used to be a lot longer than that on the external TA side. And so, we did want to hear from folks what sort of factors ought to be part of our deliberations and making those decisions and whether, you know, what we've listed here, which are fairly general criteria, are kind of sufficient guidance.

Our hope is to actually have these criteria published on our web site to say here's how we make the decision that we're going to do an external TA, or here's how we make a decision to refer something to the MCAC, so that's what we ultimately want to craft this document into. So we're looking for feedback on this sort of initial draft on that. And we're going to maybe talk about this for another 15 minutes or so, until Hal gets set up for the next part of the afternoon. So that's open for discussion.

Should I be chair or no? Actually, Dr. Brook, you're the vice chair. Do you mind chairing this part?

Dr. Brook: Sure.

Dr. Garber: May I speak, Bob? (Laughter.) You know, Sean, one of the observations I'd make is that it's kind of hard to envision expanding MCAC's role in the general national coverage determination process if our actual deliberations take a lot of time. I'm not referring about the delay until we get something, I'm talking about our deliberations themselves.

So our typical approach has been a panel meeting takes a whole day to do one technology. The Blue Cross/Blue Shield process is usually something like 12 or more in one day. And I think the challenge before us is the MCAC process is such that some panels, I think have never met, and none of them have met frequently, and there's a lot that gets done more efficiently if you meet often enough to maintain your familiarity with the whole process and so on.

I would like to see some kind of approach that would enable and expand the role for MCAC where we could preserve the public input but have things move along more quickly. And I'm wondering whether there's some technologies where there could be written input and some time at the meeting, a limited amount of time for public discussion. To make sure we definitely get the public input, which I think is crucial, but are able to handle something more akin to four or five or six technologies at a time. I think that would make the whole process move more quickly, and there may be some issues that you would like MCAC to evaluate that don't really require the same kind of in-depth approach that we've used to date.

Dr. Brook: Can I just call the committee's attention to the criteria. I think what Sean wants to be answered is the second page, and you're going to post this on the web site, the bottom half of that page is the criteria you want our input on right now?

Dr. Tunis: And actually the top part of the page too, both of these, criteria for sending something out for an external TA, and then the criteria for -

Dr. Brook: Then can we divide this into two parts? Are people comfortable with the first part, or does anyone want to make any modifications either to one or two under criteria for requesting external TA assistance?

Dr. Francis: It does seem to me that the question about disagreement and the interest of seeing different perspectives, the kind of question Barbara raised, is something that isn't captured in one.

Dr. Brook: Would you want to suggest a modification to that?

Dr. Francis: I don't know how to put it but it could be something like the need for a methodological evaluation or methodological disagreement.

Dr. Brook: Well, should we restate - if you go down to one under the next bar, is that more like what you had in mind?

Dr. Francis: Well, the first question was sending it out to TEC assessment and part of what - I mean, it's not just if it's complicated but if it's contested, it seems to me to be a reason to send it out.

Dr. Ferguson: Or controversial?

Dr. Francis: Controversial.

Dr. Brook: So do you want to add controversial after too extensive?

Dr. Francis: Uh-huh.

Dr. Brook: Or too extensive or controversial to be reviewed internally within a reasonable time frame?

Ms. Richner: I have a question and that is, since we just received this this morning - I mean, I think I have a lot of concern about many issues about this in terms of how they're triaging or the criteria, how they're essentially - I don't know what this document is for. Is this going to be something that we're going to decide that this is a permanent document?

Dr. Brook: Sean.

Dr. Tunis: Yes. We want to eventually post publicly on our web site criteria of factors that we consider in determining something needs an external TA or something will be referred to MCAC. So this was just to get a start on the conversation.

Ms. Richner: The first draft.

Dr. Tunis: It's the first draft, it's to kind of get the juices flowing. And if what you all come up with today is you know, if we want to kind of lay some thoughts out there to sort of send you back and get another draft, you know, we're fine to do that. We don't have to come up with a final draft today.

Dr. Brook: What's the sense of the committee. Would you like to right now - Wade?

Dr. Aubry: It seems to me, Sean, what you're asking for is some comments on number two for selected national coverage decisions, external assistance will be applied and what criteria should be used for that.

Dr. Tunis: Uh-huh.

Dr. Aubry: The other parts, the bullets under the first part of the page are basically what you would do for anything, but for some of those you would need additional assistance and you would send them out. So I think there should be more under number two than is written there.

Dr. Brook: Two, you mean the second half of the page?

Dr. Aubry: Yes.

Dr. Brook: The criteria for MCAC review?

Dr. Aubry: No. I'm on the first page, number two. That's basically the item that deals with which things should be sent out.

Dr. Tunis: Yeah, so I think if you just go to the second page and ignore the first page -

Dr. Brook: The second page has the actual criteria on it as far as I could tell.

Dr. Tunis: At the top we've got number one and number two. That's as far as we've gotten in terms of the criteria for reviewing something for external TA. So that's our starting point right there for what you were just looking at as number two on the first page.

Dr. Brook: We could handle this process the following way. We could all give comments, go around the room, give comments on - give people two minutes to read this, put comments in the record for everyone who wants to make a comment to Sean. Sean could take it back and revise this draft. Or we could try to get to final wording today. How much time do we have to do this?

Dr. Tunis: We probably ought to take 10 or 15 minutes.

Dr. Brook: Ron.

Dr. Davis: My comment is, I don't think we should consider what we do to be final wording. These are just recommendations to CMS.

Dr. Brook: So what if we just went around the room and everyone just gave Sean some of those comments so we had some on record, would that be wonderful? And then you're going to come back with another draft for us, is that what's going to happen?

Dr. Tunis: Yes.

Dr. Brook: Is that an acceptable process?

Dr. McNeil: I'm not sure that going around the room is the most efficient way. I think if somebody has a comment that it might benefit from another comment that may not necessarily be in order.

Dr. Richner: Are we commenting on one, just number two on the second page? Because I have a lot more beyond just number two on the second page.

Dr. Tunis: Why don't we maybe, just to keep things a little bit organized, start on the criteria for requesting an external TA, and then we'll have a separate discussion on the criteria for removing to MCAC? Is that all right?

Dr. Brook: Do people care if we start and just go right around, and you can pass if you have nothing to say. And if somebody else wants to interrupt, just raise both hands and that person will be given right to add something immediately. Linda.

Dr. Bergthold: I was just trying it out.

Dr. Brook: No, you're first.

Dr. Bergthold: I'm passing.

Dr. Brook: Ron.

Dr. Davis: I don't have any comments.

Dr. Brook: Wade.

Dr. Aubry: Pass.

Dr. Murray: I'm concerned that the - I'm not concerned. I'm confused by when this process is going to kick in for when there are laboratory issues. You know, the laboratory generally generates a great deal of very specific data with precision analysis, sensitivity and specificity of the various tests that are proposed. And I'm not aware that a TEC assessment is the best way to go under those circumstances. So I guess my only comment is, I'm still a little bit confused by this as it applies to the laboratory, so probably I should have just said pass.

Dr. Brook: Well maybe, Sean, you're going to have to clarify how this applies to the laboratory, that's what Robert asked.

Dr. Tunis: Right.

Dr. Brook: Randel.

Ms. Richner: Okay. We're just looking at number two on the second page?

Dr. Brook: The first two on the second page.

Ms. Richner: The external TA assistance may be considered when the scope and magnitude -

Dr. Brook: Yes, and the second one.

Ms. Richner: And in terms of how that is defined in terms of scope and magnitude, I'm concerned as

well. You know, how are they going to decide when something should go for, you know, even further back for a technology assessment. That's my first concern.

The second then is who are they going to be - which body are they going to refer it to and why, and what is their criteria in deciding which technology assessment group that they will be using. What is the extent of the technology assessment they are going to be requesting. So I mean, I have a lot of questions associated with that.

Dr. Brook: Alan.

Dr. Garber: I have a big comment. I just think, Sean, it would be a good idea if this is correct, after the semicolon in one, to put "or". Is that the correct interpretation, that either condition one or two would be when you would consider external assistance?

Dr. Tunis: Right, it is implied or, yes.

Dr. Garber: Okay. I would make it explicit. Thank you.

Dr. Brook: Leslie?

Dr. Francis: I just want to reiterate that I think that it's not just when it's complex but when it's controversial.

Dr. Brook: Tom.

Dr. Holohan: I don't have any real comments. I'm presuming that external TA assistance includes looking at currently available assessments that are basically on the shelf prepared by reputable -- that would be considered part of external TA assistance, okay. Then I have no comments.

Dr. Brook: Barbara.

Dr. McNeil: I have two comments. The first is, I would actually separate out from one and two, and maybe make it two, or two-three, the issue of conflicting reports from well-respected technology assessments. So I would have the scope and magnitude, blah, blah, blah, and then I would have number two saying something like existing technology assessments from outside groups lead to different decisions, the sort of thing that Steve was talking about earlier. And then I would leave three the way it is.

The only other comment I'd make is I think I disagree a little bit with Randel on her concerns about scope and magnitude. Personally I don't think that this group should micromanage too much how CMS does things. I think if you say scope and magnitude and reasonable time frame, to me that's good

enough. I'm willing to buy what you know is within a reasonable time frame and what is within the scope and magnitude of your internal resources. I just don't see how we could add value to that.

Ms. Richner: The reasonable time frame issue, though, was again, you said 200 days seemed too long but maybe that's not too long for the Agency considering their workload. So I mean, reasonable time frame in itself should have some kind of a definition. At least with the FDA we know exactly what the timing is and those are very well defined, and we don't have that.

Dr. McNeil: But that's a different, that's a more specific comment.

Ms. Richner: Right.

Dr. Brook: Is there anything else the two of you want to say?

Dr. McNeil: No, I'm finished.

Dr. Brook: Okay, John.

Dr. Ferguson: I think basically this is just more or less like the part we used to have for having a consensus conference, but one thing is missing. There has to be sufficient data out there, there has to be sufficient data and it has to be controversial. That is, there is not conclusion from all these data, there is debate on what these data mean. And that's when you need the external, to send it to somewhere else. You know, if there's three studies of four patients each or something, but I might say that in the laparoscopic colososectomy business there was one or two large series of 1,500 patients or something like that, but they were a series with no controls, or historical controls.

Dr. Brook: Michael.

Dr. Maves: The only thing I would add is actually to build on what Leslie and Barbara said. I had a criterion number three which was objectivity, which I thought would be particularly important in areas where you had conflict. Perhaps you have two camps, one is sort of pro, one is con, so it's a very difficult decision. Or it may be a controversial area. Sending it out for an external technology assessment, as opposed to CMS doing it internally, might in fact lend an ear of objectivity. There are research that emanates from Federal agencies, NIH, NASA, so on, where there may be a call as to whether this has been objectively determined, and so I thought that might be one other criteria that you might want to put on.

Dr. Brook: Anybody else, especially people that passed, first of all, that wanted to say anything or comment or provide any other advice?

Dr. Ferguson: I guess we shouldn't say anything about the governing bodies of this country, should we, having their say in this, like Congress.

Dr. Brook: President.

Dr. Ferguson: The President, Senate, and so forth.

Dr. Brook: Tom?

Dr. Holohan: Just to comment with respect to Randel's question about when you decide to do a technology assessment. The Institute of Medicine published a document in 1992 at the request of what was then AHCPR that gave parameters as to when it was appropriate to do a technology assessment. Even though that is ten years old now, I don't think it's obsolete, and I think it gives some very very good guidance.

Dr. Brook: Sean, do you have enough guidance, or do you want something else from us on this one?

Dr. Tunis: Thanks.

Dr. Brook: It gives you a number of things to consider.

Dr. Tunis: It's a good start.

Dr. Brook: Anything else? Let's go to number two, the second part of this. This is when to refer things to us. We'll start the other way going around the room. Michael.

Dr. Maves: I am going to pass. I am pretty comfortable with this.

Dr. Brook: Okay. John.

Dr. Ferguson: Well, I think there ought to be something that says when you guys, CMS can't decide for yourself after having the external review and looked at all this stuff, then that's when you need yet another assist from an external body, namely us. And I'm not sure, does it say that there? And no internal decision could be made, or something that says why would you go beyond what you did up above. You need more help, you can't decide. Isn't that what you're doing now? No?

Dr. McNeil: No, that's not what that says. They're saying sometimes they won't even start with number two, they will go right to number three, or option number three because there was controversy, complexity, specialized methods, blah, blah, blah. So I don't think - isn't that right?

Dr. Ferguson: I'm just saying that when the issues are complex, numerous, however you want it, that no internal decision could be made. Am I reading this wrong?

Dr. Tunis: I think it's whether or not -

Dr. Ferguson: Because why would you refer it unless you couldn't make up your mind?

Dr. Holohan: I think we've been talking about that for a while.

Dr. Bergthold: Yeah. I mean there's reasons for needing or wanting more public input. There's reasons because Congress made us do it. And I actually think that, I mean, the MCAC for six months last year was totally doing this because of PET and the political stuff, and we spent an inordinate amount of time on it.

Dr. Brook: Linda, would you like to see some words added, or at least Sean think about words like, the reason they go to us, as opposed to just one alone, is because of public controversy, something that relates to public controversy, constituency pressure or something else in there? I don't see that in there. Is that what you're saying?

Dr. Tunis: Yeah, we could be more explicit about controversial. I think that probably is the tip toward, you know, heavy interest and powerful constituencies or whatever, you know , the need for more.

Dr. Brook: And John said, Sean, if I could paraphrase it a little, he was concerned that the distinction in reading the language between one and two at the top of the page and one and two on the bottom of the page is not crisp enough to allow somebody to decide why you would do one of the above ones versus the second one.

Dr. Ferguson: That's very good, Bob.

Dr. Brook: And if he's confused, that means most everyone that's going to read it is confused.

Dr. Davis: Can I chime in on that point?

Dr. Brook: You have to raise both hands.

Dr. Davis: On the point that John makes, there's a potential problem in doing the top part of the page and then if CMS can't decide, then going to the bottom half, i.e., MCAC. And that is that the process that we laid out a long time ago envisioned that one or two panel members would be involved in the preparation of the evidence report. And if we do this sequentially, then the evidence report is produced before MCAC and an MCAC panel is involved, so that would violate the process that we outlined earlier.

Dr. Brook: Ron wants to make sure that this is not a sequential document but a one or the other, for your consideration. That you make the decision up front which of these two you're going to use, and it's not clear from the document that that's what you're doing. Barbara.

Dr. McNeil: Actually that's what I was going to say. I didn't think these were sequential, I thought they were designed to make a prospective decision whether it was option internal only, internal with outside technology assessment, or outside technology assessment plus MCAC. I thought you did that ahead of time.

Dr. Tunis: I think our current thinking too, you know, closely describes our current process. We try to make the decision of whether or not to go to MCAC within 30 days after accepting the request. At that point in time the presumption would be almost always we would also be getting an external TA. There might be some situations in which we would do it with an internal document but at that point then, we would try to follow this process of consulting with the MCAC in the development of the TEC assessment as well. So that would be how I would see these as linked, that the decision to go to MCAC would automatically trigger an external TEC assessment.

Dr. Brook: Hal.

Dr. Sox: I have a question for Sean. One of the criteria that aren't on here is enormous potential aggregate cost, such as intercardiac defibrillators in congestive heart failure. Do you think that should be on there?

Dr. Tunis: We usually phrase that as enormous potential public health impact.

(Laughter.)

Dr. Sox: Should we add that?

Dr. Aubry: Can I offer a suggestion?

Ms. Richner: It's here.

Dr. Aubry: I would suggest in the preamble where it says have a major potential impact on the health of beneficiaries, I would add or the Medicare program, or raise important social, legal or ethical issues. So major potential impact on the health of beneficiaries or the Medicare program, which I think would be in line with my earlier comments and also address your issues.

Dr. Sox: Bob, you're still - I think you should carry through and complete this discussion as the chair.

Dr. Brook: We've had this debate before. There is some people that want you considering using the C word if you can actually use it legally. I mean, you've got to make that determination, and that's some advice. You've seen how Wade referred to it and you saw Hal's suggestion, and all we can give you is advice. Tom.

Dr. Holohan: For the sub-subheading one in the second half of the page, issues not related to a specific national coverage decision can be referred to MCAC, I don't understand these circumstances. I'm not sure what the MCAC would be expected to do in terms of presentation, public discussion and clarification. I just ask that that be a little bit more explicit.

Dr. Tunis: We will make that clearer. That's actually what we're getting to, our next topic basically fits this.

Dr. Holohan: I thought that was a good segue.

Dr. Brook: Leslie:

Dr. Francis: In addition to the controversy point, I also wanted to ask you the question of whether you think you should be quicker to move to MCAC, when your ex ante guess is that the decision if it were in house would be no to coverage, or when you think that the decision in house would be yes to coverage. I'm asking that as a question.

Dr. Tunis: I honestly think that, you know, by the time we've reviewed things at 30 days, I think if we had a fairly clear substantive yes or no, I just don't think that's been a factor. I mean, I'm not sure. Maybe you're suggesting we think about whether it ought to be a factor.

Dr. Francis: I didn't know whether that had at all been a factor, that is whether in a controversial case you thought you were likely to resolve it no versus likely to resolve it yes. My own view is that you shouldn't let that play a role, but I wanted to ask the question.

Dr. Brook: Randel, do you want to comment on that

Ms. Richner: Alan was next.

Dr. Brook: I meant Alan.

Dr. Garber: I will concede the floor to you, Randel.

Dr. Brook: I'm sorry, Alan.

Dr. Garber: No, no.

Ms. Richner: There are some concerns that you know, as the Executive Committee we came up with panel operations that don't seem to match this. That's my first concern.

My second thing is that the charter of MCAC doesn't necessarily coincide with this because it essentially

says in this first bullet here that the purpose of MCAC is to allow for additional expert and public input on topics that are highly complex. Well how did we evolve to that? I mean, it is not in the charter, it's not in our public operations or our panel operations and recommendations that we came up with. Unless we want to go back and revisit this. We spent a lot of time going through this and if you remember, it was very difficult to go through looking at the expert opinion and having a panelist involved in the key questions for the evidence report to address.

Dr. Brook: Randel, let me ask you a question. I'm now confused.

Ms. Richner: Yeah.

Dr. Brook: This has nothing to do with the panel process, this has to do with what gets to us. Is there something in the charter that you're concerned about?

Ms. Richner: Yes.

Dr. Brook: What?

Ms. Richner: Because essentially MCAC is to look at technology assessment broadly and I, you know, once again this was brought up earlier, that how can we be a nimble process where we can be an expert public advisory committee for a general array of technology problems that come to CMS.

Dr. Brook: So your response is you're figuring that by doing this, by putting this - you're objecting to this whole thing.

Ms. Richner: Yes.

Dr. Brook: And you're objecting on the basis that it limits the flexibility of the MCAC process?

Ms. Richner: Yes. And well -

Dr. Brook: And therefore, it produces some additional limitations to our charter so that we should have the right to look at simple noncontroversial things that they know the answer to?

Ms. Richner: Well, that is what we were discussing earlier. I mean, when and how, what is the purpose of MCAC essentially. I mean, this is what this is bringing up.

Dr. Brook: That's correct, what is the role of MCAC in this controversial process. And you're not comfortable yet with this language at all, that we've reached a point about what role, if any, MCAC should have in this process.

Ms. Richner: Right.

Dr. Bergthold: Can someone - do you know the charter? Randel, it sounds like you probably know it by heart. Can you tell us what it says?

Ms. Richner: Linda, I don't know it by heart.

Dr. Bergthold: I mean this little sentence part of it. I haven't looked at it for a long time.

Dr. Garber: I don't recall that it had language that said what MCAC would and would not consider as to the level of detail that this draft document has. And so I don't think there's any real possibility of contradiction at this point, I'm not saying that it's not possible. And I think what Bob just said is quite correct, the operations document for this committee doesn't address the same topic. Now ambiguities about timing, sequencing, that needs to be worked out, obviously, but I don't see any inherent contradiction here.

Dr. Tunis: In your draft guidelines on evaluating effectiveness, the first sentence in the preface says, "HCFA convened MCAC to provide advice on scientific and clinical questions regarding coverage." That's not exactly, you know, the - and I guess what this is going to is, that's a fairly broad statement of function and we're trying to get to the subset of issues upon where MCAC could function most properly, functionally, or helpfully provide that advice. But again, Randel, if you have a -

Ms. Richner: I just remain concerned. Once again, it's about the public input in terms of how and when that CMS takes - you know, gets a request, and how do they triage it, and I still don't think that it's clear from these criteria. So until we know exactly how you're going to decide that something comes to MCAC, I think it's sort of hard to move forward with this.

Dr. Tunis: And are you actually making the suggestion which, it's an interesting suggestion, that the actual decision of which things to refer to MCAC should itself be a public process of comment, you know, that we would propose that something go to MCAC and then -

Ms. Richner: No, I'm not suggesting that. I think once again, there was some comment earlier that as you get requests in whether formally or informally, that they're posted and that people are aware of it, and they have some kind of limited comment period to know that that's going to be occurring.

At lunch we were discussing that like the GAO, you get constant e-mails that there's something that's going to be coming up for evaluation and it's very very public, and people are aware of that. So whether or not - I mean, I still think that there's a black box in terms of what are you evaluating and why, and how can we assist in that process more effectively. You know, MCAC still has only evaluated, as you know, probably six or seven technologies over the last few years so, is that the most efficient way to use MCAC? I don't know.

Dr. Brook: Robert?

Dr. Murray: As a document of criteria, I'm comfortable with it the way it is. I prefer to see it broad and somewhat give CMS discretion. It's filled with words like might be considered and may be referred, and I don't think that CMS is going to allow MCAC to put handcuffs on it, that must be referred, will always be referred. I am in favor of words that broaden what is eligible for referral, and I think this is fairly broad as it is.

Dr. Brook: Wade.

Dr. Aubry: I agree with the last comment. I think there may be some confusion about whether something can be referred to MCAC if there hasn't been a formal request for coverage. And I think that there may be issues that are appropriate for MCAC and CMS may want to refer to MCAC, but there hasn't been a formal request for coverage. Also, there may be, in fact there have historically been cases where a formal request for a national coverage decision has been made and then withdrawn. And in those situations, it may be important to the Medicare program or important for consistency across the country, et cetera, for MCAC to review that, even if the requester withdraws it.

And so I think there should be discretion for CMS and I wonder whether it might be appropriate to put a note at the end to the effect of the following: MCAC referral may occur at the discretion of CMS without a formal external request for a national coverage decision. So that's going to be my recommendation.

Dr. Brook: Ron.

Dr. Davis: Just one quick comment in response to Randel's comments. I agree with her that the words "the purpose of" implies that this goes back to our charter and we don't have the charter in front of us. But assuming that this goes beyond what the charter says, then I would suggest some other language like CMS generally uses MCAC to allow for, or CMS preferentially uses MCAC to allow for. And then we can debate whether it ought to be that way, but at least it gets us away from implying that this is our charter.

Dr. Brook: Yeah. Linda?

Dr. Bergthold: Well, just a little reminder of the historical. Do you remember one of the first meetings where Grant I think, or somebody put up on the wall that flow chart. Remember that flow chart about how decisions were going to come in and how they were going to be triaged? I haven't seen that flow chart for - I don't know where it is, but one of the things we talked about at the time, as I recall, was what would be the criteria for deciding whether something comes to MCAC. And we then settled on exactly where we are now, which was things that were complex, things that were, you know, that were contradictory, where evidence was contradictory, and we had a couple of other factors.

So you know, I think at the time there was a great desire on the part of some members and perhaps industry to know exactly what was going to happen. It was really clear then that nobody could prescribe it in advance, we wouldn't know. And now what, two years later, three years later, I don't know that we know any better. We've got a body of cases that we can look at and sort of extrapolate from those as to what has come to us, and I think we can see that things have come to us that had complex scientific evidence, things have come that were political pressure, and I don't know what other categories. Can anybody think of any other categories? I think it's those two buckets.

But I'm perfectly comfortable with CMS using us as they see fit.

Dr. Brook: Hal.

Dr. Sox: Just use us.

Dr. Brook: Hal, do you have any thing else? What the committee had asked is it sounds like, you take a crack at rewriting this, since it's your request of us. And Sean, do you want to pick some people to help you look at that and review it before the next meeting or would you like to just present it to the group as a whole at the next meeting, or whatever your druthers are.

Dr. Tunis: Yeah, we will I think - I know Randel has a particular interest in this, so we'll certainly if she's willing run some language by her, and if there's anybody else that in particular would like to give us some feedback.

Dr. Brook: I would ask Ron and Leslie if they would also be willing to, with the different points of view that I heard to look at this thing. I think if you can get all four of them to agree, we will be in good shape.

Dr. Tunis: And again, you know, it may be that it's not possible to do a lot better than what was come up with two years ago. I think there's continued to be interest in our further specifying and you know, that's why we wanted to move forward with putting a document out there to get some feedback with the intention of posting it. And even at that point where we post it, my guess is we will continue to get feedback to further refine and specify.

It is -- I'm sure you appreciate this -- the decision itself of whether or not to refer something to MCAC or for an external TEC assessment, like I said, is by itself an extremely important decision for the individual companies who are affected by that. It means, you know, six more months prior to a decision and during that period of time functionally what happens is many of the contractors put their own policy development processes on pause. So it's a nontrivial decision and I understand the interest in us being more explicit and more accountable for it. Dr. Brook: I just want to say one brief thing and then I will turn it back over to Hal. The only thing that this really argues is whether it's time now to do some sort of an evaluation of the importance of MCAC in the coverage program, if we have enough maturity in this. Somebody ought to do that, which might include both looking at some quantitative data from the office as well as questioning the members of the group and trying to write a helpful document of what's next

for MCAC. As opposed to, you know, tens years from now we'll have the exact same format and be around the same time.

So maybe we have enough experience that it's worth doing that. Somebody should try to do that because you're hearing that part of the confusion in recommendations to you relates to what really is the mission of this organization relative to the other things that you're doing.

Anyone else who has anything else to say about either one of these topics? Okay, Hal, it's yours.

Dr. Sox: I just would comment that the evaluation ought to take into account how well the procedures that we've developed are actually serving your purposes. Here's the preview of coming attractions. I just want to let you know that Sean has to leave for about a half hour at 2:00 to be in a conference call, and at that time Steve Sheingold will move into this seat so that he can advise me.

The topic we're going to turn to next is how to deal with rare diseases, and we're going to pick a subset and we're going to pick a subset to try to make this really concrete, we're going to pick a subset of technologies, and that is diagnostic tests for rare diseases. It's partly to focus it, it's partly because there may be something on the horizon for the coverage group in relation to tests for rare diseases. And in general the process I want to follow here is to go over our procedure for dealing with diagnostic tests and kind of stop at several points and say is this element of that procedure sensitive or not to whether the disease is a rare disease, with the goal of eventually getting to the point of being able to decide whether we need to change our procedures for diagnostic tests to take into account the fact that the disease for which they're intended is a rare disease. So that's the goal.

We're going to have public comment from Dr. Robert Wahl speaking on behalf of the Academy of Molecular Imaging, and he's going to be telling us a fair amount about some of the characteristics of a particular subset of rare diseases, which are rare cancers. And we can then listen to what he has to say and see whether that's really going to modify the thinking that we come up with. So that's a preview of coming attractions.

I'm going to lead this discussion from the podium with slides being advanced in the back of the room.

Dr. Francis: Are the slides going to be up on the screen?

Dr. Sox: They should be.

(Discussion off the record.)

Dr. Sox: Sit wherever you want. So, could we go to the first slide please? This is a reminder of our procedures for diagnostic tests, and you'll remember that we hope that we will get high quality studies that provide direct evidence that results in new tests improved healthcare outcomes. An example of that would be a randomized trial in which women are assigned to usual care or regular mammography where

you then measure the death rate from breast cancer. So that would be an example of direct evidence.

Most of the time we're not going to find evidence like that and so then we have to evaluate the indirect evidence that the new test makes a difference, and the procedure that we developed was a mixture of empirical evaluation of the accuracy of the test together with some rather simple and straightforward disease modeling. Next please.

So first let's just remember together the technique for evaluating a diagnostic test. Basically the types of results you can have are true positive results, that is, positive results in patients who got the disease in question, false positive results, and then their complement false negative and true negative results. Next please.

And the measures of test performance are sensitivity and specificity. Sensitivity is simply the number of patients with true positive results divided by the number of patients with the disease. Specificity is the number of true negative results in patients who don't have the disease. Next please.

Just a reminder that a simple measure of test performance that relates test performance to what tests do, which is to change probabilities, is the likelihood ratio. The likelihood ratio is given by a base theorem which is illustrated here. Post-test odds equal pretest odds times something called the likelihood ratio, and the likelihood ratio for a test that's positive is sensitivity divided by one minus specificity. The likelihood ratio when a test is negative, which tells you how much the odds change when the test is negative is one minus sensitivity divided by specificity. Next please.

Now just a reminder. Whenever you make a measurement, you have some uncertainty about the true value and that can be expressed by putting a 95 percent confidence interval around the sensitivity, specificity, or the likelihood ratio. And doing that is important because it tells you, it enables you, particularly with the likelihood ratio, to get some idea of just how good the test or how bad the test might be under the conditions of sampling that you've got given the size of the sample. So we have to take into account not only the point estimates of test performance but also the degree of statistical uncertainty about those. Next. Now here's how you measure test performance. To determine whether the test is positive or negative you have to do the index test, which is to test your study. And to find out whether the patient has disease or not, you do what is called gold standard test or diagnostic reference standard, which basically tells you the patient's true state, and that gives you whether disease is present or absent. And in the ideal study of test performance, you do both the index test and the gold standard test on a large number of patients, enough so that the confidence interval around your measurements are relatively narrow.

Why don't we stop here just for a second to think about, since this is one of the key measurements that we use in evaluating the effectiveness of diagnostic tests, whether the fact that the disease is a rare disease is likely to affect our ability to evaluate the test performance in detecting that disease. Does the fact that a disease is rare make it impossible or very difficult to make this fundamental measurement upon which our evaluation of the test depends? Any thoughts about that, anybody have an opinion?

Leslie.

Dr. Francis: Well, it affects the numbers, but also the technology is changing. We get shifts in the technology itself.

Dr. Sox: How would that situation, how would the fact that technology changes -

Dr. Francis: That makes it even harder to get the numbers.

Dr. Sox: Okay. So it's harder to get numbers in a timely fashion when the technology takes a step upwards.

Dr. Ferguson: If the prevalence is really low, it takes longer to get any good results.

Dr. Sox: Well, if the prevalence is low, the test could still be very good at picking up the disease. It is true that sometimes there's a relationship between clinical suspicion of a disease being present and the accuracy of the test, and in patients with low pretest probabilities, reflecting few clinical manifestations of the disease, often those patients have smaller cancers that are harder to detect. But I personally don't think that the fact that a cancer is rare necessarily makes it more difficult to detect with a diagnostic test. I don't agree with that but others may wish to contradict me. Bob.

Dr. Brook: I think it has a profound effect in the following way. To do it with the same set of numbers that you do with a common disease, the fixed costs of doing it increase dramatically because you have to have institutional cooperation, and cooperation among people that may not want to cooperate to do it. So the cost of doing it, of getting that data, probably increases at some number that we haven't predicted.

And the benefit to the group that is actually providing the new test or service, or use of it in somebody is decreased because there is so few people to market it on. So you get an imbalance between costs and benefits of doing this work. And I think that issue needs to be addressed head on. And we need to address a social equity question, do you really believe that we the public ought to subsidize this process even if it means taking money away from more common diseases to advance treatment in those areas. That I think is the implication of using the same standard of care. You could argue that we don't need the same standard because the cost implications, I know I'm not supposed to use the C word, and the health implications are so much smaller because it affects so few people, and therefore you are not going to pursue this strategy for that reason. Just like if you drive on a rural highway, the likelihood that you will get to a hospital when you've run off the road is different from if you drive on an urban highway system, and we accept that, and we're not willing to subsidize helicopters constantly flying over rural highways to pick you up off the street and take you to emergency care.

So I think that's the question that you raise here if we use the same standard. If we relax the confidence limits or our certainty, instead of using a .05 value use something broader, then you could take this back to the same level of resource investment, or at least closer than what's required in a common disease.

Dr. Sox: So what you're saying, Bob, which I agree with, is that the social cost, particularly in relation to the population benefit, changes when you have a rare disease, because it's a lot more hassle and expense. You're talking about cooperative studies all over the country to get enough patients to make accurate measurements of test performance, and they benefit fewer people.

Dr. Brook: And if you were in business, you would do everything possible to avoid putting resources there than putting resources elsewhere. So in essence, by developing the same standard we are guaranteeing inferior scientific evidence for people with rare disease, even though it can be captured, unless there's a public subsidy of that process, I think. I think that's economically a correct argument.

Dr. Ferguson: Isn't that what the drug companies and the FDA do, to give them the money to ramp up?

Dr. Brook: Yes.

Dr. Sox: So your objection, Bob, is one of social cost and benefit, not the fact that we can't do it, because you could it.

Dr. Brook: It's just a matter of extraordinarily rare diseases, but for most rare diseases -

Dr. Sox: Once a year diseases.

Dr. Brook: For most rare diseases the answer is you can do it, but you won't do it unless there's a change in policy. So the association between can do it and will do it becomes extraordinarily great.

Dr. Sox: On the other hand, if CMS decided it was important as a matter of public policy to do this, they could fund studies and we would get the data.

Dr. Brook: Absolutely.

Dr. Sox: Barbara.

Dr. McNeil: Hal, I'm sorry I was out for a minute, but did you define rare?

Dr. Sox: It hasn't been defined, and I don't know of any formal definition for rare. John.

Dr. Ferguson: When I was at NIH, by definition we used 200,000 prevalence in this population, or less.

Dr. Sox: And that's prevalence, not incidents?

Dr. Ferguson: Prevalence.

Dr. Sox: Prevalence countrywide.

Dr. Ferguson: 200,000 or less.

Dr. Sox: Tom?

Dr. Holohan: There's one practical limitation since the example we're talking about is measuring test performance in a population with a rare disease. To the best of my recollection, and this is not my area of expertise, but for a fixed sensitivity and specificity of any given test, as the prevalence of that test in the population goes down, the ratio of false positive to true positive results goes up, to the point where at - and you don't have to get down to 200,000 in our population of 280 million in round numbers, where the ratio of false to true positives may become 5 to 1, 10 to 1, 20 to 1, and the issue then is what do you do as your next step to separate the false positives from the true positives.

D. Sox: Well, you phrased an interesting question that I don't think we have the answer to and that is, are we talking about tests to detect the disease in patients with a particular symptom, in which case your point is right on. The probability of disease if the test is positive is going to be lower if the pretest probability is really low.

Dr. Holohan: Right.

Dr. Sox: On the other hand, if we're using a test in patients who have the disease and it's a question of staging or monitoring, as was the case for example in some of our work with PET scanning, then the prevalence of the condition that you're trying to detect in patients with the disease could be low or it could be high because you're starting with patients who have the disease, rather than a population with symptoms that suggest the disease might be present.

Dr. Holohan: In that case by analogy, the prevalence would be the probability of metastatic lesions in a patient with the disease known to exist to begin with.

Dr. Sox: Yeah, and that could be quite high. Barbara, did you have something?

Dr. McNeil: Hal, all I was going to say is if we are going to discuss rare in terms of this decision matrix here, I'm not sure that prevalence is the relevant number that we want. Because if we're talking about detection, which a lot of this is about, or staging or whatever, it has to be incident cases. We are seldom looking at something that is going to affect all prostate cancer patients on the diagnostic side. It might be on the treatment side there's a better chance of that, but on the diagnostic side you have certain tasks at certain points along the treatment pathway so that the number varies depending upon where you are in that pathway, I think.

Dr. Sox: So are you talking about an application that tests in patients who might have the disease but

have a symptom that could be present?

Dr. McNeil: Well, we were just talking about Alzheimer's disease for example, at the last meeting. To say that the prevalence of Alzheimer's disease is 5 million and I don't know what it is now but say it's 5 million, for the sake of argument. That really wouldn't have much relevance toward our decision to go forward with a formal technology assessment or not, because the number that we were applying that test at that time would have been suspected patients, and it would be probably far less than 5 million. So all I'm saying, I was just trying to get a sense of rare, and it's just a little bit had to know what that is when the number has a different meaning depending upon where we are in the screening, diagnostic, initial treatment or subsequent monitoring process.

Dr. Holohan: I think in fact we've all been talking about pretest probability of disease in different ways.

Dr. Sox: Well, you're right, although I was trying to draw the distinction between pretest probability in patients with symptoms that might indicate the disease is present which would tend to be quite low in patients with rare diseases, because the base prevalence or anchor point for us to name the probability would be very low. And the situation where the patient has got the disease, in which case the prevalence could be anything and wouldn't depend on the rarity of the disease in the general population, because the population of concern now is the patients with the disease. So in that sense we're using those two things differently, but if we're clear on whether we're talking about patients that got the rare disease or patients who are suspected of the rare disease, then we can avoid getting mixed up, I think.

Dr. Brook: But I think Barbara, I mean this is all feasible. Let's say a rare disease probably is somebody, in a homogeneous group of patients you're going to make some decision about, but I would guess it's 1,000 in the United States, or 2,000, because if you take that prevalence, reduce it by incidence and then begin to work through at what stage these people actually need that test, you may be talking about hundreds to a couple of thousand people. And if you say you need to put together a couple hundred people to have a reasonable answer to this question, a few hundred people, you're talking about coordinating doing that in a year, coordinating half the country or a third of the country, or a quarter of the country to do that. That's expensive.

And so if we had a web - you know, using web-based technology and a whole slew of other things you might be able to get there, but it does require the standardization, the manuals, the protocols and all this stuff so you can pool the data. And I think you're right. I just think that we ought to make the caveat that you're only right if somebody's willing to invest the resources to do this, and it's not likely to be the company or the business that is manufacturing this device.

Dr. Sox: Okay, should we go on? Does anybody else disagree with Bob? I certainly agree with him. Okay, so let's charge on. Next slide please.

To decide whether you've got valid measures of diagnostic test performance you've got to be concerned with certain characteristics of the studies themselves, the studies that measure test performance. And this

comes from our own interim guidelines. The study population should be consecutive patients with the same chief complaint, same presenting complaint. The patients who get the index test should all get the reference test. If that does not happen you end up with a biased sample of patients who get the reference test and it's the patients who get the reference test who really constitute the study population, and you tend to underestimate sensitivity and specificity when many patients with a negative index test don't get the reference test.

A third criteria in our patients, the people who interpret the index test and reference test should be blinded to all other information. That is, patients who interpret the index test shouldn't know the results of the reference test or the other clinical data, lest that bias them to make a call in a sort of toss-up case in a direction that would be consistent with the other information about the patients, which would result in overestimating test performance.

And finally, the reference standard needs to be a valid measure of the disease state at the time of the study.

And I guess one of the questions we should ask is, if these are part of our criteria for evaluating studies and deciding if they are valid, do these change because the disease - do our criteria for an excellent study change because the disease itself is a rare disease? Any opinions about that?

Most of these, at least the two middle ones are really about sort of the logistics of doing the study. It might be harder, as Bob says, to police a nationwide sample to make sure of consecutive patients, where you really observe the criterion of consecutive patients. Although again, if you were studying patients who had the disease, you wouldn't define it the same way. Yes, Barbara?

Dr. McNeil: I was actually going to say, Hal, about a number of these studies and I would guess but I can't say for sure, that meeting the second criterion is probably easier with a rare disease than with a more prevalent disease. Just because people are going to be less sure that the clinical constellation really goes along with whatever it is, and they are more likely to go on to a reference standard than if it's prostate cancer, where they are going to buy into the CT ascertainment of nodes alone rather than the biopsy of pelvic nodes or something like that. So I think it might be easier. This is not to say I disagree with Bob's point, I'm just saying for that particular criterion rarity probably helps.

Dr. Sox: Any of these where rarity might be a problem, make it harder? Barbara, what do you think?

Dr. McNeil: No, I don't think so.

Dr. Sox: Okay, let's go on. A second step in the evaluation of test performance is to decide whether the test you're evaluating -

Dr. Brook: Can I just add one other point, Hal?

Dr. Sox: Okay.

Dr. Brook: In all of the work that we get presented to us we get sort of efficacy work presented because the investigator, there's been a lot of effort that the single person controls this and you're really now almost getting effectiveness data presented. So the likelihood that we will approve something in rare diseases becomes less.

Dr. Sox: The likelihood of what?

Dr. Brook: That we will approve that this diagnostic test is valuable becomes less. If there's any difference than in the efficacy work because if you depended on people all over the country using goodwill and cooperation to do this, it's going to be very hard to standardize the MRI. They may have an older machine as opposed to the newer machine, and they wouldn't have done it as carefully because it's really not their study, and so it produces more useful information to us, but it's a different level of information than efficacy data. It almost becomes the same as effectiveness data.

Dr. Sox: What I thought you were saying and maybe you did is that with a nationwide cooperative study, following all those criteria for a good study is going to be tougher.

Dr. Brook: But even when it's tougher, you're not going to get the same controls. It's like somebody doing their own survey, you know, asking the questions and monitoring it on a small sample in their backyard versus using a national survey firm to do it where you know that 10 percent of the data is going to be totally unreliable, and you still hope the results are valid. And you're going to find the same thing. You're going to get people submit things that are, if you look at even the cooperative studies of cancers that have been, you have all these problems with finding an investigator that did things poorly and whatever. You now have one patient per doctor around the country, and this becomes a problem. That's all I'm saying. It's not that it can't be done, all of this can be done. It just is a logistic nightmare. Barbara can spend the rest of her life doing this.

Dr. McNeil: Well, actually I'm not sure I agree with what you said, Bob. I think in theory what you say it sort of rolls off the lips, but in fact I'm not sure that it's true. I've done tons of cooperative technology assessment studies that would have been called efficacy studies because they were done in large teaching hospitals versus St. Elsewhere's in addition.

And two facts: One is that there's an enormous variability in reader accuracy across even teaching hospitals, so you get a wide range in sensitivity and specificity, much wider than you would have expected even in an efficacy study.

Dr. Brook: I understand that.

Dr. McNeil: So I'm not even sure that it could be even wider with effectiveness, maybe it could, but it could be pretty wide.

And the second thing is, I'm not sure that the accuracy of data collection bears any relationship to the size of the institution collecting the data or the academic experience of the institution collecting the data. We've had actually completely the reverse experience. In some situations smaller institutions have been much more careful because this is now their raison d'etre than larger ones. I think you could argue it either way so I think the jury's out.

Dr. Sox: Alan.

Dr. Garber: I'm not sure we will ever resolve whether the data is better or worse for rare diseases, but I did study one rare disease in some detail and that was Gaucher disease, and a lot of the patients with that disease would be treated at specialized centers part of the time. They might get their routine care at a local place but they would go to one of the major centers around the country maybe once a year, like if they had bone symptoms it might be Mass General. And in fact, I think they were likely to get much more standardized data collection than you'd see in a common disease. So you know, there are probably examples both ways. I just point out that there is nothing that I saw that suggested that the data collection would be any worse or that it would be harder to study. In fact I think those kinds of studies are much more well set up to do studies of test accuracy than the cites where we routinely perform studies.

Dr. Sox: Well, let's move on. Another issue is whether - oh, I'm sorry, could you go back please? Another issue is whether a new technology that's being applied to a rare disease is better than the technology that's already in substantial use. And one of the ways you can be better is to pick up patients that are missed by - one way that a new technology can be better is to pick up patients that are missed by the established technology.

There are two ways to evaluate that. One is if the sensitivity of a new test is a lot better than the established test, then it's pretty much a given that it's picking up patients that the established test is missing, and we saw that with PET scanning, comparing PET scanning to CT scanning for some of the applications that we studied earlier. But another way you could do it is to have the two tests be pretty similar in the number of diseased patients they pick up because they pick up different patients and therefore, they complement each other. And even though the new test has a similar pickup rate in disease, it's still worth doing because by doing the two tests you pick up more patients than by doing one alone.

And in order to evaluate the ability of the new test to pick up patients that the established test misses you really have to do both tests, as well as the diagnostic reference standard in a series of patients.

Now the question before us now is we'd have to say that that is frequently not done in studies of test performance. But to the extent that it is done, does evaluating a rare disease with a multicenter nationwide study make it less likely that we will be able to evaluate the sensitivity of the new test relative to the established test and its ability to pick up patients that the established test misses. I think

inherent in the fact the disease is rare and you have to do a multicenter study to get adequate numbers of patients.

Dr. Garber: I think Barbara addressed that earlier and clearly she didn't think the answer to that was yes. She thought it anything it was a little bit easier.

Dr. Sox: And I wonder, I wish he was here, because as long as you're doing a protocol that evaluates PET scanning against CT scanning as compared with some diagnostic reference standard like surgery, is it really that much more hassle because you have a lot of centers to do both studies as part of the protocol, than to have a single study. And while you could argue that it would be just because of the logistics problem, if you have a protocol that calls for two studies instead of one study, does that really make it harder to do the study.

The question, Barbara, that you missed was evaluating the complement degree to which tests complement each other, does the fact that it's a rare disease make that any tougher?

Dr. McNeil: No, it doesn't make it any tougher. It's tough, period. And the reason it's tough if we're talking specifically about imaging is the fact that you usually start off one of these comparative studies, say CT and MR, which was the example you gave, against pathology for prostate cancer or ovarian cancer, whatever it is. So you start off with two protocols, you standardize the CT protocol and you standardize the MR protocol, and ideally what you do is then randomize the patients. In theory this is what you do. You say okay, Barbara, you're going to get the CT first and Randel, you're going to get the MR first for your ovarian cancer, because we want to make sure that the results of one don't leak in some way to the results of the other.

Well, you can hardly ever do that just because of scheduling these days. And then the other thing that makes it difficult, but regardless of the prevalence of the disease, is the fact that a standard protocol is good for a study but frequently coming to patient care the radiologist who is doing the second study like the MR really feels in good conscience that he or she has to know what the CT showed so that they can fine tune the technique of the MR. They just cannot follow a blind protocol in the best interests of the patient.

So that makes the randomization all the more important at the beginning in terms of the order, and that requires a scheduling freedom. But more than that, it requires the good will of the investigators at the site, because you can always juggle a schedule. I mean, the reality is it's never as bad as it says and you can always juggle it. But meeting that criterion or those criteria is the same regardless of the prevalence of disease, it has nothing to do with whether you're doing 1,000 patients a year or 10,000.

Dr. Sox: Let me ask a quick follow-up question. I thought I heard you say that it's difficult to evaluate the CT and MRI independently of one another, blinded to the other, because the radiologists want to use the CT to help sort of calibrate themselves in respect to the MRI.

Dr McNeil: Right.

Dr. Sox: Now if the MRI was a well established test in that setting, wouldn't that -

Dr. McNeil: I think that people will frequently tinker. It's always well established but they like to tinker and they like to go back and do it again, or just do a little more contrast or something. So that while it's well established - well, pretend they're both reasonably well established. People do tend to tinker so what the ideal arrangement then is, here's the ideal situation but you have to deal with the reality of the fact that doctors take care of patients and in good conscience they really feel as if they need to know the results of the preceding test.

So how you get around that is the following. You say okay, we are going to do our best to randomize patients, we're going to do our darnedest to do that. And at the end of the day we're going to reread our films independent of the clinician who actually did the study. And we can reread them in one of two ways. We can do them separate in which you read this and I read this independent of the other, or we can say the real question is not MR better than ultrasound in detecting ovarian masses or staging ovarian cancer, not better than, but how much more does it contribute than ultrasound. So in that particular case, instead of just reading the two modalities separately, they would be read by experts in each of those fields, and you would have the MR reader of the ultrasound be sitting with the ultrasound person and say okay, this is what you saw, here's what we saw, now tell me what you saw. Then you get a cumulative reading that would show the impact of A on top of what is known to be B, versus A compared to B.

So you can't answer both questions readily because the rereading process is so incredibly labor intensive. You're not going to get anybody to go through it twice, so you have to decide ahead of time whether you're looking at A versus B or A on top of B.

Dr. Sox: If I can just get us back to the topic of rare diseases -

Dr. McNeil: It makes no difference.

Dr. Sox: Okay, let's go on. Next slide.

So then, the second sort of big question in this is to ask whether differences in test performance are important clinically, and this involves in our case some very simple qualitative disease models in which the first step is to calculate the post-test probability for both diseases and in the second test to say, is that difference in post-test probability going to be important for the management of the patient or in some cases is the post-test probability different enough from the pretest probability for either test to make a difference. And so the question is again, with rare diseases, are we going to have more difficulty carrying out this step? Next please.

So just going back for a second to see where we've come, we evaluate studies of test performance, we decide on the sensitivity and specificity based on those studies. We then calculate a post-test probability

for the new test and the established test and then we ask, does the test result with either one change the probability enough to change management and if so, with which one does it make a difference.

If both tests have good enough post-test probability to change management then you could argue that the new test really doesn't make a difference. It's only when one does better than the other and that one changes the probability enough to change management that you really have a compelling case for doing the test. Next please. So just to remind you of examples that we had in our own experience, PET scanning to detect scar recurrence of colorectal cancer. The clinical question is in an indurated area near the original scar tissue, are you dealing with a big scar or are you dealing with cancer? Next.

So the question is, does the negative PET scan lower the probability of it being a recurrence rather than just an exuberant scar enough to alter the decision to simply biopsy the mass? And what we learned in our evaluation of that problem is that the pretest probability that this mass represented a recurrence was about 70 percent. And we learned that the PET scan has enormously good test performance characteristics for detecting a scar recurrence. So we then used base theorem to calculate the post-test probability of recurrence. Next please.

And we plotted post-test probability on the vertical axis for every possible value of the pretest probability. Now for a pretest probability of 70 percent the curve corresponding to a negative test result which is the lower curve, has a post-test probability of recurrence on the vertical axis of about 8 percent.

So the question then becomes - next please - if the negative scan lowers the probability of scar recurrence in 70 percent to 8 percent, clearly you would biopsy if you didn't do the test and the probability was 70 percent, the question is would you biopsy if the recurrence is 8 percent and if you would, then why do the test? Why not just biopsy the patients without doing the test? On the other hand, if you decided that 8 percent was a low enough probability to simply observe the patient for a month or two and see whether the mass grows or not, then the test would make a difference because it would lead you to put off doing the test and maybe never do it if the mass didn't grow.

What we found in our discussion was that this Executive Committee was basically divided on whether or not they would biopsy at a probability of 8 percent or less, and so we eventually decided that since we couldn't come to agreement on whether the test result would change our management, that we should endorse PET scanning for this application.

So that's an example of how we use the sensitivity and specificity of the test which we got from the first stage of this process, and the pretest probability of disease, and our knowledge of the clinical situation to decide if the test would make a difference or not. And we argued that if the test wouldn't change your management then we shouldn't approve the test.

So now, the question is, does the prevalence of disease in the population make a difference in applying this particular modeling situation? And I'll just remind you that we have two situations to consider. One is where we're concerned about making a diagnosis of the rare disease, in which case the pretest

probability will be a lot lower when it's a rare disease than if it's a common disease. Or are we interested in the use of the test in patients with established disease, in which case the rarity in the population at large is probably not pertinent.

So, everybody got their arms around the question? Who wants to start the discussion? Alan?

Dr. Garber: Well Hal, think the first indication is presumably irrelevant to today's discussion because when we're talking about approving a test, whether a test should be approved for the diagnosis of a rare disease, it must be in a situation where the pretest prevalence is high enough that we could say ex ante it's being considered. So it's not as though we're talking about a disease that's a one in a thousand. If we're talking about solitary pulmonary nodules, that's what we're talking about.

But as I understand it, we're really concerned about its use in rare diseases, not for -

Dr. Sox: In patients with a diagnosis of a rare disease.

Dr. Garber: Either with the diagnosis or where the pretest probability is very high because you've already done some tests to narrow it down to a few things where this rare disease now becomes a very reasonable likelihood, that's my understanding anyway.

Dr. Sox: But wouldn't it be true, Alan, that if the disease is rare and you apply a bunch of information, history, physical, lab tests to try to raise that probability, that if you start with an extremely rare disease, your probability after gathering all that clinical data is going to be lower than it would be if the disease was much more common, because you're applying the same data.

Dr. Garber: All things being equal, if you're talking about applying the same sensitivity and specificity to each piece of information and applying it to a very much lower pretest risk, that's true. But I'm assuming that we're dealing with a situation where that whole stream of pretest risk and then the stream of tests and physical findings and so on, this creates the probability. Because otherwise we're talking about the use as a screening test essentially, and we're dealing with a whole different ball of wax. Then we're talking about things like should you use a test for a solitary pulmonary nodule, or should you use it for an abdominal mass, or this or that, and it's not really a question about rare diseases.

Dr. Holohan: That's kind of what I meant when I said in different ways we're all really talking about pretest probability. At some point in the diagnostic stream the pretest probability changes from a screening test to a - confirmatory is not the right word, but diagnostic.

Dr. Garber: You know, the cases where I could see this coming up is if you have a cancer that's know to be a rare cancer and you're trying to decide whether something is metastasis from that cancer or something totally different. There is clearly where we're talking about what is the sensitivity and specificity for this rare disease. That would seem to be squarely what we're often talking about with rare diseases.

Dr. Sox: Well, I'm not sure that when you're using a test that's a diagnostic test and the disease is, we haven't made the diagnosis yet, I'm basically not convinced by your argument that the pretest probability -- if you apply the same sort of information to a prior probability that's very low as you do for a prior probability that's higher, you're going to start with a lower pretest probability when you do the test for the rare disease, I'm pretty sure.

Dr. Garber: Well I think when we're talking about CMS trying to make a coverage decision about a test for a rare disease, it's hard for me to imagine how they can make a coverage decision that's contingent upon the results of that test. In other words, if you find X disease then this test is reimbursable but if you didn't, then it's not. So that's why I think we're not really talking about something where there's a lot of uncertainty ex ante about it. I mean there's obviously got to be a lot of uncertainty, but where the pretest risk is extremely low.

Dr. Sox: Barbara, do you have any thoughts about this?

Dr. McNeil: No.

Dr. Sox: Well, I guess I would contend that in a situation where we're using this test in a rare disease to stage it, test for metastases, that that's a pretty clear case where the fact that it's a rare disease probably doesn't make a difference in the model, and -

Dr. McNeil: That's what I was trying to say.

Dr. Sox: Yeah. So at least we can all agree on that.

Dr. Garber: Right.

Dr. Sox: It sounds like we're not entirely all together on this question about using it as a test to detect a disease in somebody with a constellation of history, physical and other lab data.

Dr. McNeil: Hal, an example here in the testing would be to use something to test for Gaucher's disease, is that what we're stylizing that against? Ultrasound for staging ovarian cancer or something like that, is that -

Dr. Sox: Or somebody who has a big liver, is it Gaucher's disease or just cancer from a metastatic disease from a colorectal cancer, which is relatively common. And I guess I'd assert that if the test performance is the same for the cancer and for the Gaucher's disease, the post-test probability would be higher for the cancer than it will be for the Gaucher's disease.

Dr. Garber: Right. But the coverage issue here, if you're talking about the abdominal mass what we want to evaluate and test for, is this a good test for the evaluation of the abdominal mass. This doesn't really

have anything to do with whether a rare disease is in the differential or not. It has everything to do with can you change management in a way that improves outcome. So I don't see where the rare diseases make anything remotely different in that situation.

Dr. McNeil: I agree with Alan.

Dr. Sox: Okay, good. Bob?

Dr. Brook: Well, I was going to say sort of the same thing. I suspect that most rare diseases are made by application of a gold standard and that you're not going to make incremental because the prior probability is so low and that our tests are so imperfect even if they're excellent, that we're making most of this by almost whatever is the gold standard definitive approach. So that screening, I mean it's unclear to me that the results of an improved imaging study or diagnostic testing will lead to confirmation that you have a rare disease without taking the next step. And what we're saying is that then most of the questions about rare disease is advancing care for patients with rare disease, that you've already identified they have a rare disease, and we've had two comments.

One is Alan's comment, and I don't know the answer, whether many people with rare disease go to a few places. I don't know whether that's true and I don't know if there's any evidence to support that. That would make it easier to do some of this work because they would congregate in large numbers and places, but you have a bias of who gets there versus who doesn't, versus having to do a sample that would be much more community based which would make this work tougher, so that while we would approve PET scanning in colon cancer because you could find enough of these people, you wouldn't approve PET scanning in a cancer that affected only a few thousand people a year, or looking at metastatic because nobody would ever put the data together.

Dr. Sox: Tom.

Dr. Holohan: Are we talking about detection of a metastatic lesion or mass as Alan referred to in patients with common versus uncommon cancers, or is it much more likely that CMS is going to be confronted with a claim that in malignant dysgerminoma, very rare tumor, I have a labeled antibody that will uniquely attack to metastatic lesions of malignant dysgerminoma. It has nothing to do with colon cancer, lung cancer or anything else. The test itself is designed to detect a very rare tumor.

Because the circumstances are very different then in terms of trying to determine what the pre and post-test probability of disease is and whether it in fact changes treatments.

Dr. Sox: Well, if the tests were like most tests, it wouldn't be perfect and it would occasionally attach to colorectal cancer, metastatic colorectal cancer cells. And so the prevalence of the other diseases that might lead to false positive results would be important to know. Barb.

Dr. McNeil: Well, I - I'm sorry.

Dr. Holohan: Maybe I didn't clarify it. Given all of that, a premise that an imaging test using antibody labeled isotopes that are "uniquely" designed to bind to rare tumor cells is different than the question of location of metastatic masses in a general sense in the way Alan described it. I think that the underlying biologic principles of the two tests are quite different. And I'm trying to get at whether the issue for CMS is the use of more or less conventional imaging that's used in common tumors, in rare tumors, or in unique or allegedly unique diagnostic tests used in rare diseases. Because I'm not sure that -

Dr. Sox: Alan, can you help?

Dr. Garber: Well, it would be nice if Sean was in the room, and maybe Steve can answer that, but the - I mean I think in principle we should be ready to deal with both questions. I feel fairly confident that the type of disease-specific test that Tom is referring to is probably not the number one issue right now. It's probably a more generic test that would be used to detect certain disease, or a certain metastatic disease. I think what Tom was talking about, you can imagine having a specific monoclonal antibody or some type of wave bolt for a battery of 20 extremely rare causes of metastatic cancer and saying everybody in whom we don't have a diagnosis gets this entire panel of tests.

Dr. Holohan: I hadn't thought about that, but that's a great idea.

Dr. Garber: No, no, but I mean that's where this would lead, right?

Dr. Holohan: Yes, you're correct.

Dr. Garber: I think that's the situation you were setting up, and I imagine today that's more a theoretical than practical concern, I could be wrong about that, but the first one is a real one that's for a rare cancer where you don't have a lot of knowledge about whether the metastases are due to that or how good the diagnostic test is at telling whether the metastases are due to that cancer. I think that's something that's a very live problem. Steve?

Dr. Sheingold: I guess I would try and pose this a little bit in the way that I posed it in the paper that you were sent in your packets. Having listened to all this, if in fact we have a use of the diagnostic for which we don't have the data from credible studies that we want to base a decision on, whether that be the sensitivity or specificity, or whether that affects the outcomes of having the test, are there cases in which we would look at something else like data from another use of that same diagnostic, and try to extrapolate from those data something about this use of the test.

Dr. Sox: Well, that's an important question but it's not the question we've decided to address in this discussion which was, do we change our criteria and our process for evaluating diagnostic tests just because it's a rare disease.

Dr. Sheingold: I know. I think that there's a slight crossover, though.

Dr. Sox: There may be and that may be something that we want to talk about at some point, but personally I'd like to kind of answer the question that we started out to answer, and if we have time we can deal with that one. But I'd kind of like to get to the finish line on this one rather than cut off in another direction, which might take more preparation. Barbara?

Dr. McNeil: Hal, you know what I would benefit from, and maybe this is jumping the gun a little bit. We have the issue of rare although we haven't really defined what rare is, and we've got a couple of different indications on the table and we're talking about whether or not we can get data. So one thing that would be very helpful would be if this committee had a couple of tables that said okay, for disease with incidents - name the tumor if we're talking about tumors - the number of incident cases each year is 5,000, just make up a number, and they're going to be spread out over hospitals. And pick ten hospitals, ten big hospitals and say how many patients are in those hospitals in a given year, get somebody to say okay, we're looking at new technology X versus new technology Y. Find out what the delta is in incremental performance between the old and the new. Do a sample size calculation about the number of patients that you would need to answer that delta in a clinically meaningful fashion and say okay, you need 232 patients.

And then say okay, 232 patients; can we possibly get 232 patients in fewer than ten sites?

Dr. Brook: The answer is no. I don't need to do the calculation.

Dr. McNeil: Ten is really - no, I think that's wrong actually.

Dr. Brook: There's 7,000 hospitals in this country.

Dr. Sox: Yeah, but how many are M.D. Anderson, Bob?

Dr. McBride: I'm assuming that we're not talking about you know, some really really rare rare thing that's like - but it's an empirical thing that could be determined. If we had a list of what we're talking about as sample rare conditions and we knew what the incident values were, we got approximate distribution of those cases at major hospitals. Because I feel like we're having a discussion that is intellectually very satisfying but the rubber isn't quite hitting the road because the rubber is going to hit the road here in terms of feasibility and if something isn't feasible, we can gab all we want.

Dr. Sox: Well, I think that's good advice for CMS to do as part of trying to decide how to respond to questions, or how to respond to requests for doing evaluations of rare tests.

Dr. Garber: Well Hall, first to answer your question, maybe I could summarize my take on this whole thing, which is our approach to diagnostic tests. As far as I can tell there's absolutely nothing that's different between rare diseases and common diseases in terms of this framework. The real issue is, are there bits of information that you might need to apply this framework in the context of rare diseases that

would be very hard to get, but that's separate. I think the framework stands on its own. There is nothing about this that is different between rare and common diseases.

But that does bring us to Barbara's question. Can we say anything useful about cases where - can we first of all even define when it would be difficult to get the data? And I just want to make one general point. If you have a mistake in assigning somebody to treatment because either you didn't do the test or you did the test and it was a false positive or a false negative, that has adverse consequences. That's why we're doing the tests in the first place. If we didn't need the test we'd either just treat everybody or not treat everybody, but the whole point of the test is there's something that makes people worse off if you make the wrong decision.

And let me just point out that, let's say that it is infeasible to get enough people to assess the accuracy of the tests and we say therefore, you don't need to meet the normal criteria. Then we are basically saying make your best guess and if people are harmed because they are actually making wrong decisions and we have no way of knowing about the wrong decisions because we haven't done the right studies to find out whether the test is giving us useful answers, useful information, we are really basically saying let the patients make the wrong decisions, more people will die, but because it's hard to do the studies we're going to just live with that. And to my mind there is a big difference between diagnostic tests and treatments in this regard.

If you have somebody with a rare disease and it's not feasible to do a trial but you know that the disease is frequently fatal very rapidly, they don't have a lot to lose with the new treatment that looks very promising and I think there's a much stronger case to be made for treating empirically in that situation. With the diagnostic tests I think there is the possibility of a very real and potentially very large downside.

So that's where I see - I don't see us being able to loosen requirements to at least get reasonable information about accuracy, regardless of how hard it is to obtain that information.

Dr. Sox: Well, I think we need to move on, and I'd just like to ask whether anybody wants to differ strongly with Alan in his take that basically the framework that we've described in these slides is reasonable to apply in the case of rare diseases. That's certainly what I've heard around the table but I think we ought to give everybody a chance to, anybody who wants to dissent from that because you know, we've had a very high level excellent discussion that was meant to help you decide how to deal with these questions.

So, anybody take major issue with what Alan said?

Dr. Francis: No, but I think it's really important to underline what he said at the end about diagnostic tests possibly being different from treatments. And with diagnostic tests, if it's worse to have a false positive and not be treated than it is to have a false negative - a false - no, I actually meant a false positive and not be treated if it's a staging issue, right?

Dr. Garber: We see your point. Go ahead.

Dr. Francis: Anyway, it depends on which is worse, right?

Dr. Garber: Yeah, right.

Dr. Francis: And that might be true for diagnostic tests too.

Dr. Sox: So you're basically underscoring the notion we can't just because it's hard to evaluate these tests doesn't mean we can okay to not do it, that we in fact should do it because so much is at stake even though it's in a small number of patients.

Dr. Francis: But with treatment, if there's less downside to treatment when it's not going to help anyway, we might go ahead and treat. And that same kind of issue about how you balance the risks of not treating versus the risks of treating when you shouldn't applies to diagnostic tests too.

Dr. Sox: But if you treat and the patient doesn't respond, then you stop the medication and you haven't lost anything.

Dr. Francis: Right.

Dr. Sox: If you do a test and at least with a false positive or a false negative that gets you off the right track in terms of making a diagnosis, you've really made a mess of things.

Dr. Garber: I think there's this very simple point to be made, that let's say that with the diagnostic test, and I think this is what Leslie is saying, if it's a positive you would treat and if it's a negative you wouldn't treat. But let's say that you're in a clinical situation where it's much worse to fail to treat than to treat. In that case if you don't know a lot about the test, the rational approach I believe is to just treat and not do the test. So no case can be made for the test. You could have a similar lack of information about the treatment but that's really different because then again, you would just treat if you're in a situation where the potential upside is reasonable and there's virtually no downside. So it's very hard to construct an argument to say that you would use the diagnostic test when you knew very little about its test performance, its performance characteristics. Dr. Sox: Now I want to point out that we need to allow some time for public comment and because a number of people are going to have to leave at 3:30, that I guarantee we will be done at 3:30, we just have to kind of move along. I think we've given CMS a lot to think about and a pretty strong opinion about the application of this framework for rare diseases, so let's go ahead and move into the public comment phase.

Ms. Anderson: We're going to do the scheduled public comments. We have one public commenter and it's Dr. Richard Wahl. Will you please come forward, state your name for the record and your affiliations, and any conflicts.

Dr. Sox: And if you could also tell us about how long you think you're going to take so we can set a schedule.

Dr. Wahl: Pretty briefly. Richard Wahl. I'm at Johns Hopkins University, director of nuclear medicine. I appreciate the opportunity to speak here today. I actually am representing the Academy of Molecular Imaging. I have several conflicts. I have received research support from GE and Siemens, as well as honorarium from both of those organizations who make high tech imaging devices. I've also consulted and received honorarium from Adak and CTI. I'm a shareholder in CTMI, who makes devices. I've been a scientific advisor to Diatide, Burlex Laboratories, and hold patents on diagnostic and therapeutic agents. So, I'm thoroughly conflicted but hopefully conflicted in a balanced way.

I enjoyed the discussion and found it fascinating. I'd like to comment about a couple of the points, particularly Dr. Garber's points about no downside to therapy in a minute. Could I have the first slide please?

One thing I did want to say just in discussion, it is fascinating and just to let you know, I'm professor of radiology and vice chair of radiology. I deal with imaging patients with cancer all the time and in a tertiary care hospital like Hopkins, we see a lot of common cancers, but you see a lot of uncommon cancers referred in who are reallly diagnostic challenges. Because the data on the therapeutic options and the date on the diagnostic tests are simply not of the same quality as the more common illnesses because you don't have the number of patients in every study. So your confidence intervals on sensitivity and specificity are less robust, and the power of therapeutic studies is less good.

One thing I did want to mention is that the concept of the reference test for diagnostic imaging is quite difficult because in fact the reference test for diagnostic imaging tests often would include a whole body cryosectioning with immunohisto chemical assessment, which is thoroughly invasive. It's actually very hard to have truth as to where cancer is located within the body and not be destructive in your testing. So having the reference test is highly desirable, but if we really had reference tests in all these conditions which are difficult to manage, and which were not invasive, we wouldn't need new tests. So the reference tests are a problem, and often our best reference test is an autopsy with additional sectioning. So that's a problem with the approach discussed. Could I have the next slide please?

So the diagnostic questions we're often faced with in cancer imaging is whether cancer is present or not, which is detection. Usually in patients in our institution they are highly selected so there's a high probability of cancer. They may have a mass, trying to determine if cancer is present or not. Whether cancer is localized or whether it has spread, that is whether it has disseminated of not, because this makes a big difference in terms of the decision whether to do surgery or not, and surgery is a treatment that carries with it substantial risks and costs, and whether or not cancer has spread or not makes a big difference. And there's certainly a downside to doing unnecessary surgery. Also, we like to know whether therapy is working; if it's not then we'd like to abandon it or change it and determine if more treatment is necessary. And diagnostic studies with new agents or new methodologies often look at these questions. Next please.

Unfortunately, the standard method for assessing cancer for the last hundred years has been anatomy, which though useful and widely applied without limitation to prevalence of disease, doesn't really adequately characterize most aspects of cancer. It doesn't tell whether masses are malignant or benign, because many have water density and overlap. It doesn't tell if lesions have spread to regional lymph nodes because small lymph nodes contain metastases often not detected. It has substantial difficulty with small lesions, small lesions near normal structures. And if a patient has had surgery or radiation, the performance is altered when anatomy is altered. Next please.

Further, in terms of planning therapy, our current imaging methods don't predict response to a given therapy, often show little information about tumor biology, are slow to change in response to therapy, certainly difficult in terms of assessing cytostatic therapy, you have to give it for months to see if the tumor doesn't shrink, and then new imaging techniques which are being proposed and looked at in other diseases have the opportunity to address the limitations of anatomic imaging. Next please.

Now the prevalence of cancer and the incidence of cancer are shown in this particular graph which is in your handout, and basically in my opinion, the quality of our diagnostic imaging data is related to the frequency of these diseases. And as you go down the rate of incidence, obviously the quality of the studies to some extent declines. Next please.

These are death rates, which are really quite similar, showing lung and breast to be predominating, and you notice that lung, breast and colorectal are diseases where we have fairly significant diagnostic imaging data. But as you get down to the less frequent ones, we run into more problems. Next.

So what is a rare cancer? The issue of an orphan disease, whether the prevalence is less than 200,000 a year or less than 100,000 a year. There are certainly some diseases that I've called very rare that may be less than 1,000 cases per year. In my opinion, these are sufficiently infrequent, particularly for the very rare, that doing prospective large blinded clinical trials, even in several institutions, is difficult, time consuming or impossible due to low rates of presentation. Thus, there is less evidence in a specific rare disease of the efficacy of diagnostic tests and it's very difficult if you have rare diseases to take consecutive patients who have the same presenting complaint. They may present with different complaints. I think the technology will evolve and I think Dr. Brook's points about the logistics of doing these studies are real, that there are significant economic costs to doing very large multicenter studies.

I would say they are insufficiently frequent to justify to the NCI in the current environment and other funding agencies the tests. More pressing and common diseases win out. For instance, the recent Akron decision to fund the 50,000 patient study for lung cancer screening was done over, because of a limited budget, over some smaller studies that would have been in less prevalent diseases. Next please.

So what are some examples? I think virtually all pediatric cancers, and I think the non top-10 adult cancers. Anatomic imaging is routinely used in managing these diseases but limited disease specific high quality evidence supports this. Next please.

I knew you'd like this list. I think it illustrates that there are a lot of different kinds of cancers, and I think that many healthcare economists, to do proper studies in all of these I think would exceed the capability of the healthcare enterprise in the U.S. and probably our budget to do proper trials. Next please. So how do we deliver advanced imaging methods to patients with rare diseases, specifically cancer? I mean, the evidence based medicine and disease specific approvals don't occur for some diseases under current CMS guidelines. Thus, if you are a patient with a common cancer and good evidence for a method, you're okay. But if you have rarer cancer and limited evidence, newer imaging methods are basically not available. So in my opinion, because current CMS guidelines don't allow some methods to be applied, this discriminates against children and patients with rarer diseases. It's clearly in my opinion not desirable nor appropriate that if you have a rare tumor, that you not have access to methodologies, especially when therapeutic methodologies are known to carry significant downside risks such as surgical procedures which are not indicated. Next please.

So how do we rationally deal with these? And I think an excellent discussion just preceded but I personally believe the mechanism of evaluation for these reasons must be different than for more common cancers and that we actually don't have to treat every new cancer, every rare cancer like we don't know about anything about imaging and that it's the first time we ever had an idea that imaging could be used. For instance, if I wanted to drive my - let's say I drove my automobile to Detroit and then I drove it to Boston. Is it likely I'll be able to drive it to Charlotte, Virginia? I don't know, I can't tell you for sure but you don't have to necessarily run the study. The question really is, do we know about the mechanism of uptake or of imaging, do we understand the biology well enough that we can perhaps use existing data to predict whether the method would be biologically reasonable and whether the technique itself is sufficiently robust physically and technically. Next please.

So perhaps a database from studies of more common cancers could show, or showed efficacy of the technique could be used to guide smaller studies on the cancer in question, showing the expected sensitivity. So if you showed reasonable sensitivity in smaller studies, albeit with less good confidence intervals, and the technology worked in a lot of other diseases, do you want to deny patients for 10 years the coverage and let them have inferior methods, or do you want to go with the clinical judgment and use a more rational method or potentially better method? For very rare cancers it may be very difficult to have supportive data, including histology, and in the extremely rare it seems to me the studies are going to be impossible to do. Certainly expert consensus agreement on general utility would make sense and use in compliance with FDA labeling would also be reasonable if CMS guidance is more restrictive. Next.

So there's a whole lot of biological alterations in cancer that are well recognized, and some of them are targets for imaging. I would suggest that if we know for instance that C-Kit is altered and you have an agent that images things with C-Kit mutations, that you may not need a thousand new patients with a new disease to show it if it worked in the first thousand of the related disease. Next please.

For instance, this is just - I hate to show a scan but here is an example of where we know the biology with glucose transport that on the lower right and upper right is a very cellular glucose transporter rich cancer. And we know in vitro and in vivo that high glucose transporter levels and high cellularity are

associated with high glucose uptake.

And on the left is a less avid, less visible tumor with less glucose transporter presentation with less cellularity. So we know something of the biology that links some of our imaging tests to the actual histology of the tumor. Can some of this data be used in rarer tumors? I think the answer is yes, and some of these data are easier to come by than multicenter prospective trials. Next please.

And again we know in a lot of tumors that uptake of certain tracers, in this plate FDG, and the number of living cancer cells are related. This is true across large numbers of cancers in vitro and maybe some of this data could be used in helping to guide a panel to a decision on a specific coverage. Next please.

So in more common cancers, I think that the approach works reasonably well. You can learn if the test is generally accurate. You can learn if there's FDA labeling, is the test generally safe and effective, what are the limitations. In more common cancers you can learn physical size limitations, what causes false positives, and perhaps do a meta analysis of larger studies. Next please.

In rare cancers smaller case studies might be considered sufficient, showing the expected accuracy of diagnostic methods versus the experience in more common cancers. I'm convinced that it will be difficult to meet the highest standards of evidence in these tumors and that perhaps estimates of sensitivity versus histology could be provided. And outcomes data obviously would be useful but not commonly available. Next please.

So for example, small cell lung cancer is not an approved indication right now for one imaging methodology but it's approved for non-small cell. Similarly, I saw a case just before I came over here of mesothelioma, which of course isn't an approved indication but non-small cell is. Other more common tumors, endometrial and cervical cancer. Next please.

And the very rare tumors, I mean also at a conference we were looking at a patient with a thymoma who was helped by an imaging test so that he had the right area of his thorax biopsied. Similar testicular cancer, I know we did a prospective study of an imaging methodology in testicular cancer and it took us six years to get 15 patients. It worked great but the numbers and confidence intervals are still marginal. Rare cutaneous tumors that are hard to find and do large studies. GIST tumors, it's know that gleevec turns off GIST tumors and stops the growth of these tumors. But to do a large prospective trial of gleevec and these tumors is going to take quite a while. We also know that imaging works well in these tumors and gleevec turns off molecular imaging signals. Are case reports, or a few case reports sufficient? Next please.

So I think that patients, adult and pediatric with rare tumors do not currently benefit from some advanced imaging methods under current approval methodologies for Medicare. They are able to get anatomic methods but not all molecular imaging methods. I think rare diseases and rare cancers could benefit from these methods. And I personally believe from what I see on a daily practice basis and at fairly, I think advanced academic medical centers, that the approval process maybe should be different

and less stringent to make the methods available but still evidence based, but perhaps using more source of evidence. Next please.

And that the biological basis of the imaging procedure, the body of evidence supporting the methodological validity, solid data for more common cancers, and small studies showing efficacy, perhaps limited sensitivity or feasibility studies should be useful in making a decision. Certainly studies should not be showing that a lack of efficacy occurs; that would certainly be a red flag to an approval. Use consistent with FDA labeling and expert consensus may be helpful. And a broad approval of a method, with exceptions if it's known not to work well might be a more rational approach.

I do thank the committee for their deliberations on this and for allowing me to speak. Thank you.

Dr. Sox: Thank you, Dr. Wahl. We have a few minutes for questions for Dr. Wahl. Why don't you stay up there. Bob?

Dr. Brook: I was trying to understand in what clinical areas in these rare cancers, because we've been having this intellectual discussion. The one specific example you used was where to biopsy for somebody that you knew had the disease, and that would go towards the discussion that we had about looking for metastatic disease. Is there limitations on using this in the rare cancers for identifying the disease that you can come up with or think about?

Dr. Wahl: Well, I think one thing I wanted to make clear to the committee is that usually an imaging test is used in concert with a bunch of other imaging tests, biopsy and all clinical information to make a decision about how to manage a patient. Imaging tests will have false positives and false negatives, and usually they have to, things have to sort of line up to make sense in terms of a decision made on a patient. A simple example of where a more sensitive imaging test would be useful is in exactly the situation you say. A patient may present with a mass that on biopsy is confirmed to be a relatively infrequent tumor, let's say it's a thymoma, maybe 2,000 cases a year in the U.S. And the question is, is it localized? Especially in the situation of recurrence.

Let me just tell you about a case I saw today. The question is recurrence. There's all kinds of radiographic changes with pleural thickening and masses and so on, and you wouldn't know where to biopsy. You would actually have to take out an entire lung or potentially both lungs to figure out where anything is. If you do a scan, and I think I can say in this particular case the patient wasn't a Medicare patient and was allowed to have a PET scan and their private insurer paid for it, and it showed two specific foci that had increased glucose metabolism. The patient went to thoracoscopy, you know, minimally invasive surgery, was evaluated and a biopsy was taken of one of the spots that had high glucose uptake. There was also an area of pleural thickening on CT but there was lots of pleural thickening. So in this particular instance this patient clearly had recurrence, it was directed to the recurrence by having the complementary information. So the situation we usually face is that it's -

Dr. Brook: Can I ask you a question? If you biopsy that one lesion and it was negative, would you still

go to the full procedure?

Dr. Wahl: Perhaps I don't understand you. If you're asking me would I -

Dr. Brook: You were just saying that one of the problems is that you have pleural thickening all over the lung.

Dr. Wahl: Yes.

Dr. Brook: And you don't know where to biopsy, and you did a test based on some biology that showed a hot spot. And if you biopsy that hot spot and didn't find the tumor there, would you then do your original course of action anyway or would that obviate the need - are you certain enough clinically with the old way of using these tests that you would then have concluded that none of this was, that this was not a malignant pleural process but it was just scarring?

Dr. Wahl: In this particular instance the concern was that it would have all been scarring, yeah. I mean, they may not have even been biopsied. So that the person might have had a delay in therapy because they didn't know it was metastatic. I mean, a more common situation I think would be patients, let's say a kid comes in with a mass and you get the biopsy of the mass and it's a relatively infrequent tumor like a Wilms tumor. If they have metastatic disease the treatment I think could be substantially different than localized disease. And you know, the decision - you know, not all therapies, most cancer therapies in phase one and two don't work and have toxicity, and surgery works very well but it carries with it the risk of death, morbidity and substantial cost. So cancer therapies are not in my opinion to be dispensed with lightly. In other words, you would like to use the right therapy and using the wrong therapy carries with it costs and consequences to society. So if knowing if there's more disease is going to make a big difference, then a test that shows you that there's something additional that you should biopsy and then you biopsy it and confirm it, that can only help the patient. So I'm concerned as a physician right now in my practice that I offer two standards of care, one for patients who have highly frequent diseases where evidence based medicine has been able to provide enough confidence that we can use their diagnostic test with a good sensitivity and specificity, and another group of patients where a lot of data suggests it would work but there's just not as much evidence. And then the question is are they better off having a diagnostic test that's not as well known or are they better off having a surgical procedure that may not be indicated. And I would argue that in some of those patients they are better off having a diagnostic test followed by a biopsy to confirm it than an inappropriate surgical intervention.

Dr. Sox: I would be interested in - I think you're suggesting that in some cases where diseases are rare, that we take test performance where we've got the numbers, and we apply it and we assume that it applies to the rare disease. I think that's what I was - maybe you should repeat it. You know the performance of PET scanning in a common disease, you have no idea what its performance is in a rare disease. So assume that the performance in the common disease applies to the rare disease as well, and if it's good performance then you approve it, you would urge that it be covered.

And I guess my question relates to what kind of principles is reasonable to apply if you make that, if you make that decision to take that approach rather than the more challenging and restrictive approach that we discussed? What kind of principles apply? What kind of confidence interval should you place on measures of test performance when you're making the leap that performance in one disease is the same as it is in another? What kinds of diseases are, is it more reasonable to make this leap than others?

Dr. Wahl: Well, I think that linking it - first of all, this stupid thing about driving the car, I mean, if you know your car goes one place then it will probably go another place, even though you haven't scientifically proven it, the confidence, you know you - as a physician you do look at stuff, and I mean you guys obviously look at methodologies, and I think at a minimum your technique has to be fairly reproducible from center to center in some of the common illnesses. So our imaging tests should work and they shouldn't be so complicated that nobody knows how to use them and you can't - one guy in Institution A gets a 30 percent sensitivity and somebody in Institution B claims 90 percent. I mean, in the multicenter studies you should have some level of a fairly good kappa among readers, that it's a test that's mature enough, and there aren't enough technical issues that it can't be disseminated. I mean, I think that's one fundamental thing

.

Then secondly, I mean, the issue is if you understand the biology of the test, is there evidence short of large clinical trials that the biology of this particular illness is similar to the other illnesses in which the tests work? Now there is an increasing tendency in diagnostic radiology to try to link and understand why our tests work based on a biological basis instead of just whether there's a lump or not. The lump or not issue is also important, though, because if you know your test is no good for under 5 millimeters it would be silly to approve it for micrometastatic disease. You know, as I think the approval in breast cancer didn't cover staging of the axilla because of known sensitivity issues in that setting. But I do think the tests working in other diseases is critical, having some linkage in biology, and ideally I would feel better if there were a case report or some cases, a small paper showing that at least a performance was as expected versus other imaging methodologies or pathology that would be better. But in some very rare tumors, you might not have that. And then the question to the committee becomes do you want to deny people access to this or is it - you know, which is more important to society? Do you want to deny people the opportunity to have the test because you don't have enough data, or is the data in everything else so good that if you realize that you don't make absolute decisions based on imaging tests, you use it as an adjunct to other methodologies and other procedures in your total clinical judgment, is it a huge risk to approve it?

I mean, I think those are the issues.

Dr. Sox: Good. Alan?

Dr. Garber: I'll actually pass and wait for general discussion.

Dr. Sox: Leslie?

Dr. Francis: I'd just ask, if the worst case scenario is that somebody doesn't get an effective treatment, or a treatment that might have significantly helped them, and the next worst case scenario is that somebody gets an intervention that could hurt them, as a result of the diagnostic test, aren't you going to want to avoid the worst case scenario?

Dr. Wahl: Well, I don't necessarily agree with you that in our current state of management of cancer care that - I mean basically, the worst thing you can do to a patient for a disease that isn't so well treated is to give them aggressive therapy that kills them when it's not indicated, I think. So I think that therapies do carry with them risks, but you don't want to deny a patient the chance for cure under any circumstance. So most institutions, especially on the issue of detecting metastatic disease, would want to prove the first site of metastatic disease by a method other than imaging. They would want to prove it by biopsy. And generally they need to be guided to that spot to know where to biopsy.

I mean I hate to give anecdotes, but I saw a case yesterday of a patient who was supposed to have a radical cystectomy. However, we did a very sensitive imaging test, we actually did a PET CT scan and this patient had multiple lymph nodes up and down the iliac lymph node chain and I - it was certainly a recent case anyway, and the decision was made that this would be inappropriate to do a radical surgery procedure. And so doing the thing that theoretically would cure the patient in the situation where there's no chance of cure is not a good deal. So both carry risks, but I mean any thought, in my opinion, that phase one diagnostics being offered to patients with very poorly treated diseases are a great alternative to - they're not great. I mean, you don't want to -

Dr. Francis: But as your confidence intervals get bigger, it matters more whether a failure to treat is worse than a risky intervention.

Dr. Wahl: Yeah, but I guess what I'm saying is that I don't view diagnostic test outcomes as completely dichotomous. In other words, I don't view it as a go/no go just based on a single test. And generally your single test is looked at with other tests, and if it's a huge decision, generally tissue is obtained, pathological samples, and one of the key issues is where do you obtain those. Biopsy is not a perfect procedure either.

As an example, in solitary pulmonary nodule, which is something you've approved, the diagnostic performance of biopsy is less good than of a PET scan in terms of sensitivity because the sampling error associated with biopsy. So we have problems with our "gold standards" too.

Dr. Sox: Barbara?

Dr. McNeil: I just have one comment because I'm going to have to leave in a minute. It seems to me we're talking about two different kinds of issues here. See if you agree with me, Rich, on these. One is that the general construct has been largely on FDG, just for the sake of argument let's talk about that as

one class, and another class would be other molecular imaging markers. So the reason we're talking about FDG right now is A, it's available, and B, the mechanism seems to involve glucose metabolism, it is glucose metabolism. So the question is, do we believe that tumors regardless of their prevalence or incidence, have the same propensity to show increased glucose metabolism? And if so, then the question that Steve raised would be a reasonable one, that if all tumors show increased glucose metabolism then we should have a sensitivity of 85 percent for everything, regardless of the sensitivity. I mean, that would be how that logic proceeded.

So the question for FDG, is that a reasonable assumption? And then if it is, fine, then we need to talk about what the steps are. But we would have to be pretty darned sure that that was a reasonable assumption to follow Steve's line of argument. Otherwise, then we really have to think about the implications of false positive and false negative decisions on treatment options on the one hand, and the feasibility of getting data on the other hand. So, I don't know the answer to that constellation of comments, but that really is one, and it strikes me as a little hazardous to say that glucose metabolism is glucose metabolism, because if you're talking about the lung, bronchoalveolar tumors in the lung aren't nearly as - PET is not, what is it, 60 percent, 40 percent sensitive, compared to -

Dr. Wahl: You have to basically read them with a different cutoff, but you can still have increased glucose metabolism but there is variability. Less cellular tumors, as I showed you, could have less uptake.

Dr. McNeil: I don't want to get into details here, but that is an example where you can't make the extrapolation on the basis of known data. So I think we need to talk about that particular marker. And then if we talk about other metabolic markers, which I certainly think is where the field is going, it seems to me we are talking about a whole different set of issues. We're no longer talking about the extent to which these different metabolic markers have the same affinity to different tumors because presumably they're going to be developed to be more tumor specific, to have a mechanism of action that's more appropriate for one tumor versus another tumor. So I think we're talking about not only the difference between rare and non-rare, but between fundamental mechanism of action, or between the mechanisms of action of the two. So there are a whole bunch of things that are going on at the same time, consequences of false positives and false negatives, the ability to extrapolate on the basis of the pathophysiology of the tumor and the mechanism of the imaging agent, and rarity and non-rarity, and it's just hard for me to get my hands around it.

I think it was a very complicated and interesting discussion, but I have no answers. I was just trying to put out a topology of some of the issues that we need to cover.

Dr. Sox: Plus the clinical situation.

Dr. McNeil: Oh sure.

Dr. Sox: Where your management decision, special probability may differ from that situation to

situation.

Dr. McNeil: Yeah, I think that's what Rich was saying.

Dr. Garber: Can I add one brief footnote to what Barbara said? She talked about varying sensitivity and whether there's a uniform sensitivity for the tumors, but even if there were uniform sensitivity it would be extremely unlikely that specificity could be uniform because that's dependent on the other conditions that you need to distinguish from the tumor that's under consideration, and there's all kinds of reasons why you would expect specificity to vary all over the place.

Dr. McNeil: You're right. I didn't think of that.

Dr. Sox: Bob?

Dr. Brook: There's a way of leveling the playing field here that would combine some of the thoughts of this afternoon. It might very well be possible for CMS to say we level the playing field and we will require a set of information of data that would allow you to have this covered for a rare disease so that we simultaneously try to meet our objectives, which may take some time, and at the same time we approve coverage. And since these are rare diseases, if we really keep it rare, we won't have a big impact on cost or health of the Medicare population. Certainly it's not going to affect kid's tumors.

So the bottom line here is that maybe there is an opportunity here to leave two charges at once, to solve some of the equity issues about rare disease but at the same time increase the onus on the fact that doctors and patients who really want this have got to be part of a process by which we collect the data to test these hypotheses that are only hypotheses at this moment. That may be a way of doing it together and may be attractive in a policy process.

Dr. Sox: Well at that point I think we'd better call the discussion to a close so I can come through on my promise to get us done by 3:30. So at this point I'll ask Steve to say whatever parting remarks he wishes to make, and then Janet has some things to say as well.

Dr. Sheingold: Thank you, Hal. This being my first full EC, I'm not sure how Sean sends you off but I just want to thank our public speakers, thank Deborah Zarin for being here to answer some of the questions about technology assessment, and most of all thank the committee not only for being here but for just such a high level discussion today on all the topics on the agenda, and particularly the last two, which weren't on a specific technology. You have given us a lot to go back with and we will come back with another draft of those referral criteria. So I wish you all safe journeys home and look forward to working with you in the future. Janet.

Ms. Anderson: Very briefly. I do have to mention to the audience, for continuing information you can visit our web sites, either at www.cms.gov/coverage or there is a click that says coverage process on the left-hand side on the same web site, www.cms.gov.

To conclude today's session, would someone move that this meeting be adjourned?

Dr. Garber: So move.

Dr. Bergthold: Second.

Ms. Anderson: And it's been seconded. Thanks to all.

(Whereupon, the meeting adjourned at 3:21 p.m.)