CENTERS FOR MEDICARE AND MEDICAID SERVICES

MEDICARE EVIDENCE DEVELOPMENT AND COVERAGE
ADVISORY COMMITTEE

- - - - - - - - - - - - - -x
:
IN RE: FDG-PET AND :
:
FDG-PET/CT :
:
- - - - - - - - - - - - - -x

Woodlawn, Maryland

Wednesday, August 20, 2008

The above entitled matter came to be heard before the Medicare Evidence Development and Coverage Committee in the CMS Auditorium, 7500 Security Boulevard, Woodlawn, Maryland.

COMPOFELICE REPORTING SERVICES, INC.

6671 Farbell Row

Columbia, Maryland 21045

(301) 596-2019 (410) 381-2755

(800) 464-3019

FAX (410) 290-7249

PANELISTS:

ACTING CHAIR:
SATY SATYA-MURTI, MD, FAAN

PANEL MEMBERS:

I. CRAIG HENDERSON, MD
NORA A. JANJAN, MD
J. LEONARD LICHTENFELD, MD
ANDREW SLOAN, MD

CONSUMER REPRESENTATIVE:
LINDA A. BERGTHOLD, PhD

INDUSTRY REPRESENTATIVE:
PETER JUHN, MD, MPH

GUEST PANEL MEMBER:
RICHARD I. WAHL, MD

CMW LIAISON:
STEVE E. PHURROUGH, MD, MPA

EXECUTIVE SECRETARY:
MARIA A. ELLIS

C O N T E N T S

P R O C E E D I N G S

MS. ELLIS: Good morning and welcome, chairpersons, members, and guests. I am Maria Ellis, an executive secretary for the Medicare Evidence Development and Coverage Advisory Committee, MedCAC.

The Committee is here today to discuss the evidence, hear presentations and public comment, and make recommendations concerning the oncologic indications of FDG Positron Emission Tomography, PET, for nine cancers: brain, cervical, small cell lung, ovarian, pancreatic, testicular, prostate, bladder, and kidney. The meeting will discuss the various kinds of evidence that are useful to support requests for Medicare coverage in this field.

The following announcement address conflict of interest issues associated with this meeting and is made part of the record. The conflict of interest statutes prohibit special government employees from participating in matters that could affect their or their employer's financial interest.

Each member will be asked to disclose any

financial conflicts of interest during their introduction. We ask in the interests of fairness that all persons making statements or presentations also disclose any current or previous financial involvement in a company that manufactures or provides devices or other tools for the research of PET. This includes direct financial investments, consulting fees, and significant institutional support. If you haven't already received a disclosure statement, they are available on the table outside of this room.

We ask that all presenters please adhere to their time limits. We have numerous presenters to hear from today and a very tight agenda and, therefore, cannot allow extra time. There is a timer at the podium that you should follow. The light will begin flashing when there are two minutes remaining, and then turn red when your time is up.

Please note that there is a chair for the next speaker and please proceed to that chair when it's your turn.

For the record, voting members present for

today's meetings are: Dr. I. Craig Henderson, Dr. Nora Janjan, Dr. J. Leonard Lichtenfeld, and Dr. Andrew Sloan. A quorum is present, and no one has been recused because of conflicts of interest.

The entire panel, including non-voting members, will participate in the voting. The voting scores will be available on our website following the meeting. Two averages will be calculated, one for voting members, and one for the entire panel. I ask that all panel members please speak directly into the mics. And you may move them -- you may move to the mic since we have to share.

If you require a taxicab, there's a sign-up sheet at the desk outside of the auditorium. Please submit your request during the lunch breaks. And lastly, please remember to discard your trash in the trash cans located outside of this room.

And now, I would like to turn the meeting over to Dr. Steve Phurrough.

DR. PHURROUGH: Good morning and welcome. A particular thanks to the panel members who have agree

to be part of this event today. The panel is a bit smaller. We had a few more who, over the last couple of days, had some issues that prohibited them from being here.

We have a topic today that we've engaged in over the last several years, looking at appropriate indications for the use of PET scanning in oncology conditions. And we look forward to a vigorous discussion around those particular indications.

In addition, the panel is going to have some broader discussions around the concept that we used in our last decision to collect some information, this coverage with evidence development, and have some discussions around the validity of this kind of data collection and its -- in making decisions, not only around coverage, but around use of technologies at the physician/patient interface. So we look forward to that discussion today.

Just a note as to the -- for those of you who are veteran MedCAC attendees, we have a fairly open voting process. Questions are asked, and the

panel gives their vote. In this case, typically, we hold up the score. Most voting questions are Lichert scales one to five.

However, two of the -- the first two voting questions -- the first two voting questions are charts where a lot of information is requested. The panel has worked on those charts prior to this meeting. They will finalize these charts as the day goes on. And at the voting time for those two questions, they'll be given a few minute to finalize that, and those charts will be turned in.

There will not be a specific question where they raise their hand or vote on each specific block within that table. They will be asked to comment broadly around what they thought about the evidence as they filled in that table.

Our staff will take those individual charts and hopefully in a short period of time, assuming that we can make technology work, we'll have all of those averages available on the screen for you to look at as we proceed with our other questions. Trying to ask

for a vote on each block in each chart would have kept us all here long past the time that we would have wanted to be here.

A couple of logistical kinds of questions. There is only the lunch break for both panel and guests. If you need other breaks, take them as you need them.

We have today as our recorder a new individual doing that. She has permission to tell you if you aren't being heard. So for both panels and speakers, if you're not speaking into the microphone, you'll will be chastised so that we can get a good accurate record of what's going on today.

So I've given her permission to be very straightforward. And she tells me she has no trouble with that. So please use the microphones, and I'll try to remember to do that also.

So again, thank you for all your participation today. I unfortunately will have to step out for about an hour shortly after we begin. My boss thinks I need to be somewhere else for a short

period of time, and so I'll do that. And Dr. Jacques will be the designated federal official during that period of time.

So with that, I'll turn the meeting over to Dr. Satya-Murti who is our chairman today.

DR. SATYA-MURTI: I was able to convince Dr. Phurrough that in view of the caffeine withdrawal we are undergoing this morning, we will be having a short break between 10:30 and 10:45.

I am Saty Satya-Murti. I am a neurologist. I do part-time voluntary teaching and clinical work at an inner city non-profit, and I consult for industry. I have no conflicts of interest.

DR. HENDERSON: I'm Craig Henderson. I'm a medical oncologist from the University of California, San Francisco. I have no conflicts of interest.

DR. JANJAN: Nora Janjan, M.D. Anderson Cancer Center. I have no conflicts of interest. I'm a radiation oncologist.

DR. LICHTENFELD: Len Lichtenfeld, American Cancer Society. No conflicts of interest.

DR. SLOAN: Andrew Sloan, University Hospital Case Medical Center. No conflicts of interest.

DR. BERGTHOLD: I'm Linda Bergthold, and I'm the consumer member of the panel. And I have no financial conflicts of interest.

However, I am doing some advising to the Center for Medical Technology Policy on private sector coverage with evidence development. So I'm going to lay my biases out on the table and say I think it's a really good idea, and I'd like too see it work.

DR. JUHN: Hi. Peter Juhn, Medco Health Solutions. I'm the industry representative. No conflicts of interest.

DR. WAHL: I'm Richard Wahl from Johns Hopkins. I'm Director of the PET Center at Johns Hopkins, so in that sense, I have work that deals with PET and part of my clinical income is derived from PET interpretation.

I also have a patent that was developed at the University of Michigan some years ago on

radionuclide guided biopsy, which has been licensed by a dedicated breast PET company which, so far, hasn't resulted in any sales. But there's a potential conflict of interest.

And in the past, I have received -- or Johns Hopkins has received research grant support from General Electric, who makes PET scanning devices.

DR. SATYA-MURTI: All right. We'll start the meeting. It's a weighty topic. The first presenter is from CMS, Stuart Caplan.

Stuart, are you ready?

MR. CAPLAN: Yes, sir.

Good morning, everyone. Thank you. I hope everybody is comfortable, and that your travel here was easy. Chairman Murti, panelists, members of the public, and invited speakers, once again, let me welcome you to today's Medicare Evidence Development and Coverage Advisory Committee meeting.

As Maria mentioned a few moments ago, today's topic is FDG-PET or positron emission tomography for use in nine cancers. Now, they are

brain, cervical, ovarian, pancreatic, small cell lung, testicular, prostate, bladder, and kidney.

When we refer to PET today, we're also to referring to PET/CT. The terms will be used interchangeably.

The CMS analytic team for today's presentation includes myself, Stuart Caplan, as the lead analyst; Katherine Tillman as the co-analyst; Dr. Shamiram Feinglass who is our lead medical officer on the project; Dr. Jeffrey Roche, who is the co-medical officer. We've heard from Dr. Steven Phurrough, the Director of the Coverage and Analysis Group. Also part of the team is Dr. Louis Jacques who's the Director of the Division of Items and Devices and our executive secretary, Maria Ellis, whom you all know well. And I'd like to also thank the other people in our staff who helped with today's presentation.

The panel has received the following materials today: a draft technology assessment provided by the Agency for Healthcare Research and Quality, presentations of scheduled presenters. Note

that that written testimony bullet is incorrect, although transcripts of today's meeting will be available on our coverage website.

The panel members have a summary of evidence, definitions of PET indications from the National Coverage Determinations manual, and questions for the panel.

PET is nationally covered for a variety of oncologic, neurologic, and cardiac indications. But it's important to note that the national coverage determination for PET is an only policy. Unless a specific indication for PET is nationally covered, then it is nationally non-covered. F18, Fluorodeoxyglucose -- Fluorodeoxyglucose or FDG is the only nationally covered radionuclide for PET's use in oncology.

Beginning in 1998 and through January of 2005, coverage of PET has been based on cancer site and indication for the study. FDG-PET indications in oncology are defined in Section 220.6 of our National Coverage Determination Manual. And these indications

are diagnosis, staging, restaging, and monitoring.

I'd like to read an excerpt from Section 220.6 that defines each of these indications. And we've provided the panel with a copy of these definitions. This will take a moment.

Diagnosis is defined, "PET is covered only in clinical situations in which: (1) the PET results may assist in avoiding an invasive diagnostic procedure, or in (2) the PET results may assist in determining the optimal anatomical location to perform an invasive diagnostic procedure. In general, for most solid tumors, a tissue diagnosis is made prior to the performance of PET scanning. PET scans following a tissue diagnosis are generally performed for staging rather than diagnosis. PET is not covered as a screening test (testing patients without specific signs and symptoms of disease). So that's the definition for diagnosis.

For staging out of Section 220.6, "PET is covered for staging in clinical situations in which: (1)(a) the stage of the cancer remains in doubt after

completion of a standard diagnostic workup, including conventional imaging, such as CT, MRI, or ultrasound, or (1)(b) it could potentially replace one or more conventional imaging studies when it is expected that conventional study information is insufficient for the clinical management of the patient, and (2) clinical management of the patient would differ depending on the stage of the cancer identified. That's the staging definition.

For restaging, PET is covered, "(1) after completion of treatment for the purpose of detecting residual disease, (2) for detecting suspected recurrence or metastasis, (3) to determine the extent of a known recurrence, or (4) if it could potentially replace one or more conventional imaging studies when it is expected that conventional study information is insufficient for the clinical management of the patient. Restaging applies to testing after a course of treatment is completed and is covered only subject to the conditions above.

Finally, the definition for monitoring is a

short one. "This refers to PET to monitor tumor response to treatment during the planned course of therapy (for example, when a change in therapy is anticipated).

As part of the CMS's in-depth evaluation of PET over the years, CMS has commissioned technology assessment through the Agency for Healthcare Research and Quality. In September of 2000, there was a technology assessment that addressed 22 indications for PET. In May of 2001, there was a MedCAC -- I'm sorry -- a technology assessment for breast cancer. In April 2002, we commissioned one for soft tissue sarcoma, and in April 2004, there was one commissioned for six cancers; brain, cervical, pancreatic, small cell lung, and testicular cancers.

In its evaluation of FDG-PET over the years, CMS has also convened MedCAC meetings in October of 2000 and June 2001. Also, a MedCAC Executive Committee for Diagnostic Imaging reviewed evidence on PET for breast cancer in 2000.

In May of 2006, CMS implemented a national

coverage determination that expanded coverage for FDG-PET for all cancer site and indications with limited exception, as long as data was collected at the time of service to assist in patient management.

CMS identified the National Oncologic PET Registry or NOPR to prospectively collect these data on the expanded cancer indications. We'll hear more about NOPR in a later presentation.

Let's go on to the panel questions. The scale for each response to the questions, one, no confidence, two, little confidence, three, equivocal, four, moderate confidence, and five, high confidence. The panels have numbers to hold up so their votes can be recorded.

Please note, panel members, that if you have no comfort level responding to a question, you have the option to not vote on that question.

Question one, how confident are you that the evidence is adequate to conclude that FDG-PET imaging improves physician decision making when used for the following indications of each of the nine cancers

mentioned earlier, diagnosis, staging, restaging, and monitoring? And again, the same five point scale of confidence.

Question two, how confident are you that the evidence is adequate to conclude that FDG-PET imaging improves proprietary oriented clinical outcomes when used for the following indications in each of the nine cancers, diagnosis, staging, restaging, and monitoring response to treatment? Again, no confidence, little confidence, equivocal, moderate confidence, or high confidence.

Question number three, how confident are you that these conclusions are generalizable to other cancers? The same voting scale.

Question four, how confident are you that these conclusions are generalizable to non-research PET facilities in the general community? No confidence, little confidence, equivocal, moderate confidence, high confidence.

The last question, number five, how confident are you that these conclusions are

generalizable to the Medicare beneficiary population?

Here are contact information. My email address, if anyone wants to find me, is stuart.caplan@cms.hhs.gov. Here's the web address for the Medicare coverage database. Thank you everyone, and be well.

DR. SATYA-MURTI: We'll start with the first TA presentation. I don't know who is presenting.

DR. GULENCHYN: Good morning. My name is Karen Gulenchyn. I am the chief of nuclear medicine in Hamilton, Ontario. But for the purposes of this presentation, I have been working with the Evidence-based Practice Group from the University of Alberta.

And I may be the most relieved person in the audience at the moment because Maria is my technical expert, and I had this horrible feeling that she wasn't going to be here on time. I've been here for a while and nervously watching for her. I am certainly not the technical expert on technology assessments. So I was very, very grateful to see her in a audience as I stood up.

I'd like to acknowledge that this work was done by the University of Alberta Evidence-based Practice Center and supported by the Agency for Healthcare Research and Quality and Centers for Medicare and Medicaid Services.

Today we are going to present very briefly some background information as background information has already been presented, and I do not want to be redundant; review with you the key questions that we were asked to address as part of this technology assessment.

We are going to outline the methods that were used in the technology assessment such that you are able to determine whether or not those methods were, in fact, robust and valid; present the results or a brief summary of the results of the technology assessment and summarize the data and provide you with the opportunity to ask questions.

As you know, positron emission tomography has been under study for a long period of time. It is a technology that became widely used in cancer over

the last ten years with the development of whole body imaging devices.

And now with the introduction of PET/CT, we will be using terminology quite interchangeably here in terms of PET and PET/CT. But we have identified interested technology assessment which technologies have been applied to specific papers. Concerns have been raised about the cost of this technology and also about the cost of imaging in general with the rapid growth of expenditures on imaging services, which has outstripped other areas of medical practice.

And this has resulted in really a flurry of technology assessments that have been performed around this technology, beginning in 1997 and extending now to the present one. Certainly this has occurred in multiple jurisdictions and not just in the United States. There has been an international focus on this particular technology.

And I'm going to turn this over to Maria now to continue. We're going to flip back and forth as we go through the clinical and technical portions of this

presentation.

MS. OSPINA: Thank you, Karen. Good morning, everyone.

So this report is kind of a continuation of previous report that was made by the Duke Evidence-based Practice Center in 2004. That report evaluated pretty much the same questions that we did, but it was related with six types of cancer; brain, cervical, small cell lung, ovarian, pancreatic, and testicular.

The conclusions that this report arrived was that PET was beneficial in helping physicians with clinical questions that were related with staging and detecting metastasis disease and recurrences.

They also identified some limitations in their literature, and they were based on the fact that most of the PET technologies were older generations, the group of patients that were included in the studies were very heterogeneous, and the results were not presented by significant clinical subgroups.

So part of this work has continued by this technology assessment that was commissioned by the

Centers for Medicare and Medicaid Services through the Agency for Healthcare Research and Quality. The objective is to try to summarize or synthesize the evidence available for the use of FDG-PET for nine different types of cancer, which are bladder, brain, cervical, kidney, ovarian, pancreatic, prostate, small cell lung, and testicular cancer.

This report is focused on four main questions. The first question is related with the diagnostic performance of FDG-PET or FDG-PET/CT compared to other conventional modalities for diagnosis or other procedures such as biopsy or other tumor markers. The idea was to evaluate the abilities of and the accuracy of PET for different clinical situations which were related with diagnosis, staging, restaging, and monitoring response to treatment.

The second question has to do with the diagnostic thinking impact of FDG-PET and FDG-PET/CT for the same clinical situations.

The third question has to do with the impact of FDG-PET and FDG-PET/CT as part of a management

strategy in terms of the impact that this technology and including this technology in the decision-making process might have on patients and their outcomes.

And the fourth question is about an economic evaluation more precisely, a cost effectiveness analysis of FDG-PET and FDG-PET/CT for the same clinical situations that we've been considering throughout the report.

Basically, we've used recognized and standard methods for synthesizing the evidence. And I'm just going to make a brief presentation here of the process. So first of all, we refined the key questions along with HRQ and CMS. After that, we reviewed some guidelines on diagnostic reviews and prepared a protocol for the technology assessment.

Then we completed a comprehensive literature search and retrieved the articles. Then there was a process of selection of the studies according to a set of eligibility criteria. Then we made a quality assessment -- methodological quality assessment of the studies. We made that extraction, and then we

summarized and synthesized the evidence.

So I'm going to go a little bit more -- in a more detailed way in each of these steps of the process. We reviewed some guidelines and basically, the framework of this report is based on the six tiered efficacy model of the technology assessment that was proposed by Fryback and Thornberry some years ago.

Basically, we focus our report in level two to level six that are described by these authors. That means that we are not including any information about the quality of the imaging and the techniques themselves into this report.

Our searches covered a period from 2003 to 2008, and were conducted in main electronic databases such as Medline, Embase Central, and Scopus. We used both mesh terms and free word terms. The terms are described in the appendices of the full report.

For the selection of studies, this was a two-step process. The first one was screening titles and abstracts. And this work was done independently

by four reviewers. And all the decisions -- final decisions were made by a consensus.

Then we applied a set of selection criteria, and I'm just going to go briefly over them. So we were including studies that were published in English that were not duplicate, meaning that multiple publications will not be considered, but they will be considered part of a main study. In terms of the population, we were going to include studies that had more than 12 human participants, adults. The age limit that we established was older than 16 years with any of the 9 primary cancers that we were considering.

The test in terms of the type of studies that we will consider were those including either FDG-PET or FDG-PET/CT. In terms of the comparator, we did not make restrictions, and we were open to include any type of comparator such as biopsy or the use of other imaging technologies.

We did not make restrictions in terms of the study design, meaning that both perspective and retrospective studies will be considered. But we

tried to keep all of them separate in our analysis. And, of course, in terms of the outcome, those studies that provide numeric data that might allow us to provide simple estimates will be included.

Because we have four different questions here, we could not use a single quality assessment instrument for this technology assessment. And we tried to adapt some available quality assessment instruments based on the type of questions that we were addressed.

So for questions one and two, we used the Scottish Intercollegiate Guidelines Network tool which is based on the QUADAS instrument. This is widely used in systematic reviews of diagnostic tests for the types of questions related with diagnostic accuracy. So for questions one and two, that's the instrument we used.

But for question three, because question three about the impact of FDG-PET as part of a management strategy resembles more like an efficacy or effectiveness study, we tried to use an individual

component approach that will incorporate the fact that the best scenario that we can find to address this question will be like in the realm of a clinical trial or analytical cohort studies. So we developed a checklist based on these characteristics in trying to assess how the analysis and how the design of this study was planned.

For the question about the economic evaluation, we used the Consensus on Health Economic Criteria. This is the checklist. And it's, I guess, with the drama (phonetic) criteria are the most widely used instruments for economic evaluations nowadays.

But also, we tried to do a grading of the evidence. And we used the Veterans Affairs Technology Assessment Program grading scale to assess the evidence. This is a system that goes from A to D where A are the studies with the highest quality of evidence and D are those studies with multiple flaws in the design and analysis. So each of the studies was rated also independently by two reviewers, and the final decisions were made by consensus.

We extracted different types of data. But I'm just going to present pretty much the type of information that we're going to use in this presentation. So study design and methods, we tried to establish whether the study was retrospective or prospective in the data collection; what was the duration; what was the type of primary cancer; the inclusion and exclusion criteria for enrollment of participants; and any demographic characteristics.

We were also interested in describing the characteristics of PET in terms of the technical details, all the procedures about administration, what reference standards were used for comparison, and what were the criteria for interpretation, whether SUV was used or it was more like a visual interpretation of the data without consideration of any quantitative information.

Again, the data was extracted by one reviewer and was double checked and verified for accuracy by a second reviewer.

For the summary and synthesis of the

evidence, we made a narrative analysis and constructed evidence tables. For the questions about the diagnostic accuracy of PET, we provided two by two tables and individual values of sensitivity and specificity. But also, and based on different common characteristics on the studies, we tried to provide pooled estimates for the positive and the negative likelihood ratio. And we also produced a summary receiver operating characteristics curve.

For questions two to four, because it was hard to find a common denominator to try to combine the evidence, this was presented in a narrative way in the evidence tables.

The decisions, of course, of making or not making a qualitative or quantitative analysis is pretty much guided on what type of information we find. We wish we could have provided a more quantitative approach for questions two to four. But unfortunately, due to not that many number of studies and the different characteristics in terms of the PET technology itself and the populations and the

characteristics of the studies, it was not possible to get pooled estimates for all the questions that we intended to.

So I'm going to present the results for the entire technology report and the results for the first question about the diagnostic accuracy. And Karen is going to present later the results for the questions from Q2 to Q4.

So we started with 12,568 citations that were identified through the electronic searches to end up with the number of 112 studies that were included into the report. You can see here on the flow diagram the reasons for exclusions of the studies, and they are basically related with the set of eligibility criteria that we were considering. The main reason, as you can see, sometimes the studies did not satisfy the criteria for two reasons at the same time, but we chose only one of the reasons to make the final decision for exclusion.

You can also see the distribution of the number of studies in general. And we can tell that

the majority of the studies have been produced, of course, from 2003 to 2008, in the areas of cervical cancer, ovarian cancer, pancreatic cancer, and small cell lung cancer, basically. And we can see that that's a common denominator for each of the questions. But I will present this information in a more detailed way right now.

So for the question about the diagnostic accuracy, I'm going to present the results by type of cancer. For the questions about the diagnostic thinking impact, incorporating FDG-PET as part of a management strategy and the economic evaluation, we're going to present the results in general.

So for bladder cancer we identified three studies. Two of them were prospective and were evaluating the use of PET, not PET/CT, and the main use for staging purposes. You can see the sample sizes of the studies range from 35 to 55.

We were able to provide a pooled estimate of the efficacy of PET for staging purposes compared with any reference standard. This analysis was based on

two prospective studies, including 88 participants. And we found that the positive likelihood ratio of 4.88 was not statistically significant.

I'm just going to give you some guidelines now so you can see for the next positive and negative likelihood ratios in the next slides, how can you interpret the data. Every time the positive likelihood ratio crosses one, that means that the measure is non-statistically significant. If it's not including one, it means that from a statistical point of view, it's a significant finding.

The other thing that you might want to keep in mind is that the largest positive likelihood ratio for a test, the better. The smaller the negative likelihood ratio, the better for the test. So you can see here that not even the positive or the negative likelihood ratios are statistically significant.

I tried also to provide some sense about the sensitivity and specificity values in the individual studies. And you can see that there is a wide variation in sensitivity. It goes from 50 percent to

94 percent. And specificity values for FDG-PET in bladder cancer goes from 72 to 100 percent.

For brain cancer we identified five studies. Three of them were prospective studies. All of them were evaluating the use of FDG-PET. They are basically mostly related to staging of the disease. Sample size varied widely from 17 to 81 participants.

In this case, we were not able to provide any data about the accuracy, pooled estimates for any of the clinical indications that were considered. Sensitivity values, you can tell, this is a dramatic change from 7 percent to 63 percent values in sensitivity reported in the individual studies. And same with specificity, ranging from 14 to 100 percent.

As I mentioned before, most of the evidence has been produced for cervical cancer. We can see that 21 prospective studies have been produced, and the majority of them have evaluated the use of FDG-PET. The indications -- the majority of them, again, are addressing the problem of staging of the disease. And this is a finding that is consistent between the

studies about FDG-PET and FDG-PET/CT.

Sample sizes are -- I mean, they have a wide range of variation from 14 to 517. And the same can be said about the sensitivity and specificity values. Sensitivity ranging from 40 to 100 percent and specificity from 50 to 100 percent in the individual studies.

These are the results of the Meta-analysis, so basically we were able to produce Meta-analysis for two indications, for recurrences and for staging. For recurrences, only information about PET and for staging for both PET and PET/CT.

You can see the comparators there. And what we can see here is, in the first case, there were -- and this is the result of three studies that provided information by site of the body where recurrences might be identified. So what they -- and I mean, it was interesting that the three studies were reporting the data on the same types of sites of the body.

What we found is that the positive likelihood ratio for identified recurrences in the

mediastinal region was statistically significant, and the same to identified recurrences in liver or spleen.

On the other way, PET showed statistically significant results to identify negative -- sorry, for the negative likelihood ratio for peritoneal lesions and again for mediastinal lymph node. The values are good from a statistical point of view for both the positive and the negative likelihood ratio.

The same results and statistically significant findings were identified for staging purposes, both for FDG-PET and FDG-PET/CT. But the best value in all these indications was for the identification of recurrences, in particular mediastinal lymph node lesions.

For kidney cancer, we found eight studies. The majority of these are retrospective studies. And the distribution of the purposes range from staging to primary diagnosis, basically. Sample sizes range from 15 to 60 patients. And we were able to produce one meta-analysis on the use of PET for primary diagnosis and staging. And again, for the positive likelihood

ratio based on prospective studies, the results were statistically significant.

It's interesting that for the positive likelihood ratio when we compare this information based on retrospective information, this result was non-statistically significant. This is a very interesting finding because you would expect the other way because retrospective studies are more open to bias, and they might produce positive outcome bias. But in this case, what we found is that the positive finding was related to prospective studies, and it was not confirmed by retrospective evidence.

Sensitivity values range in individual studies from 47 to 90, and specificity values range from 66 to 100 percent.

Ovarian cancer was another condition that produced -- where a lot of evidence has been produced with 14 prospective studies and 6 retrospective studies. Most of the studies evaluated the use of PET/CT as compared to the use of PET only. And in terms of the purposes, recurrences again are the most

studies indication. Sample sizes are very dramatically from 13 to 101.

We were able to produce some meta-analysis about the use of PET for recurrences, both PET and PET/CT. And what we found here is again statistically significant results. It's interesting to see that in some cases, the confidence intervals are very wide. And in those cases, what is recommended is to -- I mean, you need to have some reservations on how to interpret this data because the wide confidence intervals might be due to small sample sizes because the rate of events to complete the two by two tables is low. Those situations might affect the width of the confidence interval. And again, the confidence that you can put in the significance of these results.

Pancreatic cancer, again 17 studies, most of them prospective, related with the role of PET as compared to PET/CT where only two studies were produced. The main purpose for the use was primary diagnosis and in some cases, a combination of primary diagnosis and staging.

We were able to produce two meta-analysis; one when primary diagnosis were considered together, and the other one was when the studies only considered primary diagnosis separately. For the first case, meta-analysis were produced for PET and PET/CT separately. And the results, as you know, were statistically significant, meaning that FDG-PET and FDG-PET/CT seems to be useful for primary diagnosis and staging purposes in pancreatic cancer.

Prostate cancer had only four studies, and we were unable to produce a meta-analysis because no common denominators in terms of the comparison in terms of the populations were found.

But I can say that sensitivity values, well, they showed a huge variation. Really nothing can be said in terms of an approximate value of sensitivity and specificity.

The same can be said with small cell lung cancer where ten studies have been produced, the majority of them prospective studies. And again, no pooled data were obtained to evaluate the accuracy.

And any conclusions that can be made about FDG-PET for this indication are based on individual studies, but not on the pooled summary.

Same for testicular cancer. Any decisions that can be made should be based on the results from individual studies.

So I'm going to leave you now with Karen. She's going to present the results for the questions 2 through four.

DR. SATYA-MURTI: May we ask a very brief question, very focused?

MS. OSPINA: Certainly.

DR. SATYA-MURTI: You mentioned the larger the LR positive the better.

MS. OSPINA: Sorry. I don't see it. Okay.

DR. SATYA-MURTI: The larger LR is better, and the smaller LR is not so good. We are aware of that. But how did you again calculate the statistical significance of positive and negative LRs?

MS. OSPINA: Well, the indication -- let me see if I can go -- I'm just going to show you in one

of the meta-analysis here. The positive and the negative likelihood ratio, they behave like a Knott's (phonetic) ratio. And the value of one means that the results are not statistically significant. Every time the confidence interval around the estimates include that value of one, that means that from a statistical point of view, the result is not significant. So that's the criteria to -- does that --

DR. SATYA-MURTI: Yes. Definitely. Thank you.

MS. OSPINA: So if you -- I know, I mean, this is a lot of information to be presented in 45 minutes. But if you go over the meta-analysis in the report, I guess that's something that can help you to interpret the results from the meta-analysis. Keep in mind that every time the confidence interval crosses one, the results are non-statistically significant.

DR. GULENCHYN: As Maria indicated at the beginning of her section of this presentation, the answers around questions two, three, and four are much more descriptive in nature because of the relatively

small number of studies.

When we look at the diagnostic thinking impact of FDG-PET or FDG-PET/CT, we eliminate two of the tumors as there was no information for those tumors. The relative small number of papers that describe this with respect to other seven tumors are discussed -- or displayed on the next two slides.

And I'm going to sort of talk about these in aggregate, although I will display the slides separately as there was just too much information to put all on one slide. It would have been absolutely illegible.

I think the first thing to note is that the sample size varies, but is relatively small in all of these papers, ranging from a low of 24 to a high of 102, when one looks at the first paper by Bang under pancreatic cancer.

The second is that there is a mix of technology being described here, both PET and PET/CT with PET/CT, of course, being much more common in the more recently published papers. In ovarian cancer,

I'd like to point out that all of the papers utilize PET/CT technology. And in kidney cancer and testicular cancer, all of the papers are using only PET technology.

There is a mix of indications that are being explored, from staging through to restaging or recurrence and then through to looking at diagnosis with respect to the kidney cancer.

The changes in management is the large -- is the endpoint that is most frequently reported. And they range from 7 percent in testicular to 69 percent in pancreatic, with a wide range extending from tumor to tumor.

There is one where diagnosis was looked at, and that is again in pancreatic cancer, the paper by Ruf in 2006, which indicated that there was an improvement in diagnostic interpretation in 25 percent of the cases.

Most of these studies are either of grade B or of grade C, i.e. there were methodological issues associated with these studies. There was one

prospective study, therefore giving it a grade -- I'm sorry -- grade A study in the ovarian by Simcock, 2006. There were 56 patients and that study was performed prospectively and was felt to meet the criteria for grade A.

There was one grade D study, and that was in the area of testicular cancer. So we have a retrospective grade D study by Karapetis in 2003.

So in summary, these two slides are showing a wide range of impact of FDG-PET/CT from tumor to tumor, with the numbers ranging between, as I said, a change in management of 7 percent, lowest for testicular, and 69 percent in pancreatic. There is no ability to further synthesize this data in order to come up with sort of a more comprehensive statement regarding FDG-PET.

In Q3 where we're being asked to determine whether FDG-PET or FDG-PET/CT can be used as a part of a management strategy, the number of papers is even smaller. This is probably why I volunteered to do this part of it and leave Maria with those huge

numbers to be able to deal with.

And here we are dealing only with five papers. And the results are summarized here. And I'm going to try to take a little bit of time to go through this slide. The first paper around brain was looking at the ability of PET to, in fact, predict survival. And it divided tumors into those with high uptake and tumors with low uptake. And tumors with high uptake of FDG, indicating and associated with therefore a higher grade CNS neoplasm, had shorter survivals. And there was an indication that PET could be used as a survival predictor

This was a retrospective study. It has major flaws associated with it. It was graded as a grade D study.

In cervical cancer there were two papers that looked at this particular issue. Both of them actually used a historical control group. So this was not a randomized trial, although it was a prospective trial. And in patients in whom -- in the first paper by Chang, in patients whose treatment was managed with

input from the PET study, there was some evidence that, in fact, the survival rate was longer, i.e. the treatment was more appropriately prescribed and patients lived longer at 22 months versus historical controls of 12.7 months. And the 95 percent confidence intervals are there. They overlap very slightly.

Lai 2004 also looked at this time restaging. And, of course, trying to separate restaging and recurrence is very difficult because you get information on restaging any time you assess a patient for recurrence, if the study is indeed positive.

And in this case, the two year overall survival rate as compared to historical controls was, in fact, not statistically significant. So there is a little bit of -- there is difference between the two studies that do exist in cervical cancer.

Ovarian cancer has been studied in a very similar way. SLL stands for second look laparotomy which was the comparator that was used. So this was studying whether PET could be used as opposed to

second look laparotomy and predict progression. And therefore direct treatment and therefore -- and what we were looking at was in the first study progression-free interval, and the second two disease-free interval in patients in the first place with negative results and then in positive results.

And so those -- the data has been looked at in a number of different ways. And in fact, there does not appear to be a huge difference between the two. Of course, the interpretation of that could be that PET could be used in place of second look laparotomy. So one study, 55 patients, retrospective design, but a little bit of evidence there that there may be some utility.

And finally, the pancreatic study, again using PET to assist in primary diagnosis, staging, and therefore direction of therapy -- sorry. And on this occasion, looking at responders versus non-responders to treatment based on FDG-PET/CT. There seems to be an indicator that PET may be able to identify patients with an improved prognosis. So sort of tantalizing

bits of evidence here that are beginning to tell us that PET and PET-FDG may be valuable in some tumors as part of a management strategy.

Cost effectiveness was the easiest one to do because there's one study in the literature on this particular group of tumors. And that is in pancreatic cancer. And this was a prospective study given a grade B. And this really is a fairly rudimentary cost effectiveness study and would be classified, I think, as a cost minimization study indicating that there may be net savings from the use of PET/CT at the rate of about a thousand dollars per patient when it is used to identify metastases and therefore avoid unnecessary surgery which is, of course, very similar to the use that it's being put to in non-small cell lung cancer. And that was the only study in this particular group of tumors that was identified that provided any evidence on cost effectiveness.

So I'm going to turn this over now to Maria to summarize, and then both of us will be available for questions at the conclusion of the presentation.

MS. OSPINA: So I'm just going to present a brief summary of the main findings of the report. So for the question about the diagnostic test performance, we found that the quality of the evidence in general for the nine types of cancer is moderate to poor. It's interesting to find that -- well, if we do some analysis by year, then most recent study seems to be improving the quality of the evidence.

Taking each of the types of cancer individually, what we can find is that for bladder cancer, the likelihood ratios were non-statistically significant for the indication of staging. For brain, we could not obtain pooled data. So any conclusions should be based on the individual studies.

For cervical cancer, the largest positive likelihood ratio was for FDG-PET to identify recurrences in liver and spleen. And the smallest negative likelihood ratio was for FDG-PET dedicated to detect recurrences in mediastinal regions. The positive likelihood ratio was statistically significant for both FDG-PET and FDG-PET/CT. And the

negative likelihood ratio was non-statistically significant for staging for both of the technologies.

For kidney, the findings of the positive and negative likelihood ratio were statistically significant for staging. In ovarian cancer, the largest positive likelihood ratio was for FDG-PET for the indication of identifying recurrences. And the smallest negative likelihood ratio was for FDG-PET/CT, again for recurrences.

For pancreatic cancer, what we found is that compared to FDG-PET/CT, FDG-PET seems to show some advantages for the diagnosis and identifying the initial stages of the disease.

For prostate, small cell lung cancer, and testicular cancer, we could not obtain pooled data. And again, results should be based on individual study data.

For the question about the diagnostic thinking impact of PET, the quality of the evidence again is mainly poor to moderate. I would say again, more moderate than poor. And the largest amount of

evidence for the use of PET was for cervical, ovarian, and pancreatic cancer.

We found that the treatment management strategy changed after PET and ranged from 7 percent in a study that was of a very, very low quality to a 69 percent in a moderate quality study on pancreatic cancer. And again, I guess it needs to be kept in mind that there is a lot of variation in these results by type of cancer.

For the question about PET as part of a management strategy, quality of the evidence ranges from moderate to very, very poor. And the evidence that is available is limited to cervical, ovarian, and brain cancer as Karen showed you.

Cost effectiveness data, there is limited evidence for pancreatic cancer only. Again, I emphasize that these are results from 2003 to 2008.

This technology assessment has strengths and some limitations. Some of the strengths are that it's based on a sound methodology. We tried to provide quality control in every single step of the review

process. We are using sound methods that have been endorsed by other groups such as Cochran Collaboration, the diagnostic test groups.

Some of the limitations are we restricted our inclusion to studies that were published in English. So there is some potential of publication bias in this review. And also, there might be the possibility of some grey lit bias. We did not include results from abstracts from scientific meetings into this report. So that's another limitation of it.

So basically there are some important points to finalize this presentation. And the quality of the evidence on FDG-PET for the nine cancers has been poor to moderate. But again, I think there is an improvement in this area as most of the prospective studies have been produced in the most recent years.

The evidence as we can see is largely confined to diagnostic accuracy as opposed to the other questions that were considered in the report. And the largest amount of evidence for the questions in this report was found for cervical, pancreatic, and

ovarian cancers. And the evidence about the cost effectiveness is restricted to pancreatic cancer with only one study produced since 2003.

So this is what we we've done. Thank you very much.

DR. SATYA-MURTI: Thank you very much. We have about two to four minutes time for very focused questions. Not general comments, but very focused. I was keen on finding out what the statistical significance on ovarian is. Your ratios are pretty good, positive and negative, but they still were not significant. Were they?

MS. OSPINA: Yes. It was. Yes. Yes. Well, you can see. And these are the largest -- these are the best estimates that we were able to find. Both of them are significant and --

DR. SATYA-MURTI: They are significant, the ovarians?

MS. OSPINA: Pardon?

DR. SATYA-MURTI: The ovarian was significant?

MS. OSPINA: Yes. Yes. It was.

DR. SATYA-MURTI: Okay. All right. Thank you.

DR. JUHN: I have a couple questions. The first is that I saw that in your inclusion and exclusion criteria, you limit it to studies that had more than 12 subjects. What was the reasoning behind that? Why was 12 chosen?

MS. OSPINA: That was -- actually was one of the criteria that one of the clinical investigators proposed to make a cut-off. It's an arbitrary cut-off. That's all that I can say. I guess that the impression he has is that -- and we tried to put it a little bit lower. Traditionally, what reviews do is to make a cut-off of more than 10 participants.

Again, I don't think that there is any evidence to show that studies from 10 or larger or smaller than 10 makes a difference. I'm not aware of that. But it's like a convention in the area that you try not to include the smallest studies.

I think that also might have an impact in

terms of the evident rates that you are calculating. So that might be one of the reasons why it was chosen.

DR. JUHN: Actually, my concern is that the number is too low because my kind of thinking in this area is that those numbers are generally higher. You want larger sample sizes in order to actually draw conclusions from individual studies.

MS. OSPINA: I understand your point. We were just trying to keep a balance between -- I mean, there are some advantages. And of course, these -- I mean, making a decision based on the sample size is a very hard decision when you are trying to do a systematic review because precisely one of the powerful characteristics of a review is that you empower the small studies if you are able to provide a full estimate with all the small studies.

DR. JUHN: Regarding the pooling of the estimates, did you include the grade of the study? So did you include all the studies, or did you only include grades A, B, or C? Or did you include all of them?

MS. OSPINA: Well, in -- we included B and C in the analysis. We did not make any restrictions in terms of quality in the pool estimates of the meta-analysis. And that would be a matter of more sensitivity analysis itself. Of course, the impact of combining studies with different quality, that will reflect in the measure of heterogeneity of the results.

In some cases, because these analyses are based in the best of the cases on three or four studies, the numbers are too small to make a sensitivity analysis by the quality of the evidence.

DR. JUHN: My last question really has to do with, I think there were probably what, over 200 studies? Or what was the total number of studies?

MS. OSPINA: 112 studies were included for all the questions.

DR. JUHN: Okay. So I was -- I was actually stunned by the fact that only eight of them actually had a grade of A in the quality of study. So if you could comment on kind of what was the -- why the

general level of the quality of all the studies were so low.

And then secondly, I guess I'm a little concerned that the investigators who are designing these studies, you know, all of these criteria for how to evaluate are kind of readily available, you know, quite evidence based. Why they're not actually incorporating that into their study designs.

DR. GULENCHYN: I'm going to try to answer that question because I think it perhaps wants to this group and a broader group as a whole to answer that question, not just me.

I will identify first of all that I come from Ontario in Canada, which has probably taken the world's most rigid view of how PET should be introduced and is introducing it through studies because studies that have been sponsored by the Ontario government, because they determined that in fact the type of work that was being done was not sufficient for them to make a decision.

I think what we're witnessing here, speaking

as a physician, and I've been in practice for 30 plus years, is that we're seeing a transition in how the diagnostic imaging world thinks about the introduction of technologies. And so the way that technologies have been introduced in the past, which is for usual single site investigators to apply them to a group of patients under the authority of their own REV's and to look at how they impact on patients, thus resulting in relatively small sample sizes and publications initially of case reports and then slightly larger publications is in transition.

I think it's changing. But the change -- and the change has really only occurred in the last certainly five to ten years. We do now have publications of standards through the various different initiatives that have been undertaken by professional societies. But it takes time to do those studies. And it takes time to get them out and into publication.

In the case of the Ontario studies, they've been in process now since 2003. And we've just had

our first abstracts presented in June of this year. So I think what we're witnessing is a change in the way things are being evaluated. And that change is underway at the current time. And we're only now beginning to see some output from those changes.

I think Maria referred to that when she said the more recently published papers have been -- have tended to be the ones that have achieved the higher quality ratings.

DR. SATYA-MURTI: Thank you. Yes. One last question.

DR. WAHL: You said that one of the advantages of the way your approach was that it was well done in terms of the design. And I was just wondering, though, in terms of the approach to evidence based medicine, when you have a mechanistic underpinning to a process, is it -- and this is increasingly becoming important as we understand the genetic abnormalities associated with specific cancers and targeted therapies that target those abnormalities -- is it reasonable to slice and dice accelerated

glycolysis into all of the disease states and treat them as if they're not linked biologically?

Because they are very obviously linked biologically. This is a consistent and rather common process across cancers. Do you think the methodology you're using adequately addresses the underlying biology and physics which seem common? In almost every case, the likelihood ratios that are positive are very positive. The likelihood ratios that are negative probably reflect that microscopic disease is undetected. And they're not ever zero, which would seem to be true across all of your analyses, but with different confidence intervals.

And I'm just wondering, how does your analysis deal with the underlying biology? It seems to ignore it.

DR. GULENCHYN: Okay. I'm going to answer that, first of all, with a question -- with an answer that relates to how this study was designed and what we were asked to do. And then I'll give you my personal opinion after that if that's okay. All

right?

So I'm -- this technology assessment was designed by CMS and AHRQ. And this was what the University of Alberta Evidence Based Practice Group was asked to address and asked to do. And therefore, the questions -- the approach that was taken was as a result of that particular design.

But I don't want to come across personally as sounding tremendously defensive, which is what I feel like I sound when I say that. And I think that there are some arguments to be made for looking at this more as a biological process and in summing data. Where that may fall down would be where there is particular interference in terms of the interpretation of the results.

And that would occur in cases, as we all see, with many nuclear medicine tests, not just this one, where you have urinary excretion or normal accumulation of the material in gut or in brain or around the heart and, therefore, that may interfere to some extent with the interpretation of the test.

So there may be an argument that can be made to look at things in a more comprehensive view. But that is my personal opinion and not related to the work that we were asked to undertake.

DR. SATYA-MURTI: Thank you. I think we should go on for the next two presentations. And then -- we'll just slide from one to the other, and then hold off on questions at the end of Dr. Mankoff's presentation. Dr. Hillner is presenting next.

DR. HILLNER: Good morning. I'm Bruce Hillner. I'm the lead investigator of the working group of the National Oncology PET Registry. During the question sessions my fellow co-investigators will not -- will certainly not be bashful and be able to involve themselves in the discussion.

Before I spend a few moments and my first six or so slides reviewing the background of the design of the Registry and then the subsequent results, I think it's important from the get-go for my discussion to say that we're presenting only in our slides the results that are in the public domain of

peer review publications, which are the one-year results of the summary findings across all cancers within the Registry.

We have -- I have prepared subsequently, either as a hand-out or additional slides that will provide information on the treatment monitoring findings of the NOPR Registry, as well as the findings for the nine specific cancers that the questions were put to you to address. So those information I'm not going to show you here. But I can either show subsequently or as a hand-out.

This project was sponsored by the Academy of Molecular Imaging to provide the funding to get started. But it's important I think also for you to know that the Registry is self-sustaining with the $50 dollar user fee per case, that it's a self-sustaining project that the data management is coordinated through the American College of Radiology, and it's imaging network. And you see the endorsements as well as that the Center for Medicare Services was an advisor in our design.

Stuart briefly touched upon the idea that as data accumulate, additional cancers may be covered for PET scanning and that the Coverage with Evidence Development program provides a clinically appropriate compromise of data collection via the Registry that provides important care information, at the same time striving to have -- to minimally impact the flow of patient care.

NOPR covers all the cancers that are neither specifically covered or non-covered by CMS up so far as 2004. We have detailed evaluation for 18 different cancer types that we can share with you. For low prevalence cancers, there likely will never be an adequate quality evidence to support a coverage decision outside this mechanism.

This is a national program. All Medicare eligible PET facilities can participate. The program is entirely funded by user fees, as I said, after the start-up process. CMS reimbursement depends on timely reimbursement of data submission. No data, no payment.

The objective of the Registry is to assess the effect of FDG-PET on referring physicians' plan of -- and here's the key message -- intended patient management as a surrogate for what their actual management is across the spectrum of expanded cancer indications for PET.

There is other registries and other questionnaire legacy studies that have showed, not a perfect, but a high correlation between intended management and actual management.

Within that primary objective, secondary objectives are to look at the role on intended management when looking at referring physician, does that change by specific type of cancer, the specific indication, that is for example, initial staging, diagnosis versus suspected recurrence. I specifically pushed to have inclusion of the following two characteristics to be able to see if patient performance status and the role of if the physician was the provider of the service versus a diagnostician, if that influenced their decision

making and the stratification of PET, the specific technology. But with time, that's gotten to be less of an issue. Currently, 90 percent of the studies are PET/CT.

Any PET facility in the United States that is an approved to bill Center for Medicare Services for the technical or global charges can participate. They're responsible for collecting the data and sending it via an internet web-based data collection to NOPR.

Patients -- the study is open to all Medicare beneficiaries. There is no consent necessary to submit data to NOPR for the required transmittal to CMS. Patient oral consent is necessary for inclusion in our research data set, which is the data that I'm going to show you.

Let's review four slides about process. If you're the referring physician, you have to complete a pre-PET form of five questions and return it to the PET facility prior to the PET scan. Once the study is done -- I will subsequently show you a graphic of this

sequence -- there is a complete -- they have to complete a post-PET form which is four to seven questions and return it to the facility within 30 days. Consent on the post-PET form is requested to the referring physician to have their data in the data registry.

So here's the time flow of what would happen if you were getting a PET scan within the Registry from left to right. The referring physician believes the PET scan is appropriate for patient management. They can -- they need to complete the pre-PET Registry form shown by the red arrow.

The patient is asked for consent on the day that they're at the facility to get the scan. The PET scan is done. After completion, the PET scan in interpreted and reported in standard operating process. And that report is sent to the referring physician. That referring physician is then sent a post-PET form including a questionnaire about their consent that asks them questions about, in light of the PET findings, what is their intended management.

Thereafter, at the far end, the ongoing patient management occurs. But we do not collect data on what actual management is beyond the 30-day period.

So the pre-PET form asks five questions; the specific cancer indication which the five are listed here; the cancer type; the disease stage using classic summary staging; the performance status using ECOG classification; and intended patient management plan.

After the PET scan, there are questions that are customized for the clinical indication. Questions related to diagnosis are different than for treatment monitoring. There's three to six for each indication. Most of these are check box yes, no answer form. And the consistency across the pre- and the post-PET question is, what was your intended management plan before the PET scan was done, and now in light of the PET scan, what is your subsequent intended patient management plan? And in addition, in light of the PET, are you going to continue to -- what is your role in the cancer care management?

What I am subsequently going to show you is

the results from the first year results that were published in the Journal of Clinical Oncology in the spring of 2008. This is a snapshot of the abstract. This is the first year of data. We can subsequently, as I said, give you individualized cancer results for the first two years.

So from the first year, there was 22,000 consented cases from over 1500 facilities. The flow diagram shows you the eligible cases to the actual study cohort. 88 percent of patient who were eligible had both physician and patient consent for participation.

The technology profile which was 84 percent was PET/CT. This is not a university medical school's study. 71 percent of the studies were non-hospital based. 76 percent of these were at fixed facilities. Currently, there are over 90,000 completed cases of PETs that data are available for evaluation. So the numbers here are just a logarithm different than what you heard or two, compared to what you heard in the preceding speaker.

Before getting to the results, we thought that it would be relevant to give you a sense of the number of scans that are done per year compared to new cases of cancer in individuals over 65 to get a sense of the magnitude of PET scanning per new cancer as far as the level of penetration across the country.

So here is shown eight common cancers, the number of NOPR based scans that were done in 2007. Again, this is not -- this is age-adjusted incidence of individuals estimates over age 65. It shows that the level of penetration ranges from about 3 percent of new prostate cancer patients are getting -- got a PET scan within the Registry to our estimate of approximately 38 percent in ovarian cancer. Across the 8 cancers, most of the strong cluster tendency is about 10 to 12 percent of cases -- excuse me -- of cancers had a PET scan done in individuals over age 65.

Here is the main result that is shown. And I'm going to take a few minutes to walk you through it. The columns represent the cancer indication which

were quite balanced as far as the number of cases of diagnosis, staging, restaging, and suspected recurrence.

At the top it's showing you where there was a concordance between the pre-PET plan and the post-PET plan, and shows that there was approximately a 62 percent -- but they markedly differed between the indication. Change in the intended management plan was dominated at three to four times as frequent of moving from a non-treatment plan to a treatment plan versus going from a treatment to a non-treatment plan. The overall summary of a 36.5 percent change in intended management.

An alternative way of thinking about this is to think about the pre-PET plan as far as the type of plan that it was. Approximately 40 percent of patients would have had an alternative imaging CT, MR, ultrasound, et cetera. Within that universe of patients, slightly more patients went to initiating new therapy, 47 percent, while 37 percent of patients would have had subsequent watchful waiting, supportive

care.

Our group thinks some of the most powerful findings are the other red areas of attention. Approximately 30 percent -- excuse me -- about 15 percent of patients would have had biopsy. Instead of -- in light of the PET scanning, only 24 percent of patients continued to have biopsy as their intended management. 76 percent had a change to either initiating therapy or observation, which were almost -- which was an equal distribution.

In patients who a watchful waiting or supportive care strategy was the plan intent, approximately ten percent of cases, approximately one quarter of patients, the physician changed his or her management plan to initiate therapy.

Within the individuals that treatment would have been done if PET was not available, 15 percent would back away, usually because of evidence of more advanced disease than anticipated to supportive care or watching.

Eight percent of the time, there was a major

change in the treatment modality. We don't have time to go into details for you with that, but the classic would be to move a shift of category of therapy from chemotherapy to radiation, surgery to radiation, et cetera. That occurred approximately nine percent of the time.

As I noted that the left column, 40 plus percent of the time the intended management was alternative imaging. As a worse case scenario, we suggest that including the cases where the pre-plan was alternative imaging may overestimate the impact of PET. If we take as a lower boundary for assessing the intended management, we analyzed the data assuming no benefit from those case in the numerator. But we include all those cases in the denominator of the evaluation. The imaging-adjusted change in management was just under 20 percent of 19.4 percent with a very tight confidence interval as shown here at the bottom.

What about impact on other tests or procedures? This was a specific question that was asked of referring physicians. It's their opinion.

We cannot tell you what tests they subsequently didn't actually do. But in their opinion, 76 percent of referring physicians indicated that PET would impact their future care by avoiding additional tests or procedures.

In the working group's opinion, the strengths of the Registry are that this is from the real world. This is timely. This is data over the last two years. These numbers are unprecedented in their size. They're representing current technology, 85 percent PET/CT. In 2008 the data is 92 to 95 percent PET/CT.

We believe without having to drag this out, that there is other evidence that good observational studies usually match controlled studies in magnitude and direction of effect. These results are similar to a study that I did a couple years before under the funding of AHRQ of a single center more tightly managed institution that looked at our first year results of what PET was introduced at my university.

The limitation, we have said over and over

in our publications is that we collected on a surrogate endpoint of intended management, not actual management. We don't know if the management changes were in the correct directions. This was not a clinical trial. We do not yet have information on if it improved long term outcomes.

NOPR cannot address the question and was not designed to address the question if PET should be used in lieu of or as a complement to other imaging technologies. The optimal sequencing of PET, CT, MR is needed. And those sorts of evaluations require clinical protocols with imaging adjusted modulations of clinical strategies. We cannot address that question. We also cannot tell you incrementally how much better PET is compared to the next best legacy imaging technology.

Our global summary is that intended management associated with PET changed in previously non-covered cancers. And this NOPR Register was similar to similar institution studies of covered cancers.

One-third of older patients undergoing PET for cancer types covered under the Coverage with Evidence Development program had a major change in their intended management, including type of treatment. The relative impact of PET on intended management was observed across the full spectrum of cancer indications and showed minimal -- minimal variation across the 18 different individual cancer types.

As I mentioned at the start of my remarks, there is other work that we have prepared or have in press. Our work related to PET for treatment monitoring during a course of chemotherapy with or without radiation involving 10,000 NOPR cases is in press. And we are prepared to show you slides showing the breakdown of that data either as summary or for the nine cancers of the question that was put to the panel.

We also have submitted for publication, but only by a few weeks -- don't have accepted yet -- the assessment from just under 50,000 cases of individual

cancer types for 18 individual cancer types of the 4 categories of cells that we are also prepared to show you.

Those results show -- sure, there's a little variation from 32 percent to 42 percent here. But in the bigger picture, the NOPR Registry reflects the biology of and not individual cancers.

We are fully prepared, and I am eager to do evaluations where the NOPR data, the individual identifiers, because we know the identifiers, we know when the technology was performed, to be linked to CMS claims data to track subsequent clinical outcomes.

We are also interested in using that data because we have physician opinion on prognosis in light of the PET scanning, and we have the PET reports. So when the PET report is abnormal, for example, we could look at SUV values as far as that being a predictor of prognosis.

Our request to CMS in this reconsideration is specifically for coverage of diagnosis, staging, and restaging. But our working group opinion is that

the nine cancers is too narrow. This should be for all the cancers that are covered within NOPR. We do not find that there are specific individual cancers that stand out as being exceptionally richer or exceptionally distinctly clinically different.

We therefore eliminate the correspondence for continuing going forward with the Coverage with Evidence Development for data collection for these classic indications, but continue the CED indication for treatment monitoring.

Thank you for your attention. And now or later I'll be happy to answer your questions.

DR. SATYA-MURTI: We should go on with the next presentation before we open up for questions.

DR. MANKOFF: Okay to start with the next presentation?

Good morning. My name is David Mankoff. I'm professor of radiology medicine and bioengineering from the University of Washington and Seattle Cancer Care Alliance in Seattle. And I've been asked to present some additional information and data from the

prospective of the four societies that you see represented here. I should point out that this represents a combination of societies that are experts in cancer imaging, but all societies that are experts in cancer care and treatment as well.

My take home message is simple. And it echoes the last statement that was made by Dr. Hillner, is that FDG-PET and FDG-PET/CT is a broadly applicable technique to cancers in general and that, again, I think the evidence in the NOPR and some of the additional data I'll present to you this morning support the broad applicability of this test to cancer.

While we would expect a need to be able to customize the approach for individual cancers, and we would naturally expect some variability in performance across different indications for different cancers, the underlying biologic principles and the potential impact upon clinical care is broad enough that, again, the argument is that this is a broadly applicable technique.

And so I'm going to present three sets of data to help support that claim. Number one is that I'm going to show you -- at least summarize some studies and data that suggest that FDG is a biologically robust tumor marker. Basically, the glycolysis is a fundamental and important portion of tumor biology.

Number two is that based upon reanalysis and ongoing data collection and studies in previously covered indications for FDG-PET and FDG-PET/CT, remembering that in prior meetings we had similar presentations of technology assessment showing you fairly sort of moderate level of evidence, that subsequent ongoing studies have actually continued to support the accuracy and the utility of these studies for these indications. And in fact, we're beginning to use some of that data to refine our approach in some of the previously covered cancers.

And then finally, I wanted to highlight some studies, many of which were included in the technology assessment, but perhaps not completely brought out in

some of the new cancers that are being considered at this meeting.

So it's very important to go back to the fundamental process here which is the radiopharmaceutical that we're using to generate the PET and PET/CTs fluorodeoxyglucose, which is a very close analog of glucose. It has a single substitution in the two position. And it very closely mimics the kinetics and biochemistry at the two critical and rate limiting steps for glycolysis. That is the delivery and transport into the cells and at the phosphorylation step at hexokinase.

And there's been fairly extensive validation. We're dating all the way back to the foundation work of Dr. Sokoloff (phonetic) in the 70s showing that deoxyglucosine (phonetic) and more recently fluorodeoxyglucose is a fundamental marker of glycolysis. So we're measuring tumor glycolysis.

Now, I think some of the most exciting data that have come across in the past ten years really and some of them earlier are data supporting the very

early observation by Otto Warburg in 1931 that aberrant and accelerated glycolysis is a very fundamental process associated with cancer and the cancer cell.

A body of work -- this review comes from Dr. Peterson's (phonetic) group at Hopkins -- has shown that this is not simply a question of nutrient availability, that accelerated glycolysis occurs under aerobic conditions when plenty of nutrients are available. And that it is really actually a very fundamental part of the way that cancer cells operate biologically.

And there have been a number of studies showing that aberrant glycolysis and in some cases aberrant uptake of FDG is associated with processes that we know are fundamental to cancers and that we know are associated with more aggressive cancers. This includes work coming from Dr. Wahl's (phonetic) lab that showed that hypoxia would contribute to uptake as you would imagine.

I think some very exciting work coming from

a number of labs, including Craig Thompson's (phonetic) lab at the University of Pennsylvania beginning to unravel how aberrant glycolysis is associated with fundamental growth and other important properties in cancer, very likely mediated through the molecule AKT, which takes a very fundamental role in the life and death of a particular cell.

And actually some of these more recent observations are going back and being able to provide the biologic underpinnings for some speculation that was done in the mid to late 90s suggesting that aberrant glycolysis may be a fundamental way where the cancer cell avoids cell death that is normally programmed into non-cancerous cells and allows it to grow under sometimes stressful conditions without undergoing what would normally be naturally programmed cell death in normal cells.

So increasing data on both a clinical basis and again on a molecular and scientific basis showing that aberrant glycolysis, which is what we're measuring with this technique, is a very fundamental

portion of cancer biology.

Now, I think one of the other clinical data sets that I think tends to get underemphasized in the technology assessment approaches is that there's a wide variety of publications coming out showing that for a number of different cancers that are treated in a very wide range of fashions with everything from surgery to chemotherapy to radiation therapy to targeted treatment, that FDG uptake is prognostic. That the level of uptake tells you something about the aggressiveness of the tumor and the likelihood that the tumor is going to progress and kill the patient.

And again, these are tumors with very widely variant clinical presentations, sometimes widely varying biologies, and certainly widely varying treatments. And this association which occurs across all of these different tumor types I think provides further evidence of the fundamental association between FDG uptake and the malignant phenotype.

Now, my favorite example comes from a study that was done at Sloan Kettering on thyroid cancer

patients. These are iodine negative or iodine refractory thyroid cancer patients which is one of the approved indications.

I like this 'cause I, as a practicing nuclear medicine physician, actually see these patients because they often get treated with radioiodine. And when I run out of steam with radioiodine, we very much turn to PET as is the indication to be able to go back and identify those tumor sites which failed to take up iodine in thyroid cancer, but may be seen by FDG-PET. And in fact, the data continued to support that in about 70 or 80 percent of the cases, we're able to identify sites of disease that we can't see when those tumors redifferentiate and begin to lose their ability to take up iodine.

But I think one of the under-appreciated points here is that it's not only that we're getting staging information, we're really getting very fundamental information about the cancer because we can go back and say, you know, we missed 20, 30

percent of the cancers in this situation because they don't have FDG uptake. But what these data show you is that in a very advanced patient population, none of those patients who had absent FDG uptake, even if they had widely metastatic disease died. And in this otherwise indolent cancer, we had a rather profound death rate with about 50 percent survival or 50 percent death rate at 3 years in the FDG positive peeps.

So not only is this a staging tool that helps up with restaging and making clinical decisions, it really is fundamentally associated in many studies with the cancer phenotype and provides us information that I think is actually very difficult to get in technology assessments that are focused appropriately on things like sensitivity and specificity.

Now, the other thing I wanted to point out is really the fundamental change in the technology that has been cited before because a lot of the literature we have is based upon FDG-PET, but more recently based upon FDG-PET/CT. And here we're again

combining two rather fundamental approaches to cancer. One is a functional approach based upon glycolysis as I just showed you. But our ability to identify shape, location, and their densities has been a fundamental part of our approach to cancer diagnosis since the inception of diagnostic imaging.

Now we're allowed to marry these two techniques together. And this is an example from a study that looked at the improvement in diagnostic accuracy, showing that the combination of PET and CT even versus separate reading of the PET and CT improved the accuracy in correct staging.

And there have been a number of other publications, including many folks in this room that have contributed to this. But this is the kind of thing that you see. If we were trying to read this spot showing up in the pelvis as a nuclear medicine physician, saying, there's a hot spot over in this corner of the pelvis, yeah, we can identify that that's an abnormality. But we may not have confidence that that's an abnormality related to cancer as

opposed to a normal variant.

When we can all of a sudden marry this with an anatomic technique which has its own limitations that didn't really show this particular lesion well to begin with, but then marry them together, we can say, aha, that's a hot spot in the middle of a bone in a patient who may have lytic bone metastases that might not show up on CT. And I can have quite a bit of diagnostic confidence that this is an abnormality. And in fact, many of these studies will support that.

I now want to move on to show you some highlights of ongoing information from previously covered trials that continue to support the fact that this is a very useful technique and that it, in fact, shows that based upon prior covered indications, the field has not stopped investigating these approaches since they've been covered, and in fact, has used data to be able to refine their approaches to diagnosis and sue of the test.

So one of the more I think powerful indications is that on the NCCN guidelines which have

more and more become some of the standards for practicing cancer diagnosis and treatment, many of the cancers, both covered and non-covered, are beginning to be included in those guidelines and take an important and appropriate role in the diagnostic workup and, in some cases, treatment direction for these individual cancers.

Now, in the case of lymphoma, this is perhaps gone the furthest, where the presence or absence of FDG uptake at the end of the disease has been recognized as being fundamentally important and prognostic to the disease management. In fact, so important that the latest generation of lymphoma trials that are out in Europe and beginning to happen in the States as well are really using FDG-PET as a ploy for making adaptive therapy decisions.

And that led to this position paper which was published and widely cited from the Journal of Clinical Oncology that not only talked about the utility of this technique, supporting some of the early data that had been presented at previous

meetings like this, but also said we need to develop actually lymphoma specific guidelines so that we can begin to use this in clinical trials and apply those to clinical practice. And so this is an example of work that has gone on subsequent to a prior coverage decision.

This is a study that was presented by Dr. Gulenchyn, the Ontario study that she mentioned in the Society of Nuclear Medicine 2008. And I hope in the question and answer period, if I make a mistake Dr. Gulenchyn will correct me on this.

But this is a multi-center randomized control trial, so an approved indication. We now have randomized controlled trial going on in the Canadian health care system in Ontario. And it was a randomized control trial between conventional imaging in lung cancer and PET/CT. And again, although the numbers may be slightly different than what was published in some of the retrospective studies, it is showing a significant improvement in upstaging and avoidance of inappropriate surgery in this randomized

controlled trial.

And very importantly from the standpoint of cost effectiveness is that a single FDG-PET/CT may be able to replace a number of other components of conventional imaging in cancer staging. This would lead to a considerable simplification and actually lead at the same time to some cost savings.

At this year's ASCO in the clinical science symposium where I actually was privileged to be able to moderate it and Tony Shields who is in the audience was one of the discussants, there was a Dutch study that was presented that was again another randomized controlled trial in restaging for metastatic colorectal cancer, one of the approved indications, the randomized FDG-PET versus no PET prior to a planned laparotomy for apparent isolated liver disease.

The primary endpoint was the futility of the resection. This was a relatively small study, but empowered to identify the primary endpoint. And in fact, they found a 38 percent relative risk reduction

in futile laparotomy that was statistically significant and came to the conclusion that, again, the addition in this case of FDG-PET -- most of this was done prior to PET/CT -- was helpful in reducing the number of futile laparotomies in this fairly common patient population.

Now, an area that I've worked on in particular is breast cancer staging. And this is an area where I'd say we've used some of the data that have come around to help us to refine the indications and use of this particular technique. So there are compiled data, and actually again, I think supported by another abstract that ASCO presented by Dr. Gulenchyn, showing that FDG-PET/CT as a systemic staging approach was not particularly helpful in early stage breast cancer. I'll get to areas again where it was helpful.

And perhaps this is not surprising because prior using convention imaging have also shown a low true positive rate and high false positive rate in trying to stage very early stage breast cancer. We're

talking about stage one and early stage two disease.

On the other hand, the data continue to show very promising results for more advanced disease, for stage four disease where FDG-PET/CT has been very helpful. So as an NCCN guideline panel that was assembled, the specific recommendations coming out of this was that PET was not recommended for systemic staging of early stage disease.

And again, we're using ongoing evidence to help refine the approach to using these techniques for these common and previously covered cancers.

Finally, I want to highlight some of the data that relate to the covered cancers that we discussed that were perhaps not brought out by some of the previous discussions and the technology assessment. So one of the technology assessment cancers where it was difficult to draw conclusions, in part because there were not a lot of specificity data, was in small cell lung cancer.

Now, here sensitivity and specificity are obviously an issue. But the primary clinical issue is

trying to determine limited stage disease versus extensive disease, which has a fairly profound impact upon management. And there are three papers -- and I believe all of these were included in the technology assessment -- that very strongly showed that, number one, PET upstaged patients in many cases and upstaged them in a highly accurate fashion and had a, at least modest, but fairly significant impact upon therapy. And then importantly, at least two of these papers came to the conclusion that this approach could simplify staging.

So, yes, in a strict sensitivity and specificity basis may have been difficult to assess. But the application of this to the particular clinical diagnostic test at hand has actually been pretty well supported by the existing literature.

And this is just an example to show you why that is. I'm going to see if I can get that slide to move. Nope. Sorry. If you could see this, there would be spinning diagram on the bottom. But what you can see is that this is -- oh, thank you. Nope.

Didn't want to do it. Thank you.

Part of the reason this is such a powerful technique for this is small cell lung cancer is one of the more glycolytic cancers that we've seen. And in fact, this provides a very high contrast study so that you can get very good information with a lot of contrast, sometimes not necessarily a lot of detail. But it's ideal for doing these whole body surveys and trying to determine limited versus extensive disease.

Thanks. Maybe we can go back to the slides now.

So one of the areas discussed was cervical cancer restaging and for documentation of recurrences. This is data from Washington University, Dr. Siegel's center, that looked at perhaps an important -- a related but perhaps -- and certainly is a more important question related to the restaging.

This was a prospective study done in two stages. And what they did in this situation was to look at treated cervical cancer patients, look at their presence or absence of metabolic disease at

three months post-therapy, and then use that as a comparator for subsequent progression free survival.

So to put this in slightly different terms, we're doing an FDG-PET in the sense of a restaging standpoint, and we're identifying disease that may be there, it may not be there. But very importantly, we're trying to identify whether the presence or absence of that disease predicts what happens to the patient.

And really very striking data that you see up here, and you see the curve is represented down below, showing that the presence or absence of FDG post-therapy and to the extent in which we see new disease versus persistent disease was very strongly predictive of survival.

Now, what does that mean? Well, a very interesting portion came into the second half of the study where, as these data began to emerge and the Center was identifying these patients at very early stages as being failures of convention therapy, they were able to institute salvage therapy in some of the

patients that had PET persistent disease. And in fact, there are eight patients within the study that had a successful response to salvage therapy and had improved outcome at that time.

So not only do we do a good job of staging in this particular instance, but again, coming back to the fundamental properties of PET as a marker of cancer, we see that this predicts outcome and can actually be used to make critical decisions at an appropriate time point.

Another area that was covered and actually shown to be beneficial in the technology assessment was in the restaging and staging of ovarian cancer and suspected recurrence. There was an additional study that was again cited in the technology assessment, but looked prospectively at a very problematic area, which is a pelvic mass in a patient who otherwise has a high index of clinical suspicion, a study from Risum and colleagues. It was published in Gynecology Oncology.

Here's a PET/CT showing a supercuvicular (phonetic) recurrence in that situation. And it had

very high sensitivity and specificity in a diagnostic application in a symptomatic patient population, a patient with a pelvic mass and a high suspicion of disease where really the only appropriate -- or the only accurate approach thus far has been a surgical approach.

And PET provided a very appropriate diagnostic alternative in that setting. Again, not necessarily covered in the technology assessment, which shows the broad applicability of these different techniques.

Finally, I would cite the data coming from Sperti. Again, this was also cited in the technology assessment and felt to be one of the higher quality studies for these particular lesions which are intraductal papillary mucinous neoplasms -- I hesitate to say that quite this early in the morning too fast.

Now, this is a very interesting lesion. It's actually not the most common lesion in the pancreas. And it's a little bit of an artifact of the fact that we've got good cross-sectional techniques

that we're beginning to identify some of these cystic lesions in the pancreas that we didn't necessarily see before. And it presents a tremendous diagnostic dilemma because sometimes you have needed a surgical approach to be able to make a decision about whether or not these are benign or malignant.

And in fact, what this study shows, that especially in the symptomatic patients, but even in the asymptomatic patients, the FDG-PET in a diagnostic application was highly predictive of benignity or malignancy in these lesion and, in fact, was very helpful in these patients in trying to guide the surgical approach.

Now, why do I cite this study in this relatively rare cancer? Again, it's the broad applicability of this technique. We saw data supporting the use of FDG-PET in pancreatic cancer. Here's a relatively unusual tumor and in an early study, again, the approach to diagnosis looks very promising. Again, underscoring that this is a broadly applicable technique to cancer.

So to summarize, I think there are good data that FDG is a biologically robust marker of cancer. And maybe equally importantly, rapidly evolving data showing that aberrant glycolysis is a very fundamental part of the cancer phenotype. Clinical data in prior approved indications continue to support the use of FDG-PET and, in fact, are being used to refine some of the guidelines that exist in practice. And then as we look at some of the other cancers that are now being investigated, we find other forms of data and other aspects that continue to support this application.

So again, there are going to be individual variations in how we use this technique, individual variations in its performance across cancers. But I think based upon what we've seen in the NOPR data and based upon these other data that I presented today, we would recommend broad coverage for FDG-PET and PET/CT for cancer diagnosis, stating, and restaging of suspected recurrence as was highlighted at the end of the last presentation.

Thank you.

DR. SATYA-MURTI: Thank you very much. We have about ten to twelve minutes. If we have questions, we can pose them at that time, and then move on for a very short break.

I have a very brief question. In the Registry, the participants knew that their future -- their answers might influence future coverage by Medicare. Did they or did they not know that the pre- and post-PET answers would alter the potential for coverage?

DR. HILLNER: I don't think explicitly they knew that. I have no way of speculating about the average provider in the community. There was nothing in our consent forms, for example, that said this data will subsequently be used for subsequent reconsideration. So that is a reasonable speculation. But I have no data to ground it.

DR. SATYA-MURTI: Okay. Thank you. Dr. Bergthold and then Dr. Sloan and Dr. Henderson.

DR. BERGTHOLD: I have a question for Dr. Hillner. Given that there's so many limitations in

the NOPR study and there's so little as far as I can tell that you can really say about kind of the outcome, the actual outcome instead of the intended outcome, I was a little surprised to see your request that CMS cover all of these cancers for PET scanning and stop the CED data collection.

It just didn't connect for me. I mean, I don't see how you make the leap from sort of a pretty poorly designed study that doesn't really tell you what happens and has so many limitations and now, therefore -- and same here, we recommend broad coverage.

The quality of the data that was in the tech assessment was really stunningly poor. Only nine -- or eight or nine rate A quality studies. And yet as a result of that, we recommend broad coverage. Can you connect those two for me?

DR. HILLNER: Well, we respectfully disagree with that conclusion. But my individual opinion is that if the concern of actual patient -- actual management and actual outcome is the coin of the realm

that would require a change in coverage decision, that the design of the Coverage with Evidence Development process would need to be changed.

That continuing getting additional data within the structure of where we've collected data I think provides no additional new insight. The confidence intervals across so many of the cells are so narrow that the profile of what I would present to you would not change. So the fundamental design would require reconsideration.

DR. BERGTHOLD: Excuse me. So you're not suggesting -- so you're not suggesting that CED be dropped, but that it be changed in design. Right?

DR. HILLNER: I am suggesting to your question that if you view that our results of intended patient management are insufficient for the request, that continuing the Coverage with Evidence Development with the current NOPR design for an additional year will not change your -- will not change your reluctance because the confidence intervals are -- the data are not evolving. Getting an additional case is

not going to change what we already are finding.

DR. SATYA-MURTI: Dr. Sloan?

DR. SLOAN: Yes. I have a question for Dr. -- forgive me if I mispronounce your name -- Gulenchyn.

A number of the studies cited -- so Dr. Mankoff just gave us a talk where his conclusions were quite different from yours and cited a number of papers that you did not cite, that it would seem to me that, you know, they were published in the time frame that you looked at. I'm wondering why some of the papers he cited were not listed and do not appear to be reviewed by your group.

DR. GULENCHYN: Okay. I'm going to try to answer that, but I may have to go back and actually look up the contents of the paper. I believe two of the papers that he cited were, in fact --

DR. SLOAN: More than that.

DR. GULENCHYN: Were in the technology assessment?

DR. SLOAN: Okay. Well, some were, but some

weren't. For example, just looking, you know, at small cell, Brink and Fischer, Schwarz was not cited. I didn't -- I didn't go through all of them. Risum was not cited. So at least four of the studies he cited, you did not cite.

DR. GULENCHYN: Okay. I would have to go back and look at the specific reasons why those studies were excluded. I --

DR. SLOAN: 'Cause they have an N greater than 12, and they were published during the time frame that you reviewed.

DR. GULENCHYN: Okay. Let me -- Maria is going to address that.

MS. OSPINA: I guess we can do that if you want. There is a list of excluded studies by reason for exclusion at the end of the report. So we will have to go and look at those references and see where they were, in fact.

There is another possibility. And it's that the study was not captured even by the filters that we used for the search strategy. And that's something

that -- I mean, there is no perfect systematic review in terms of capturing all the available evidence that was produced.

So there are two possibilities here. One, the study was excluded on any of the grounds that we mentioned, and then we will need to look at -- if you give me the full references, I might be able to track the studies right now.

The other option is that the studies were not captured by the filters in the search strategy. And in that case, now that we know that those studies are out there, we will have to incorporate that data into the report.

DR. SLOAN: I wonder, it looks like you looked only under specific cancers. So a paper that covered more than one type of cancer probably would have been excluded. Is that correct? Is that assumption correct? So if a paper covered more than one type of cancer --

MS. OSPINA: Oh.

DR. SLOAN: -- would it have been excluded?

MS. OSPINA: There was one reason is we required that the studies provide data by type of cancer. Right? And the results would have to be presented by the type of indications that we were considering.

In some cases, for example, there -- and it comes to mind, I don't know exactly the references. But I remember there were some cases where they presented, for example, gynecological cancers overall. But there was no specification about ovarian or cervical which were the conditions that we were looking at.

There were other cases, for example, where let's say information was presented by type of cancer. But the numbers for that type of cancer were not enough and were less than 12 participants. So you see, there were a lot of conditions that we required for the studies to be included. And although it might look weird that those three or four studies were not included, that's another option that these studies might have been identified, might have provided

information by the type of cancer. But maybe the numbers were less than 12 participants.

So those are the factors that we need to look back and -- look back at the report to see what the reasons were for exclusion, if the study was, in fact, excluded.

DR. SATYA-MURTI: Dr. Henderson?

DR. MANKOFF: I also wanted to point out that there are situations in which the questions that were posed to the technology assessment may not be appropriate for the cancers. So in an area that I think you're likely to be very familiar with, in brain cancer, systemic staging doesn't make any sense. It's a disease that only metastasizes to the neuraxis, and so staging is not an appropriate question.

However, the appropriate question there which would I think be documentation of recurrence and prognosis was not one of those questions. There are a number of studies that weren't cited in that, including some work from our institution showing that the predictive capability of the technique exceeds

conventional imaging like MR.

But I think the question in some cases weren't -- wasn't posed for the technology assessment. And that may have been one of the reasons it wasn't there. And I did try to cite some of those important indications to indicate broad utility.

DR. HENDERSON: I have two questions for Dr. Hillner. The first, you emphasized appropriately that there was no follow-up after 30 days. But you didn't explain why that decision was made in the design of the study.

Second question is, I was rather surprised at the high frequency in which physicians and patients refuse to give consent, particularly since that was a requirement when they requested the PET scan. Why such a high percentage of patients eliminated for that reason?

DR. HILLNER: Approximately eight percent of patients did not participate and four -- three to four percent of physicians. I'm looking to my colleagues if we have -- we have not -- I mean, it's suspicion of

big government is the extent of my answer. That if they chose not to participate in the research project, they still got the clinical service and that their scan would be -- would be paid for.

But other than fear of big government, I don't have an answer for that. What I can tell you is that we have looked over time for trends of that. And in the first six months compared to -- in six month cycles that that's been minimal variation over the 24 plus months that we've been open. It's been very, very stable.

The core primary endpoint of change in intended management was negotiated. That was negotiated between the Center for Medicare Services representatives, the whole concept of evidence -- of coverage with evidence development. And that was negotiated honestly before I was even recruited to participate in the project.

The balance was felt that the burden of -- burden to place on participating centers and referring physicians that there was no carrot for the level --

for -- you know, the referring physicians are the key to the project and that their time was felt to be respected and what was the minimum data set design that could get us relevant information.

You can agree or disagree on our design. But this was not retrospective. This was -- you know, before the design they had to indicate what they intended to and what they subsequently planned to do. And it would have been a different story if we were presenting today if each of these cases was pre-identified so that we could give you data on the claims assessment on that. But that was -- that's the history of how the process worked out.

DR. SATYA-MURTI: Thank you, Dr. Hillner. We have copious amounts of time in the afternoon for further discussion. So if I may, Dr. Lichtenfeld, you had one question.

DR. LICHTENFELD: Well, since I'll be limited to one question 'cause I do have some comments. But I'll save that for this afternoon.

I want to go to Dr. Mankoff. And Dr.

Mankoff, you know, we're on the horns of a dilemma here today because we've had the technology assessment that says one thing. We have your comments which say open it up. I want to make sure I understand the basis that your society consensus statement, your inner society statement, is really that last sentence on that last slide.

Do you basically recommend that PET scanning be opened up for everything?

DR. MANKOFF: Right. You know, I think we're --

DR. LICHTENFELD: That I don't misinterpret that because --

DR. MANKOFF: No. And if I might, you know, add a little bit of interpretation to that, is that there will be cancers and indications where this will work better and not work better for other areas.

However, I think the documentation of looking at what's happened to the prior covered areas and the refinement of the field and the fact that when we look broadly even at some of the smaller studies in

some of these rarer tumors, we see very similar results in terms of diagnostic performance and outcome.

And I think this is somewhat of a practical decision. We can -- you know, part of the reason we saw less data in the technology assessment is that perhaps other than things like cervical and ovarian cancer, some of these are rarer cancers.

So we have two approaches. We can do a randomized controlled trial for every cancer and every indication that's out there, including these rarer cancers. We'll spend a fair amount of money doing that, and we'll take a while to collect those data. Or we can try to identify trends that look broadly applicable within the cancers and do what we do with the rest of the practice, is to continue to refine our application of both therapy and diagnostic techniques as we appropriately use them in the field as professionals with guidelines from our professional societies and ongoing support through a variety of mechanisms for continuing to test these and refine

these.

And that's why we spend a fair amount of time looking at those previously covered indications 'cause, in fact, that's what's -- that has been what happened, at least in some of those indications.

DR. SATYA-MURTI: Okay. Thank you, Dr. Mankoff. Dr. Phurrough has a comment, and then we'll break for 15 minutes.

DR. PHURROUGH: Just to clarify based on some of the comments and questions over the last several minutes. In making coverage determinations on diagnostic tests, we have in the past and I suspect will continue in the future, used patient management -- change in patient management as a sufficient condition of coverage.

We don't necessarily like to do that. We need to be confident that changes in patient management is a good surrogate for improved outcomes. But, in fact, that's why we commissioned this particular CD process to look at patient management issues. And that's why we have this first question,

which really focuses on is there a change in patient management.

So while in diagnostic tests we like to have patient outcomes, and there are some decisions that we've made that have required patient outcomes depending upon the maturity of that particular technology, it is not unusual that a coverage decision on a diagnostic test is largely based upon change in patient management.

DR. SATYA-MURTI: Thank you. About a 13 minute break, and we'll come back.

MS. TENENBAUM: Good morning. I'm Cara Tenenbaum. I'm with the Ovarian Cancer National Alliance. Before we get started, I do want to disclose -- excuse me -- disclose that we don't have any financial ties to the Medical Imaging Technology Association. But we have done two awareness and outreach projects with them. So I'll be happy to talk to anybody about those in more detail, if you'd like.

I also want to let you know that our organization has a very clear policy that we represent

patients and not necessarily our donors.

Also, before I get started, I know other people have said this. But I'm going to use PET scan as shorthand for FDG-PET and FDG-PET/CT scan.

Also, a disclaimer. I am not a doctor. I'm not a statistician. I'm a patient advocate, and I'm here representing patients and their families.

So all of you can read this. And I'm not going to -- I'm going to try not to say what's on the PowerPoint slide. Ovarian cancer, as many of you know, is probably the deadliest gynecologic cancer. This year 21,000 -- over 21,000 women will be diagnosed with ovarian cancer. 15,000 of them -- not of them, but 15,000 women will die.

I apologize for the typo. There's actually a 45 percent relative five-year survival rate. So more than half of the women diagnosed this year will not survive five years.

Ovarian cancer has a high mortality rate in large part because there is no good early detection test and because the vast majority of women will have

at least one recurrence at some point in their life.

As we talk about the PET scan as a marker, there is no reliable tumor marker for ovarian cancer. There is one blood test, one serum marker, for those of you who are doctors that works for some women some of the time. And I'll get into the problems with that.

Also, more than 70 percent of ovarian cancer patients have a recurrence at some point in their life, as I said. And what often happens with these recurrences is that they happen with a shorter and shorter time period. And women become resistant to platinum-based therapies. And so, in essence, they run out of chemotherapies.

I'm with the Ovarian Cancer National Alliance. We're a national organization that represents ovarian cancer patients, families, women at high risk over the United States. We do this through a partner member network of over 45 national, state, and local members. We are dedicated to conquering ovarian cancer.

You all know about the Coverage with Evidence decisions, so I'm not going to go over that or the PET Registry too much. But I do want to let you know that we get a lot of calls in the office about PET scans.

"How do I get a PET scan? My doctor says I should get a PET scan." We always recommend that women enroll in the Registry when possible. And so we're very happy that it's been available for patients.

Ovarian cancer has an issue with initial diagnosis. As I said, there's no early detection test. It's also not possible to biopsy the ovary the way you could biopsy skin tissue or breast tissue. So what we need is something that eliminates the need for an exploratory biopsy because what that does is it takes out an entire -- doctors take out the entire ovary.

So using PET scans, surgeons can appropriately be aggressive in debulking, which is really a bad term for taking out cancer, from women to

remove as much of the primary mass as possible.

The PET scans can also be useful in determining the original course of treatment. For example, I know women who have been told they need radiation prior to surgery. And a PET scan will show that actually surgery is the better course.

What happens for ovarian cancer patients also is that the tumor spreads so much, by the time you've found the fact -- diagnosed a woman with ovarian cancer, it shows up kind of like stars on the PET scan. So it's not kind of an isolated mass. And so the PET scan can help show the extent of the tumor proliferation.

In recurrence, again what we're seeing is that most women will have a recurrence. We're not always sure when. So there are symptoms of ovarian cancer. And there are women who are sensitive to the C-125 blood marker.

But this can be falsely elevated. For example, there are some chemotherapies that falsely elevate this tumor marker. And so a woman and her

doctor might think she's having a recurrence even if she's not. There are also a number of women and ovarian cancer sub-types that do not have a C-125. So for these women, they're really in the dark. And the PET scan is literally a way to illuminate what's going on inside her body.

We're also concerned, of course, with PET scans in terms of monitoring treatment, so how a woman is responding to chemotherapy. As I said, there are some women who will not emit the C-125.

And I know that for CMS, of course, you're concerned with the cost of care. And you don't want to pay for chemotherapy that's not working. And (unintelligible) oncologists are very careful about the sequencing of chemotherapies.

For me as a patient advocate, I'm concerned about patients getting toxic chemotherapy when they don't need it. I'm sure all of you know someone or have been through chemo. And as I just said during the break, I wouldn't wish it on my worst enemy.

So, of course, we're interested in reducing

-- you're interested in reducing the cost of care, and I'm interested in saving women's lives and increasing their quality of life.

So what we recommend is that PET scans be available when medically necessary for ovarian cancer patients, as we said, for initial diagnosis, monitoring response to treatment, and for recurrence. At the very least, we'd like the PET Registry to stay open.

And I do want to take this time to thank all of you for considering this issue. It's something that's very important to my patients. And I hope that you make a decision that is really in line with patient care.

Thank you.

DR. DURIE: Thank you very much. Ladies and gentlemen, I very much appreciate the opportunity to present information related to multiple myeloma.

It's not one of the nine cancers on your list, however, I want to make the case that it is a crucial cancer. It does require the availability of

PET scanning.

I'm here today as chairman of the International Myeloma Foundation, which is a non-profit entity based in the state of California, 501(c). I'm here, therefore, as an advocate.

I am, however, a physician who's specializes in the treatment of multiple myeloma based at Cedars-Sinai in Los Angeles. I am co-chair of the Southwest Oncology Myeloma Committee. And I'm currently conducting what I think is the largest trial ever performed, involving over 600 patients, to evaluate imaging as part of the ongoing management for multiple myeloma with central review of the images. I'm also involved in other trials related to myeloma.

So why should you listen to me regarding myeloma? Well, there are quite a few reasons. One is that it was included in the NOPR program that Bruce presented to you. You may not have noted that myeloma was at the bottom of his first chart. And it showed that there were 1,336 patients with myeloma registered through that program, which represented 13.1 percent

of myeloma patients.

So there is a substantial -- although this is in quotes, "a relatively rarer cancer," where it is definitely difficult to gather a lot of data, 13.1 percent of patients are currently getting scans through that mechanism.

There are several aspects of myeloma. Myeloma is a collection of diseases which can present in a variety of ways. It can present as a solitary plasmacytoma. It can present without the production of a monoclonal protein, which makes it extremely difficult to evaluate without scanning. So for this group of patients, some type of imaging is essential.

It can present within bone. But increasingly we recognize that it can present outside of bone. And in this sense, for 20 to 30 percent of the cases, it behaves exactly like a lymphoma, a plasmacytic lymphoma. And the case I'm going to make is that myeloma should be treated from a coverage standpoint like lymphoma because it behaves in exactly the same fashion.

There are also secondary cancers. As patients live longer, about 15 to 20 percent of patients with myeloma will actually develop secondary cancers which are picked up on CT/PET.

So where does CT/PET stand? Well, what was presented to you before the break is that NCCN has guidelines. There are guidelines which incorporate CT/PET for myeloma for staging and monitoring. There are also other guidelines, international guidelines, which incorporate CT/PET for myeloma.

I published one of the first large papers on PET imaging and myeloma. It was in 2002. I'm actually slightly glad that it wasn't reviewed by the Canadians. I'm not sure whether it would be an A or not.

The one thing I can tell you, though, was that it was actually picked as the best clinical investigations paper published in the Journal for the year 2002. So I think it was a reliable study.

It did show that myeloma is 100 percent positive. Myeloma is positive. It's a PET avid

cancer. If you have a benign form, MGUS, it's negative. So it's extremely helpful in that distinction.

23 percent of extra-medullary disease, so they behave like a lymphoma which is already a covered indication. The results are prognostically important. If you have extra-medullary disease, you are definitely going to die sooner. We need to know this information because we have better therapies for myeloma.

One important thing about myeloma that makes it special, besides the fact that I study it, is that it's a disease which is called multiple myeloma. There are multiple lesions. It is a disease which involves the whole body.

So if you want to look at it from a cost effectiveness standpoint, you have two choices. One is that you can use an MRI. You can MRI the spine. Now, the disease, however, is affecting the spine. It's also affecting the ribs. There's a soft tissue plasmacytoma close to the kidney here. There are a

variety of lesions.

So that the only way to effectively and efficiently and cost effectively scan the whole body is to do a whole body FDG-PET scan. MRI for the equivalent coverage would cost you $7- to $8,000.

And so because of this, there have been guidelines published in 2006, a schema whereby CT/PET could be integrated into the routine staging and monitoring for myeloma.

It's also been linked -- you heard about the linkage with the biology and the uptake of glucose and glycolysis. Well, it turns out that PET scanning does link to one of the fundamental biologic features of bone disease and myeloma. And that is the production of a DKK1 protein -- and this is from the New England Journal -- where there was clear correlation with the predilection for aggressive bone disease and this DKK1 protein production.

A very important study was just published from a French group showing that if you have a plasmacytoma, and you do either MRI or PET scanning,

MRI missed a total of 18 lesions. Whereas, with the whole body PET, you picked up 10 additional lesions.

What that meant is that approximately 30 percent of the patients, instead of getting radiation to a single plasmacytoma, got systemic therapy because they were discovered to have not a solitary plasmacytoma, but systemic disease.

So in 2008, we do need PET scanning for myeloma. 30 percent of patients are diagnosed earlier. We do need to stage them in some whole body fashion.

There are better treatments. Patients are living longer. So we need to be able to monitor them and assess the status of the disease. And we need to assess it using imaging.

Maybe I'll just show you this one. To get the point out about survival, 75 percent of patients with myeloma are not alive at 5 years. This means that this is not a disease where you have trivial decisions to make which might influence a few months of outcome. These are decisions that will impact

long-term survival with transplant possibly or not. So there are huge potential impacts on treatment decisions.

So in conclusion, I would say that the evidence does support the use of PET scanning for myeloma in a fashion similar to lymphoma. It's accurate for staging and monitoring. It does affect treatment planning. And, as was shown by Bruce, it was similar to the other cancers in the NOPR evaluation.

More options for treatment require us to have better decisions related to imaging. Thank you.

DR. SATYA-MURTI: Thank you, Dr. Durie. We have an open public comments period. There are four persons who are here to talk about this. Dr. Siegel, Dr. Coleman, Dr. Shields -- five actually -- Allison Colbert (phonetic) and Dr. Tunis.

So we'll give approximately two minutes for each of these open presentations. And if you can line up to the microphone so we would avoid the transit time, that will be nice.

MS. ELLIS: Down front, please.

DR. SIEGEL: Good morning.

DR. SATYA-MURTI: Good morning.

DR. SIEGEL: I'm Dr. Barry Siegel. I'm professor of radiology at Washington University in St. Louis, and I'm one of the NOPR working group co-chairs. I conflicted to that extent, and also to the extent that I earn my living doing PET and interpreting PET studies.

I really don't have specific prepared comments, but wanted to provide some commentary related to some of the questions that occurred during the earlier session.

First, the question related to consent. We were actually quite thrilled with 88 percent consent. And in the prospective planning of the trial and in consultation with CMS, we set 85 percent as the bar at which if we didn't hit that target, we were going to go back and redesign the approach to see if we could get better consent.

Moreover, CMS has all the data. And they

actually can look at the non-consented data and the consented data because they have the authority to do that because they're the government.

Just a point about the funding for proper randomized controlled trials of PET. We found ourselves for decades in what I have labeled a catch-22. The data aren't very robust as you've seen. And therefore, there's no coverage. There's no coverage. Therefore, we can't get the data from studies because there are really not alternative funding mechanisms.

Richard Wahl and Ed Coleman and I did a laborious study with NIH funding. We were very lucky to get the funding. It was a breast cancer study that showed that PET doesn't work for axillary staging, and it's non-covered.

But getting the kind of funding that it would take to do randomized trials is not realistic. The sponsors who make the PET scanners are not going to fund these studies. They have not funded these studies. The people who make FDG have not and will not fund these studies.

It's a commodity now. It's not a drug. It's not a blockbuster drug that you have a billion dollars a year in sales and you have to do the studies to get FDA approval.

Going to the NIH in the last decade to try to get funding for phase three clinical trials of PET, given the current funding environment and the way it's been, it's not worth the time to write the R01 or the U01 to try the do these kinds of trials.

So that really is the reality. And as co-chair of ACRIN, which is the American College of Radiology Imaging Network, which does imaging-related clinical trials, I can tell you that mounting these trials, pulling them off, is an incredibly laborious process, and maybe 50 years from now we'd have the kind of robust data that you might -- we'd all really like to see to be truly evidence based.

So what did we do? We said we've got a construct of evidence. We have evidence from previous questionnaire studies that showed a certain amount of change in management. We have subsequent studies that

demonstrated that when people said they were changing their management, they actually were doing it right nearly all of the time. PET was leading to the right decisions. And those right decisions led to improved patient outcomes.

So we said if we could take all these other cancers and show that change in management is in the same ballpark as it was for the previously covered cancers, that would be very important.

And we originally set the bar at a 15 percent change in management and frankly were delighted and surprised when it turned to be in the 25 to 30 to 40 percent when we broke it down by individual cancers.

And my final comment is in our asking for coverage for all cancers for diagnosis, staging, and restaging, and suspected recurrence, we're not blindly saying that PET should be used in every patient with all cancers.

What we're saying is that PET should be used wisely, guided by practice guidelines, guided by NCCN

guidelines, and guided by Medicare's own rules on when PET should and should not be used, which Stuart Caplan articulated at the beginning of the session this morning.

PET is meant to be a supplement to other approaches to imaging or a replacement to imaging. But not the end-all and be-all of all cancer imaging. Thank you.

DR. COLEMAN: Hi. I'm Ed Coleman. I'm a professor of radiology and Director of Nuclear Medicine at Duke University. For disclosure, like Dr. Siegel, I'm director of a PET facility. So some of my income does derive from doing PET imaging. I'm also an investigator with NOPR and do receive some support for that effort.

In addition, I have research grants from General Electric Health Care, consultant for General Electric Health Care and am a stockholder in RCOA, which is a mobile PET company.

I'd like to make a few comments concerning the CMS process for reviewing of PET and PET/CT. One

of the issues that we have identified from the start, when CMS decided to review PET cancer by cancer and indication by indication, was that some indications would never be covered that way.

Merckel cell tumor, a very low prevalent cancer, 500 patients a year, very FDG avid, behaves a lot like melanoma on PET scan, looks a lot like melanoma, and is very helpful in those patients. But at those numbers, you're never going to get a large number of publications. You're never going to meet the criteria from the Edmonton group for saying that there's data supporting its utilization.

So over the years, we've had discussions with CMS, what do we do with cancers like that. And that was some of the basis for the development of Coverage with Evidence Development. That way we can get coverage. We can get some data on the use of these cancers through the CED process. And I think the CED process has been extremely helpful in getting data on these less prevalent cancers.

It's interesting when you look at the tech

assessment that was presented today, the cancers with adequate numbers, the likelihood ratios were high and were significant. On those that didn't have adequate numbers, either you couldn't even develop likelihood ratios, or they were not significant in numbers.

So the numbers game is very important here. But if we're going to be able to cover cancers like the ovarian cancer, some others that are less prevalent, we are going to need larger -- we would need larger numbers to meet the evidence based criteria for coverage.

Thank you.

DR. SHIELDS: Hi. I'm Tony Shields. I'm a medical oncologist at Wayne State University and Karmanos Cancer Institute in Detroit. I do work with NOPR and have some stock in RCOA. And I've been working with them in designing and carrying out these studies.

I think intended management, which is the endpoint that we followed, is certainly a reasonable one. It matches what's happened before in studies

that have actually looked at outcome and in the randomized trials that have looked at what the change in management was and the benefit to patients.

The issue is, could we do this differently? And I think it's going to be hard. If you actually went to CMS and they have all the data and looked up what happened, how would that change the results?

If 100 percent of the patients got the intended management, that doesn't prove that it necessarily was the right intended management. If 80 percent got it, does that mean we were wrong?

I, really, as a clinician who takes care of cancer patients, I'd be surprised if 80 percent of my patients get the intended management. Something happens. Their kidneys fail. They get an infection. They decide they don't really want the surgery you're recommending. So a lot of things can change between the time you plan the treatment and they actually get it.

So I think the best way to really conduct these studies is the randomized clinical trials. And

the limited ones that we've been able to do as a community have certainly matched what the intended outcome is and what the planned outcome is.

But as you've heard from Dr. Siegel, to actually conduct randomized clinical trials in a wide variety of cancers, especially some that are quite rare, I don't think would be possible. So I think the outcomes that we've measured match all the data that's been accumulated so far.

Thank you.

DR. TUNIS: Good morning. I'm Sean Tunis from the Center for Medical Technology Policy. I don't have any conflicts that I know of related to PET. I guess I wish I did, but I don't.

And I was the Chief Medical Officer at CMS when -- in the early days of NOPR. So I really came up just to provide some historical perspective and perhaps say some things that Dr. Phurrough couldn't say and will probably be sorry that I said.

But this is really just to focus on one point. And I'm not trying to express an opinion about

the quality of the evidence, which is the main question before you.

One thing I did want to do is point out for folks who probably aren't familiar with the sort of inner workings of the NOPR, that this was an extraordinary accomplishment. You know, sort of comparable perhaps to the moon mission or something. I mean, it's amazing that it ever happened.

And I think, you know, Barry Siegel and Brian Carey (phonetic) and Steve Phurrough and Stu Caplan and others I couldn't name put in an unbelievable number of hours to make this thing occur at all. So it's remarkable that it exists. And you know, I just wanted to make sure people were aware of that.

You know, the other comment in terms of some of the questions related to, well, why isn't the NOPR better or why isn't the evidence better, and I guess the one observation I'll make about it is that, you know, the design of the NOPR in my personal opinion was a policy compromise to try to get as good evidence

as is possible within the constraints of wanting to provide expanded coverage, a reasonable magnitude of expanded coverage for PET scans, which was, I think, rightly considered a promising, but unproven, technology.

And everybody in this room who's involved with the NOPR knows how to do rigorous studies of diagnostic imaging. It's not mysterious. What's difficult to do is rigorous studies of diagnostic imaging while trying to provide broad access to the technology for the entire Medicare population nationwide.

And so in order to do that, to provide broad access and to collect some reasonable information on impact on patient management, hopefully that's relevant to impact on outcomes, you know, a bunch of smart people got together, and this is the best that they could come up with, given those constraints.

To have higher quality data, you would need -- you know, you would need some kind of control groups, meaning fewer patients enrolled, meaning that

you'd need some kind of funding mechanism, as Barry and others have said, for studies like this. And they don't exist.

So, you know, the mystery insofar as I'm concerned is simply that when you're trying to design scientific studies in the heat of battle while you're trying to make a coverage decision and provide access to promising technologies, you can't do a lot better than the NOPR.

And I guess the final thing I would say is, you know, given that this is an advisory group that wants to -- that, you know, has a role in terms of Medicare's thinking on Coverage with Evidence Development and evidence issues generally, the key broader points, I think, that the NOPR experience highlights is that, you know, number one, somebody's going to have to define what is minimally acceptable evidence standards for imaging. And in the absence of having some definition for that, you're going to be all over the place in terms of what kind of studies you get. And I know Dr. Phurrough knows that well.

But in the absence of having some consensus about minimally acceptable evidence and whether or -- you know, it's impossible to know whether the NOPR is above that or below that, and how is anybody supposed to know how to design one of these studies?

The second is, there's got to be some kind of funding stream for studies of this type. I agree that the FDG companies don't have the financial incentive. PET companies don't. But if you want to make evidence based policy like this, someone's going to have to pay for the studies.

And the last point is, the discussion about how to design these studies is never going to work very well if it's done in the heat of an ongoing coverage decision. It's going to have to take place outside the context of trying to make a policy decision for 45 million Americans about what will and won't be provided.

DR. SATYA-MURTI: Thank you. We are about two or three minutes ahead of time, or even four. If there are no other comments. We'll adjourn for lunch.

And when we come back, we have sufficient time for both intra-panel and panel, extra-panel discussions. And I'd like to lead off with Dr. Janjan and Dr. Wahl, who were inadvertently or necessarily cut off this morning.

Thank you.

(Whereupon, a luncheon recess was taken.)

DR. SATYA-MURTI: I don't have the authority to hit the gavel. Maria took it away from me. So if we may get started. Again, as I said, CMS has given us time for both questions to the presenters, and subsequently within panel, kind of an intramural discussion. So we have time for that, and then the voting. And Maria has housekeeping announcements.

MS. ELLIS: Good afternoon and welcome back. Just a quick reminder, if you are eating or drinking in the auditorium, please, please take your trash outside and discard it in the trash cans outside.

Also, for the panel members, you were asking questions in regards to your travel voucher. What will happen is, at the end of the meeting, you will

receive an email from me with your travel voucher. Before you even get back to your home, you'll have it. Okay?

DR. SATYA-MURTI: In opening up the questions for the afternoon, there was a question asked about left-out studies this morning. I think Dr. Sloan was asking -- I forget who -- and Dr. Ospina has the answers for those studies.

And then we'll open with that and then Dr. Janjan and -- who else? Dr. Wahl. Yeah. And then I'll restart the whole process after that. Thank you.

DR. GULENCHYN: I just wanted to answer the question that was asked this morning about studies that had been excluded, but were identified by Dr. Mankoff in his presentation. The -- all of the studies that were excluded were either because of the tumor type not being one of the tumor types that was under consideration or, in the case of the study by Patronis on brain cancer, it was out of the time period that was to be discussed. It was published in 1985 and, therefore, was prior to the time period.

And there is one study that we included, but in fact, did not reference in our presentation. And that's the study by Schwarz on cervical cancer. In that study, we are going to have to review the paper to determine whether or not we made an error in not including it under the discussion.

DR. JUHN: Was it actually in your technology review? I didn't see it in the back.

DR. GULENCHYN: It was actually in our technology review. It was one of the included papers. According to our records, it was included.

DR. JUHN: So it's not actually listed in here.

DR. GULENCHYN: There's something wrong with that particular paper. And we're going -- not wrong with the paper -- I'm sorry -- our assessment of the paper.

DR. JUHN: I looked at the abstract during lunch. And it looked like it was a prospective cohort. So it wasn't clear that there was a control group or a comparison group.

DR. GULENCHYN: But we are going to have to look at that paper. And there will be modification to the final document.

DR. SLOAN: There are a couple other ones that -- like on small cell by Brink and by Fischer.

DR. OSPINA: (Unintelligible.)

DR. SLOAN: I'm sorry. I couldn't hear her.

DR. GULENCHYN: All of the studies that were presented were in the report. They are there.

DR. SATYA-MURTI: All right. Go ahead. Questions from this morning, Dr. Janjan.

DR. JANJAN: Well, thank you. The question I really -- the conundrum we're struggling with here is, we've got a review that says one thing and other presenters that have other conclusions with much of the same data.

I guess the question I have for Dr. Gulenchyn is, how do we reconcile differences in conclusions even though some of the -- or much of the same data has been evaluated by presenters and you? Should you revise your conclusions, or should the

others be tempered?

UNKNOWN MALE VOICE: Who should revise their conclusions?

DR. GULENCHYN: I think, first of all, there have been two totally different approaches here taken. Okay? The data that we presented is from a technology assessment. A technology assessment has specific rules surrounding it. And those rules were very clearly identified by Maria when she did the first part of the presentation.

I think that the data that is being presented to you from NOPR -- largely from NOPR, but also from selected elements of the data that we presented, is simply being looked at in a very different way.

And I think that what you're trying to reconcile here is how do you evaluate these imaging technologies. And I think that -- is it Dr. Tunis or Mr. Tunis? Dr. Tunis's comments were very, very pertinent to this early on. You've got -- you've got technologies coming on board that are new, rapidly

evolving. This one, relatively expensive. And you want to know what difference do they make in terms of the patient.

There is no good way at the current time of funding these studies. In my own jurisdiction, the funding for doing very limited studies on five different cancers has, in fact, come from the Ontario government to do the studies that we're doing. The funding would not have come from CIH or our equivalent to the NIH.

And so it is a problem for you at CMS as to how to do this evaluation most appropriately. I don't think I can prescribe an answer to that, quite honestly. It's one that is a matter of debate, I think public debate, and a matter for public policy. And that's certainly not something that an individual can determine.

DR. JANJAN: I guess the question I had really was, you came at it from the point of view of what were the indications. And the other presenters were more or less stating that in their experience,

they found great value and were working backwards, where is it not indicated. And I don't know how we mesh that.

And the reason I asked this question is because we do have this disparate opinion. And for the record, when we make our votes, we have to understand where those two diametrically opposed positions are and how do we mesh them together to make a decision.

DR. GULENCHYN: I wouldn't characterize the positions as being diametrically opposed. I think what you heard from our presentation was a review of the literature that is -- a critical review of the literature that is currently available.

We were asked specifically to address specific tumors. And, therefore, we did not do a review of PET overall. We provided you with some evidence for some of the tumors for approving PET's use.

Is it from grade A studies? In large part, no. But there is evidence certainly that it does make

a difference in terms of patient management and also diagnostic thinking and determination of prognosis for at least ovarian, cervical cancer, and pancreatic cancer.

And to a lesser extent, if you look at the small cell lung cancer literature on an individual-by-individual study -- we couldn't pool the data because of the way the data was constructed in the various different trials. But there is also some data there. As to whether or not you should more broadly accept the indications, we did not specifically study that question.

DR. JANJAN: Thank you. It helps from the position of -- the clinical position. So, thank you.

DR. SATYA-MURTI: Dr. Wahl?

DR. MANKOFF: (Unintelligible.)

DR. JANJAN: Yes, sir?

DR. MANKOFF: I would like to agree with Dr. Gulenchyn that I think these are two different opinions of some of the similar data. I think the question is very different.

And I might cite as an example your Society of Nuclear Medicine study which I think is a pioneering randomized controlled trial on an existing coverage that very clearly shows a statistically significant benefit in patient outcome and avoiding futile surgeries.

But that study is actually not set up to answer the diagnostic questions. Those PET patients are not going to surgery. So we don't know the sensitivity and specificity of that result.

DR. JANJAN: That's right.

DR. MANKOFF: And so I have to -- with due respect to the Committee, I think some of the questions, especially for some of the indications -- we mentioned it in brain cancer -- are perhaps pitched towards technology assessment, but perhaps not pitched towards the way we're most likely to use this in clinical practice.

DR. JANJAN: Thank you.

DR. SATYA-MURTI: Dr. Wahl, it's your turn.

DR. WAHL: I yield this briefly.

DR. SATYA-MURTI: Go ahead.

UNKNOWN MALE VOICE: (Unintelligible.)

DR. WAHL: Okay. I guess I have sort of two questions. Hopefully that's all right. In the assessment that was done from the Canadian judges, they look like the likelihood ratios, you know, were all positive for disease being present and all relatively low for disease absent. The statistical significance was lowest in those studies that had the smallest numbers of patients.

And I was just wondering if, given the mandate to look at the data from 2002 and later, if that's really an appropriate way to make a recommendation on coverage because the Duke assessment previously, I think, only looked at six tumors. It didn't look at bladder, renal, or testicular. And the first bladder and renal papers were written in 1991. And there has been literature evolving in that period of time between 1991 and 2001, you know, 10, 11 years, where these less frequent cancers were evaluated.

And I do think, for instance, when we're

trying to make a decision on an infrequent cancer and there hasn't been a previous tech assessment on them, the process of not including the previous data I think is a potential systematic flaw and could lead to underestimates of the value of test just because of having low patient numbers.

Similarly, having done a trial on testicular cancer which was tissue confirmed and, you know, biopsy confirmed NIH funded prospective, published before 2002, it takes a long time to do that study in a single center. And in our center, we had excellent separation between patients with disease and not based on PET. But the numbers are small.

So I think it would be -- could you just maybe comment about excluding 11 years of data in those less frequent tumors?

DR. OSPINA: Well, first of all, I would like to clarify that the magnitude of the likelihood ratio has nothing to do with the sample size. The impact --

DR. WAHL: The significance does.

DR. OSPINA: Sorry?

DR. WAHL: The significance does.

DR. OSPINA: How big the confidence interval it is, yes. It might give a wide variation. But the value of the estimate you don't expect to be lower or smaller just because of the sample size.

Of course, if you have a small -- let's say a likelihood ratio of 45, right, and you have a confidence interval that goes from 0 to 1,000, what it's telling you is that that estimate is based on a very small sample size and that you are saying that the result, yes, is not significant and that might be a false negative result.

So, I mean, you are -- it would be occurring because of the low sample size on a type two error, meaning if you increase the sample size, the result might become significant. But it's -- the magnitude of the estimate itself, the point estimate, has nothing to do with that.

But I agree with you. It's the width of the confidence interval that might be affected by that.

When we received the request for the technology assessment report, it was based on an update of the Duke report covering the six types of cancer and adding three more conditions. So I would agree with you that in making a decision, you need to take in mind -- take a look at what the Duke report said.

In general, the conclusions, our conclusions are not contradictory between the Duke report and our report. Also, the gaps in research I think, unfortunately, they remain in terms of the quality of the studies. Although we've identified a little bit of improvement for the last couple of years.

But I would say that I would agree. You need to keep in mind both pieces of information for making these decisions.

DR. WAHL: My second question was for Dr. Hillner. And he referred to some data looking at change in management in these less frequent cancers, the ones other than the nine that we're asked to address later in the panel.

And I thought it would be useful at some point if he could just briefly show us what the -- or tell us what the frequency of change in management was in those less frequent cancers. And also if he could comment in any way on the changes in treatment planning, if any of those data are available.

DR. SATYA-MURTI: Now is as good time, if you want to comment on that.

DR. HILLNER: There was approximately 4500 cancers that -- excuse me. Our two year data, which is approximately 50,000 registrants, approximately 10 percent of that universe was outside 18 well-defined cancers. So they're a real hodgepodge of small little stuff. That 4500 cases has a change in intended management of between 35 and 40 percent.

DR. SLOAN: If I may, can I ask a related question?

DR. SATYA-MURTI: Yes.

DR. SLOAN: So you mentioned in your presentation that you -- you didn't present the data broken down by cancer type.

DR. HILLNER: Correct.

DR. SLOAN: It's probably -- I suspect it may be premature to do so. But can you generalize from cancer -- I mean, in other words, can you say of the nine cancer types we're looking at today, for example, can you say that the conclusions hold for -- the conclusions that you did present hold to some degree or more with all of them? Or do you have big changes in certain types of cancer and not so much in others?

DR. HILLNER: Chair, I have the options of either giving you a handout for everyone and then it can circulate back here, or I can show. I have two slides related to those nine specific cancers, if you'd like me to show additional slides.

DR. SATYA-MURTI: Maybe the slides after we are through with the initial round of questions, the pent-up questions.

DR. HILLNER: Well, why don't I at least give this so there's hard copy to the panel.

DR. PHURROUGH: Who has the slides?

DR. HILLNER: I have them on a CD or Barry --

DR. PHURROUGH: Is our slide person still back here? We'll see if we can get them on the screen while you're talking.

DR. SATYA-MURTI: All right. Yes. While we're waiting. Dr. Juhn?

DR. JUHN: I have a question for Dr. Mankoff and also maybe any of the other society presenters. I'd like your thoughts on two things. One is the technology assessment and what your kind of general reactions are to the Canadian judges.

And then the second is, if the grading of those studies -- again, I think we've had 8 A's out of 112 studies. It kind of reminds me of my organic chemistry course. So the question really is, why are we having such a deficit of good studies in this area given that the attention that's been paid by CMS to this whole PET scanning area has really been quite active over the last 8 or 9 years?

DR. MANKOFF: So I want to answer the first

part carefully because I respect my Canadian judge colleagues. And they're actually in both cases, the nuclear medicine physicians are physicians I know well and respect within the community.

I would come back to the fact that I think the questions are pitched differently. I think that the technology assessment where the individual questions are pitched the way you do a diagnostic or maybe a biomarker technology assessment. And they work very well in some instances, and they don't work well in other instances. And so I think some of it is the question that they were -- that they were forced to answer.

I want to come back to the point on your second question that Dr. Siegel made before. And that is, for those that are used to drug therapy trials where you make very nice prospective designs and maybe even those that are used to assessing technology when you're looking at a blood or tissue biomarker, this is a much more challenging and much more expensive approach to study designs.

And let me liken it a little bit to a diagnostic biomarker. If you're doing a biomarker study, you might do it in the setting of a clinical trial. You might collect tissue or blood, and then you might send that off to a central lab. And perhaps for as little as $10, $15 dollars a study, you can retrospectively go back and test the value of that study. And then you can go back and prospectively set that up with your next therapy trial.

Easy to do. The data is already collected. You're not relying on a prospective study to collect those data and validate it. And you can get very large numbers to assess these questions with good precision.

With an imaging study you can't do that. You have to collect every single study prospectively. You've got to pay for the study at the time you're putting the patient through treatment. And those are particularly pricey studies.

So I don't want to be too broad and sweeping in this answer. But I think if we look at the history

of diagnostic imaging tests, in certain ways PET has been put through a rigor that many of tests have not lived up to. We certainly haven't seen this kind of level of testing for something that we use every day in our practice like CT.

DR. JUHN: That might be next. Is that right?

DR. MANKOFF: Yeah. And maybe I'm getting my body imaging colleagues in trouble by saying that. And I do think that, again, citing the Canadian colleagues, that when given the chance to do that, after we have information from these grade B, C, and D studies, we're able to put together randomized controlled trials with proper funding that are actually, I think, pretty good trials.

DR. JUHN: So if I could just summarize. I think --

DR. MANKOFF: I'm sorry. I've droned on.

DR. JUHN: The second point is really that the degree of difficulty is higher.

DR. MANKOFF: Right.

DR. JUHN: If I could ask our Canadian colleagues here, is that, in fact, true, that the degree of difficulty for PET as an imaging modality is greater than other imaging modalities? I mean, is there a fundamental difference there?

DR. GULENCHYN: Well, okay. I'm answering from my own experience. I'm afraid I'm going to have to answer from the experience in Canada and not in the States. And I know that there are some differences. So I will stand to be corrected on anything that I say by some of my American friends here because I may be out of line on some of this.

It is a difficult technology in that it involves two components. It involves a drug, which in Canada there is a regulatory framework around. And you also have a regulatory framework around it.

The regulatory framework in Canada has, in fact, been applied more stringently than the regulatory framework has been applied in the U.S. And that regulatory framework in and of itself creates difficulties in the introduction of this technology.

And any difficulty that you encounter usually translates to dollars.

And so in my own particular case, we now have a notice of compliance for FDG for our product, which cost us something in the neighborhood of $1.3 to $1.5 million dollars. And I'm told that compared to Pfizer, I did it on a shoestring. So, you know, but it's expensive, and it's an obstacle.

Then on top of that, you have the cost of doing the installation of the imaging equipment. And in fact, that's the lesser cost. You have the cost associated with training an entire cohort to learn how to read and interpret these scans.

And then if you want to do it within the construct of a formal trial, you have the costs of designing and mounting that trial, which are not unsubstantial. I actually don't have the costs of the Ontario trial, but it's costly because it's taken us a lot of time to do it, and then the costs of the data collection and the costs of the data analysis.

And there is no -- there is no good -- there

was certainly in our case no good framework to have all of that in place. And if you're also doing this multi-centered, which you have to do, there's the cost of communication and all the rest of it.

So, yeah. It's costly, and there are obstacles to designing good trials. Now, mind you, we've overcome a lot of those in the last little while. And as Dr. Mankoff is pointing out, you're now beginning to see data emerging that are supporting decisions that were previously made by CMS.

So it is possible to overcome. But it is not necessarily easy to overcome. And this particular technology has faced challenges I think that were not faced by earlier technologies such as CT and MR.

DR. SIEGEL: Thank you. So a partial additional answer. I repeat what I said before, funding has been a major obstacle to doing --

DR. SATYA-MURTI: Can you identify yourself?

DR. SIEGEL: What? I'm sorry. Barry Siegel again. Sorry. My apologies. So funding has been a major obstacle to mounting the large really robust

category A clinical trials that need to be done.

But I would challenge one thing that you said. And that is, if you look back at the meta-analyses that have been done of CT and MRI, that literature has been just as badly dinged as has the PET literature and, if anything, in some ways, the PET literature is a little bit better because it's learned from the CT and the MRI literature, except that the studies tend to be smaller.

And so I don't think that the comparison is fair. And I think if we took CT, and we did CT for bladder cancer, or we did CT for Merckel cell tumor, we would find that the studies, if they came out B's and C, we'd be lucky. But CT wasn't put through that particular wringer. As I've said many times, PET is the whipping boy of high technology medicine.

DR. PHURROUGH: And you've managed it well.

DR. SIEGEL: Thank you. And we're seeing it here.

DR. SATYA-MURTI: This is not uncommon. Any new technology is held to a higher rigor than existing

technologies. I know several in neurology that won't pass muster.

I have a quick question, and then Dr. Phurrough has one. And then we'll start the process again.

Is there a way to quantitate the uptake, or is it subjective? Because this morning we were talking about some of that. Is PET uptake quantifiable in any validated manner? And if not, is there a subjectivity? I know some picture clearly show where it is located along with the CT.

DR. SHIELDS: I'll answer that one as the non-radiologist. This is Tony Shields, by the way.

We certainly do quantitate the uptake. We quantitate it a number of ways. For standard clinical, it's standardized uptake value which has been validated. There have actually been panels convened by the NCI and in Europe to look at reproducibility issues, to look at standardization issues, and look at measurements.

There are now studies ongoing in the United

States, which a number of us are part of, looking at reproducibility in hundreds of patients and looking at the use of the measurements to measure response to treatment.

For the most part, people have looked also at whether those measurements are helpful in the diagnosis of cancer, particularly in small pulmonary nodules. And those are somewhat helpful. But just looking above, background, in general, has been probably the best use.

But people -- and when we looked at our surveys at NOPR, most of the surveys -- most of the reports that we generate from the local community do report the standardized uptake values.

DR. SATYA-MURTI: So if quantitation as a confounder is not there, can a prognosis or management depend on have you further looked into the degree of uptake versus the prognosis?

DR. SHIELDS: Quantitation is there. It is used widely. It is reported. It has been -- as I said, there are methods that have worked to

standardize it. Those haven't been incorporated into every practice yet.

But to some extent, the fact that you can see the lesions, particularly clinically, is helpful above background. And people have looked at the intensity to provide some measure of activity and some measure of diagnosis and some measure of prognosis. So it's used for all of those.

DR. SATYA-MURTI: Okay. Thank you. Steve?

DR. PHURROUGH: Bruce, did you want to show your two slides right quick? And then I'll make my comments.

DR. HILLNER: There's four slides. Two I am first going to show, and then I'll pause, I think, related to the change in intended management of diagnosis, staging, restaging, and suspected recurrence. And then I'd like to be able to get your reaction, answer your questions.

I have two subsequent slides that deal with the treatment monitoring which have a different -- it's a variant because it uses different indicators

for change in management.

So I'm only going to go through the first two slides, which relate to what I presented this morning. Essentially taking everything I presented this morning and taking the data for the nine cancers that we -- the panel was asked to consider.

So to review the layout here, the individual cancers are at the far left. The indications are shown across the top. In the parentheses is the number of patients in the cell. And then the change in intended management is the primary result.

The dashes are, for example, diagnosis in cervical and the two in brain cancer, we did not prepare results when cells were below 50.

So in -- and I think I would urge you to separate out the diagnosis category compared to the other three, staging, restating, suspected recurrence. The message is pretty, I think, powerfully consistent here, that 32 percent at the low range up to 41 percent.

For ovarian, pancreas, prostate, small cell,

again, those cells, 32 to 44 percent, with thousands of cases in the suspected recurrence range. Testicular cancer, there was 176 across all the board.

To go back, related to diagnosis, because of the appropriate strict language currently that PET should be used only in cases where there's problems with getting classic anatomic characterization, there's both fewer cases and the relative impact is lower. As you see, 32 percent in brain, 25 in kidney, 20 to 35 percent in these remaining four cancers. There was again too few cases in testicular cancer.

I'm happy to answer any questions. Yes, Craig?

DR. HENDERSON: So the numbers in parentheses, for example, if you add up those numbers, that exceeds the total number of scans done for ovary, if I have added them right. I may not have, but it looks like it's about 4,600 --

DR. HILLNER: There's about 47- -- these are two years --

DR. HENDERSON: Oh, these are the two year

data. Okay.

DR. HILLNER: These are our two data, not one year data. These are two year data. So there was approximately 4700 cases of ovarian cancer in two years. And the vast majority of the scans within the Registry were done for suspected recurrence, slightly less for restaging. And as we know, that sometimes it's -- clinicians are not clear-cut about which of those two are.

So that's 4- to 5,000 cases in ovarian cancer. The vast majority for restaging, suspected recurrence. Approximately 40 percent change in management.

DR. HENDERSON: Okay. Can you give me an example of -- you're saying that a third of the -- I guess it's not a third of all the studies. But of all the studies where there was a diagnostic question, 306 scans, 35 percent had a change in the diagnosis as a result or the management related to the diagnosis, as a result of that scan?

DR. HILLNER: Correct.

DR. HENDERSON: And what kind of changes? Can you just give me some sense of what that means? Talk to me like a doc, you know, not like a statistician for a second.

DR. HILLNER: Well, it may be that instead of doing a biopsy, they are prepared to do an open laparotomy, you know, in the case of ovarian cancer. In pancreas cancer, the classic would be the decision for --

DR. HENDERSON: Wouldn't that be a staging? I mean, that seems to be more like a staging issue than a diagnosis issue.

DR. HILLNER: No. For ovarian and pancreas, I think are good examples of where you're getting dual information because if you do not have a histologic diagnosis of pancreatic cancer, but you have a high suspicion of it based on whatever clinical constellation and/or tumor marker, et cetera, that the PET scan not only if, for example, in the scenario I'm describing shows active mid-peritoneum activity, but also shows hot spots in the liver, that it is both

staging and diagnostic, suggesting metastatic disease to the liver.

But the indication is formally in our Registry would be diagnosis because there was not histologic confirmation of the diagnosis.

DR. HENDERSON: Okay. That's very helpful. So diagnosis means that the scan was always done before histologic confirmation?

DR. HILLNER: Correct.

DR. HENDERSON: And staging would be then after. Restaging then is after they've had at least initial treatment. Okay. So I understand.

DR. SATYA-MURTI: So that is one of the obfuscations I too had in my mind. When you say diagnosis, did that become a granuloma instead of a malignancy or some other indolent lesion that looked like a malignancy on the conventional imaging modalities? And, if so, that would be diagnosis for me. And did you look into that?

DR. HILLNER: We absolutely did. That it's of suspected cancer. Diagnosis with a category of

what was under consideration. That does not mean that the category of diagnosis means all those individuals wound up having cancer.

It was suspected ovarian cancer was the reason for getting the scan. There were other clinical indications that made that -- supported that concern. And the clinical constellation was such that felt that the PET scan could help clarify the situation short of getting a definitive tissue diagnosis.

DR. SATYA-MURTI: And how many might that be? Say for instance, in pancreas, I understand chronic pancreatitis could simulate a malignancy. So how many were there where entire diagnosis of malignancy was removed from the running?

DR. HILLNER: Could you help me again? Possibly restate the question?

DR. SATYA-MURTI: Were there instances where the entire diagnosis was revised to something other than malignancy? You may not have a number offhand. But I wonder how many were, say, retrieved from a

diagnosis of malignancy altogether and did not undergo even a biopsy?

DR. HILLNER: Barry, would you care to comment, please?

DR. SIEGEL: Dr. Siegel again. So the truth is, at the moment we can't answer that explicit question. But there are lots of patients where, for example, the indication for doing the test was a pancreatic mass discovered on CT. The PET scan was done to try to determine whether the -- and the pre-PET plan was biopsy.

PET scan was performed. PET was completely negative. The assumption was this is chronic mass-forming pancreatitis, a true negative in this particular case. And the plan now was to watch the patient. That's the kind of change in management that we're seeing in these individuals.

We're actually now preparing another paper which is going to specifically drill down and look at the diagnosis category, the suspected lesion, the paraneoplastic syndrome, and the patients who present

with an unknown primary cancer, and how well PET teases out that information.

Right now, they're all lumped together. All three of those diagnostic categories are lumped together in the data we've given you up to this point.

DR. HENDERSON: I had some more questions.

DR. SATYA-MURTI: Yes. One more.

DR. HENDERSON: I have a couple. Can I ask a couple? Only one?

DR. SATYA-MURTI: No.

DR. HENDERSON: So on the NOPR dataset, what percentage -- maybe you gave it and I missed it -- but what percentage of patients had either a CT or an MRI before they had their PET scan or PET/CT?

DR. HILLNER: No idea because the decision was made that that was too high a burden. I -- so nationally, I have no idea. I wish to tell you -- what I'm about to tell you is my study from four years ago, five years ago.

In the first year of PET being approved at this prior meeting, we audited intended management

pre, post, same design. In my study at one institution, the first 250 PET scans that were done and 76 percent of patients had prior CT or MR or both.

But I have -- we have no idea what the national trends are.

DR. HENDERSON: Was there a difference in the frequency in which management was changed in the group that did and did not have prior --

DR. HILLNER: No.

DR. HENDERSON: There was no difference?

DR. HILLNER: No.

DR. HENDERSON: So I'd like to move on to a second -- you know, what we do with what you've just told us. You've seen the forms that we've been asked to fill out with this table.

DR. HILLNER: Yes, sir.

DR. HENDERSON: Okay. So what you've presented here is very nice in some ways. So our first question here says, "How confident are you that evidence is adequate to conclude that PET imaging improves physician decision making in each of these

categories?"

We also heard this morning that when NOPR was set up, that a 15 percent change in intended management was -- that anything above that was considered an indication for PET scan. Was that written into the study, and do you agree with that statement that was made earlier today?

DR. HILLNER: In our protocol, we wrote and actively debated the statistical section related to what would be potentially clinically important changes in intended management and how many cases would have to -- data would have to collected to be able to make an inference therein.

15 percent was the most conservative estimate for some members of the working group. Others felt that it should be higher than that, 25 percent, et cetera. The point of that was to make sure that the Registry, for example, one year out, did we have enough individuals in a cell up here to be able to have confidence around the change of intended management?

DR. HENDERSON: Is it reasonable to conclude from what you just said that for each of these cells in which you have a number, that the lower bound of that number is above 15 percent?

DR. HILLNER: For 18 different cancers --

DR. HENDERSON: Well, we're interested in these nine. We don't have to worry about these nine.

DR. HILLNER: We have results for 18 different cancers. And we have for, I believe it's 68 of the 72 potential cells, 200 or more cases. And for all of those, they are above 15 percent.

DR. HENDERSON: Okay. So then that leaves the final question. Well, I have two more questions, but on the direction I'm taking right now.

So now we come back to this. So this means then that for each one of these squares, if you have a number up here, we could reasonably put five down here.

DR. HILLNER: You could put five with the exception of testicular cancer because testicular cancer --

DR. HENDERSON: That's what I said. You don't have any numbers.

DR. HILLNER: -- is so relatively infrequent in the age of Medicare coverage. And that's not a -- that's reality. That's the combination of age and epidemiology --

DR. HENDERSON: What do you mean by --

DR. HILLNER: -- that the role of PET in staging and restaging in disease confined to the brain, it's appropriately being under utilized. So we have no way of making a position statement.

DR. HENDERSON: So how do we deal with, for example, on brain, you have -- restaging, you have blank and suspected recurrence, you have 41 percent. Now, we don't have those two separate categories. We have restaging. And pretty much the way it's presented, restaging and suspected recurrence seem to go together. They seem to be the same thing by and large.

DR. HILLNER: I believe that I would defer to others about the language related to -- is it an

artificial splitting of distinguishing restaging from suspected recurrence. My understanding is is that the language -- the regulatory language should consider those two entities combined to be restaging.

DR. HENDERSON: Okay. So we can come back to that later then. And my last question then has to do with this, everything you've presented seems to apply what we have as question number one, which is really improving physician decision making.

Question number two has to do with patient oriented clinical outcome. Do any of these data apply to that and, if not, do you have any data that applies that helps us determine whether there is an effect on clinical -- patient oriented clinical outcomes?

DR. HILLNER: No.

DR. SATYA-MURTI: Dr. Lichtenfeld, and then Dr. Phurrough wants to make a comment. You go ahead. You've been waiting.

DR. LICHTENFELD: Is this working? Yeah. Okay. It's Lichtenfeld, just for the record.

Anyway, a couple questions and somewhat to

follow up and to help us understand. If you go back to the previous slide, what is the sensitivity and specificity of PET scan in the diagnosis of ovarian cancer and prostate cancer? Do you know? Nobody's addressed the primary -- number to my knowledge here.

DR. HILLNER: I defer to a comment from Dr. Siegel or Dr. Coleman. I don't feel that I can --

DR. LICHTENFELD: So we don't know based on this data how many times a doctor changed their treatment and really made the right decision 'cause we don't have specificity and sensitivity data. Do we, or do we?

UNKNOWN MALE VOICE: On the evidence review we do.

DR. LICHTENFELD: No. No. I know the evidence review shows wide ranges, and they use the word diagnosis. What I want to know is, if a woman has an ovarian mass -- I just simply want to know the answer to the question, if a woman has an ovarian mass and has a PET scan prior to the time of surgery, how many times has that PET scan if it's negative,

correctly negative, and how many times correctly positive for both prostate and for ovarian as examples of the entire survey here?

Does anybody know?

DR. WAHL: Dr. Mankoff did present a slide on that in his talk from one paper, the Risum paper, which in 97 patients, 57 with cancer. This is in Dr. Mankoff's presentation. It did address that. This stated a sensitivity of 100, specificity of 93 percent with a 55 percent prevalence of metastases. So that's one study that was mentioned today.

I would just have to say that the data on sensitivity and specificity also the Antioch (phonetic) data was presented, which looked at a broad range of tumors which, of course, wasn't included because it covered multiple cancers, showing PET/CT to be quite consistently better than CT across a range of cancers in the probably 20 percent more accurate range.

So in general, my read is -- and I think the read of many others -- is that PET/CT performs with an ROC curve that is better, receiver operating

characteristic curve, than CT alone in many cancers. I mean, there are exceptions.

But -- and the data from the Canadian group also placed the -- did some ROC analyses that showed PET pretty far up in the upper left corner, which is indicating pretty good sensitivity and specificity. And one problem with sensitivity and specificity, as they pointed out, is if you dichotomize data, you may lose data on continuous variables when there are grades of intensity. So it's a little fraught with difficulty.

So the ROC data speak to it being, in my estimation, really a very accurate diagnostic method across the board compared to more conventional anatomic methods.

DR. LICHTENFELD: Then follow on. The second part of that question was we talked about ovarian. And obviously, you're giving a fairly high number for sensitivity and specificity. What about for prostate?

DR. WAHL: The Canadian I know just one

paper they quoted was out of Schoder (phonetic) that showed a 35 percent sensitivity which was much lower. And Dr. Coleman maybe has a comment.

DR. LICHTENFELD: And I guess -- oh, I'm sorry.

DR. COLEMAN: There's been a few studies looking at the sensitivity and specificity. It depends on what stage of the disease you're talking about.

DR. LICHTENFELD: I'm talking about primary prostate. Let me clarify. That's a good point. I'm not looking for metastatic disease. I'm talking about the primary diagnosis of a patient with prostate cancer. Two hundred and some thousand men a year. You know, now we're going to do PET scans on them. Okay?

DR. COLEMAN: I don't know of a study that spoke --

DR. LICHTENFELD: 'Cause if I may, just so you know where I'm coming from. That study -- and they didn't put the slide back up. But that study

said that 30 some percent of the men had their treatment changed. Now, we don't know why. We don't know if they avoided surgery because they, quote, "didn't have prostate cancer," if they got the surgery because they may have had metastatic disease. We don't know the answer.

So I'm just curious because it's a big universe out there.

DR. COLEMAN: Yeah. I don't know of any study that's looked at the sensitivity and specificity of FDG-PET imaging in the diagnosis of prostate cancer. So I'm not aware of any data for that.

DR. LICHTENFELD: And I think that that -- I appreciate your honesty.

DR. SIEGEL: Well, I think an important point to your data, there are 300- --

DR. LICHTENFELD: 321 scans.

DR. SIEGEL: 321 patients for prostate cancer. So let's look at that compared to the incidence of prostate cancer in the United States.

DR. LICHTENFELD: Over 200,000.

DR. SIEGEL: There is highly selective use of PET for the diagnosis of prostate cancer for whatever reason that those studies were ordered. We can't tease that out right now. But we know it's not being used for all of the 60,000 patients who were diagnosed with prostate cancer.

Now, it kind of -- and that leads to one of the fundamental issues where -- which is, we know that PET is not perfect for every cancer. We know that it's really good for some. It's pretty darn good for others. And there's a few where it sort of falls down.

The question is, though, should the cancers where it falls down be simply excluded from coverage, or should cancer be covered, and by proper professional guidance, by NCCN guidelines on down the line and education, the word gets out that the use of PET for the diagnosis of prostate cancer is something you only do under these unusual circumstances.

You know, we don't have coverage, non-coverage decisions for CT for the diagnosis of certain

cancers. You're trying to look for microscopic disease and you think that's what's going on. CT is not a good way to find the cancer you're looking for as well.

So I think -- what's always troubled me about this approach is that PET is being denied kind of this level playing field as use as a cancer imaging tool. Is there the potential for abuse? You bet. It's CMS's job to avoid that abuse. Is there potential for misunderstanding and misuse? You bet. It's the professional society's job and the practice to make it clear to people that this is how this stuff is used.

There's all kinds of tests out there. We don't use them in every patient because we know some things work for some things and other things work for other things. And we don't apply them across the board.

DR. LICHTENFELD: If I may, I'm not going to get to my second part of the question, which you've already responded to in a sense because I feel like

I'm heading into a final examination, and I'm trying to figure out what the answers are. And I can't get to the point I need to get to. And the fundamental issue -- you know, I understand what you're saying. But when you're giving that little bit of a discussion, you're probably talking to the folks over here from CMS because they're going to make the ultimate decision in terms of coverage of all cancer versus no cancer.

We're struggling with the question of we have a lot of little cells up there. And those little cells have very specific requests for us to rate the evidence and the quality of the evidence and the indication based on our best assessment of what we've heard today.

And so the conundrum becomes very simply -- and you've articulated it very well. We have a technological assessment that may focus on a particular number of years, may have eliminated other evidence that could have been of help to us in reaching our decision. We don't know that up or down.

But we have that information which, for prostate cancer, for example, says fundamentally we don't know. We can't answer any of these questions.

And then we come up, and we have another study up here from NOPR. And that study says in 30 some percent or 38 percent of the men who have primary prostate cancer -- I assume primary for diagnostic purposes. Maybe not. Maybe I'm wrong. But, you know, in that, 28 percent of the men had their diagnosis changed.

So, you know, there's -- and at the same time in the back of my mind is the same concern you just articulated, which is, having been around this -- pardon me for saying this. But from having been around medicine for many years and involved in regulatory and payment issues for a long time, I appreciate your honesty and I appreciate, you know, the Society is very, very supportive of that.

But I also deal in the real world because I see the other impacts. So that's sort of sitting out there as well. And we're not here to answer that.

But I don't know how we do answer what we have to answer.

DR. JANJAN: I'm sorry. The question then becomes, who and when are those studies going to be conducted to step back and say, we should not do it in this situation? Because the problem is, once it's approved and everybody's using it, do we critically later then do an analysis, and how long is that going to take, and how many patients are going to get scanned?

For example, in stage one breast cancer, it's not -- the consensus is it's not indicated. So how do we get to that point?

DR. COLEMAN: Coleman from Duke. I think it's a lot like what happened with the covered cancers. There were subsequent studies shown where its useful and not useful, for example, in the breast cancer. The early breast cancer, it's not useful.

We didn't know that at the time breast cancer coverage was approved. But subsequent data and guidelines were developed as the technology was out

there and maturing. So I think the same thing will happen with this.

DR. JANJAN: But those studies have to be done is what I'm trying to say.

DR. COLEMAN: Yes. They will be.

DR. SIEGEL: And just to follow that, I think -- I keep coming back to Dr. Phurrough's own words -- not his words, but the words in the guidance, which is that PET is rarely used for the diagnosis of cancer. That the vast majority of time, the way you make the diagnosis of cancer is you see a mass, you get a biopsy.

There are certain unusual circumstances where PET is helpful. And one would hope that physicians are generally responsible and trying to do what's best for their patient most of the time. You know, you have a patient who shows up with a limbic encephalitis, and you can't find the primary tumor, a PET scan that finds a small small lung cancer can make all the difference in how you approach that patient.

DR. SATYA-MURTI: Ms. Tenenbaum, go ahead.

And then I know Dr. Phurrough is waiting. We'll resume where we left off.

MS. TENENBAUM: Cara Tenenbaum, Ovarian Cancer National Alliance. I just wanted to directly address your question. And I do sympathize with your concern over over utilization. But I just called my office, and really quickly so you know, the best study that we have right now is that the sensitivity and specificity for PET scans for ovarian cancer is 92 to 93 percent, and for CT scans just as a contrast, sensitivity is 44 percent and specificity is 71 percent. So it's quite an increase.

DR. LICHTENFELD: Thank you.

DR. PHURROUGH: A comment and then a question. We make national coverage determinations to decide what we're going to pay for. We did not ask Alberta and we're not asking this panel to tell us what we should and shouldn't pay for.

The goal for Alberta was to use a specific methodology, the technology assessment methodology, to assess the evidence around the various indications and

tumor types in which PET scanning is used. And that's what they performed for us.

We didn't ask them to tell us whether we should or should not cover a particular indication or technology. Neither did we ask them to go afield from the typical technology assessment methodology to look at other kinds of data.

So the NOPR kinds of data is not the typical kind of data that would be assessed in a formal technology assessment. Typically that's not used, and so it was appropriate that that kind of information not be used.

Now, in what we're asking the panel to do, we're saying you've got information from the technology assessment. You've got information from NOPR. You have a few other trials that were not part of the technology assessment. You have an older technology assessment that predated the Alberta technology assessment.

And with all that plethora of information -- information may be broader than we want to say. But

with all that information, we want you to say is that enough for me to make a decision around whether in these particular cases and these indications a PET makes a difference. PET changes management, PET changes outcomes.

We're going to take your recommendations, and we're going to look at them, along with NOPR data and along with the TA data and along with a vast -- actually not so vast -- number of comments that we've had and propose what we think we ought to pay for and what we ought not to pay for.

Part of what we are looking for from you -- and you're asking very specific questions to help you answer the questions we asked you to answer. We want you to comment on, can you answer these questions with NOPR-like data? Is NOPR-like data sufficient to answer those questions? Do all the questions have to be answered with Alberta TA kinds of data?

And so that's part of the discussion here. As you arrive at your answers, you are in your answers going to be answering that question. But we want you

to, as you have finished your answers, we want you to be fairly blunt with us and blunt with the public as to whether this kind of data collection is something that the Agency ought to be invested in as we invested our time and the NOPR people and physician community invested their monies in collecting this information.

But it's not a question of what should we do with the information you're going to give us. We're pretty comfortable with deciding how to use the information you give and not always answering it the way that you'd like.

Now, with that, I have a question for Dr. Mankoff. I know he's got to leave in a few minutes. I want to put him on the spot first, and then some others may want to answer this, too.

The recommendation that we heard from you is that we should do away with individual decisions. We should have broad coverage of cancer and cancer indications with perhaps a few exceptions. And we should leave it to guidelines to tell physicians, guide physicians as to how they should practice.

What is your experience and opinion of the guidelines that are out there, how evidence based are they, how much expert opinion is there versus evidence, and how regularly do physicians who are ordering PET scans pay any attention at all to the guidelines?

DR. MANKOFF: So I'm going to take a shot at that, and then I'm going to leave time for my colleagues with additional expertise in that.

I gave a simple answer because it's nice to do that in a presentation. I think guidelines are part of that answer. I might come back to something that Dr. Siegel said, is that ultimately it comes down to educating individual practitioners. And I want to address both of those points.

The guidelines are emerging. There are some guidelines that are getting better. There are some other areas where we don't have guidelines. I've cited the NCCN example because I think many are familiar with those, and those are guidelines that are developed with the available evidence.

And the guidelines are actually fairly honest, I think, about what the quality of evidence is as they've developed those guidelines. They're updated quite frequently. And so if there is additional evidence that can be brought to bear on those kind of guidelines, then the quality of the data has changed to match the nature of the evidence.

I did have an opportunity to participate in one of the trials that was specifically convened for four cancers -- I was on the breast cancer panel -- to be able to try to update those guidelines based upon available data.

And, in fact, as a result of that panel was some of the recommendations on early stage breast cancer when we re-examined the data that had emerged in the meantime. An additional recommendation for future attention paid to areas like locally advance breast cancer, response evaluation, and bone metastases where the data looked very promising.

So I think guidelines is an important part of the answer. I would say those data are not bad --

those guidelines are not bad at the time being. And as the data emerges, we'll have better guidelines.

I think the other portion is coming back to the fact that oncologists and the physicians that are ordering these tests want to take care of their patients in the best way. They are very used to looking at the quality of diagnostic data and making decisions about whether they should get a PET scan and whether or not they should get hair-two (phonetic) testing on their breast cancer patient.

And so I think to the extent that we can provide that education to our referring physicians, I think we will get appropriate utilization of that. And that's one of the ways I've interpreted the NOPR is that the utilization is at least changing management. And we're not getting thousands and thousands of prostate patients coming in for diagnosis because it hasn't been an appropriate indication.

I'm going to stop and see -- and give my colleagues a chance to say.

DR. PHURROUGH: Dr. Shields might be helpful

as the practicing oncologist here.

DR. SHIELDS: You know, I've spent a while working on PET. But my colleagues are new to it. And for a long time they would stop me in the hall and ask me, here's the situation, does this make sense or not.

And they're stopping me less because they've learned how to do it. I mean, they're ordering them regularly. They understand the indications. They understand the limitations. And I think that's changed in the community over the last few years. And I think part of the availability of it has really helped a lot.

You have to understand that the list we're going through right now actually represents a relatively small fraction of the cancers and the patients we would scan. From our estimates, it's only about 10 or 15 percent above what's being done right now. So they really don't represent the bulk of the cancers.

And as you get down to even the other category which represented ten percent of our totals,

those are getting to be real rare, the Merckel cell tumors and my favorite, I treat ampullary cancer. You know, that's never going to have a guideline. You know, you have to look at what's going on nearby, you know, how does the intestinal tumors behave?

You know, it's approved for esophageal cancer. It's approved for GE junction 'cause that's right at the esophagus. But it's not approved for gastric. And really, the histology, the treatments are all the same. We need to sort of look at that biology, and we understand that.

As oncologists, we're used to looking at vague scans. We're used to looking at CT scans with little spots and trying to interpret and know when to use them, how to interpret them, how to act on them. We've done that for years. And now we're learning how to do that with PET.

And people know not to do it in prostate cancer for the most part. I think they understand that. Those numbers tell you that. And quite honestly, I suspect if you drilled down to those 300

cases out of a universe of prostate cancer in the United States, you'd find that those 300 were patients that had a spot in their prostate, hasn't biopsied yet, but they've also got a spot in their lung or something, or they've got a spot someplace else, that somebody's trying to say, okay, where do I go to biopsy, what's going on.

And you'll find that those are complicated cases with more than one thing happening, and that's why they order the scan. It's not somebody with a PSA of 6, and somebody's gone off and ordered, you know, a scan.

And I think you have to take that into mind when you do that. I think micromanaging it to say, well, we're not going to do it for prostate cancer, people already aren't doing it for prostate cancer, you know, in terms of diagnosis. And the ones that are doing, are probably ones that are complicated ones, and they're trying to tease out what's going on.

And most of my colleagues don't order PET every day. They order it when somebody comes in and

has -- you know, I'm treating them for breast, and they've got a spot here. And I don't know if that's real, and I don't know how they're responding. And I need to work out the details. And it's an amazing problem solving tool that we use regularly.

DR. PHURROUGH: Not to disparage physician altruism, but physicians aren't ordering PET scans for prostate cancer 'cause we don't pay for it except in NOPR. Now, if we paid for it in prostate cancer in NOPR, how many prostate cancer PET scans would we have after one years? More than 321.

DR. SIEGEL: That's 321 in two years.

DR. SHIELDS: In two years.

DR. SIEGEL: And you're paying for it.

DR. PHURROUGH: Only in the Registry.

DR. SIEGEL: No, but for physicians to fill out the paperwork --

DR. PHURROUGH: Only in the Registry.

DR. SHIELDS: Yeah.

DR. SIEGEL: -- the way we set the bar is not a tremendous (unintelligible). We have had very,

very few push-backs from referring physicians about participating in the NOPR. So we really believe we're capturing a very, very large sample. We know we've got close to 80 percent or over 80 percent of the PET facilities in the country participating. And we think we're getting a really large practice-based sample of the patients who are eligible.

So I -- I mean, maybe I'm being incredibly naive. But I really believe that we're seeing this tool used as a problem solving tool by people who have a general understanding of how to use PET, and now they have an opportunity to use it where they couldn't use it before.

DR. SATYA-MURTI: Well, very interesting. Thank you. This is a weighty topic. I was looking at USA Today at the left bottom to find a reduction snapshot of this whole topic. Didn't find any. So that would have made our jobs easier.

But CMS is telling us that if you're not confident to leave the -- we have the prerogative or liberty of leaving that block blank. So that would in

itself give you data. Wouldn't it? The not -- exactly sure. If we're not sure, we're just going to leave it blank.

I have a question which may pertain to many of these decisions. In the second table, we're asked to determine clinically significant outcomes. This is important for many number of treatments and tests anyway.

So I used to feel that survival and symptom-free survival were the most important outcomes. But surely there are other outcomes such as not doing a biopsy. So if I may have, within our panel, Dr. Wahl had some comments. And then any of you, very simply, what else would you like to see as an outcome? And then maybe based on today's deliberation, there would be a follow-up using those specific worded outcomes to see how useful it is.

DR. WAHL: One discussion we were having over lunch was really the challenge in diagnostic imaging, when you have a diagnostic imaging test where you have a gold standard such as histology where you

can actually measure the accuracy of the test. And that's been done with PET quite a bit in the last 15 years. So diagnostic studies comparing PET to a gold standard have been done showing that PET is often more accurate than the standard anatomic imaging.

And once one proves that, and if PET's a lot better, sometimes you don't even need really large sample sizes to show that. The first prospective study of PET for staging the mediastinal of non-small cell lung cancer only had something like 24 patients. But it was highly statistically significant because the comparator, CT, wasn't very good. And PET, though imperfect, was quite a lot better.

So to show, you know, effects when there's a very potent improvement in diagnostic accuracy sometimes doesn't take a whole lot of patients. Now we've gone on to reproduce that in meta-analyses and so on in lung cancer, which is very common.

But one of the challenges we face when we try to take a diagnostic test where we know pretty well what its performance characteristics are and put

it into a prospective clinical trial, is you know that you may be randomizing patients to groups who will have an inferior diagnostic method.

And there's only a certain window of time to do that because IRBs -- and in fact this -- won't let you randomize to what you know is a less good method. So this raises the question, you know, in this Fryback and Thornberry, that sort of thing, that level two diagnostic evidence is accuracy, which for a diagnostic test is really the performance characteristic, something critical.

But level three is what do the physicians think about it and what's the change in entropy. You know, those -- I think that those may not be as applicable to diagnostic tests as we would like.

So I think we do have to begin these discussions with accuracy. And I do think accuracy cannot be underestimated. And that's why those ROC things were so important.

But you know, with the accuracy, I think that the shorter indices include, you know, exclusion

of other tests, you know, exclusion of biopsies, switching therapies. I mean, the things that NOPR looked at I think are very important alternative indices to survival because it is -- it becomes very difficult to ethically deny patients good imaging tests to randomize them to look at outcomes.

This has been done early in evaluation. But there's only a certain window of time when one can do it. And as I commented at the Society of Nuclear Medicine meeting, given the efficacy of FDG-PET in lung cancer, I was surprised that the Canadian study was actually allowed to move forward with randomization to the less accurate arm of CT. But that was possible because PET wasn't widely available in that health care system.

So I think some of these intermediate indices have to be looked at very closely because unlike therapies, we have an intermediate endpoint with PET in diagnostic test the accuracy where we really -- as diagnostic tests -- it gets very hard to separate the diagnostic information from the

subsequent therapy and ignore that diagnostic information except for a limited period of years.

And that plus the funding makes it hard to do these prospective studies looking at outcomes the way we can with therapies.

I did want to also say, I wanted to see the other two slides of Dr. Hillner before we finish our discussion because we have to see about treatment changes. I think that was one of our series of boxes on the right.

DR. JANJAN: That's exactly where I was going because it seems to me that the use of PET is sort of weighted front-end, diagnosis and staging. Because when you look at Bruce's data, switch to another therapy overall for the types of cancers evaluated in this meeting, it's about 14, 15 percent switch to another therapy and only about 10 percent -- or 5 to 10 percent go to supportive care.

So really, the use of PET is not -- is not being used toward the progression and treatment. And Bruce, you can maybe expound on this. But it's not

being used late in the course of therapy. It's being really used earlier in the course of diagnosis and treatment, if that's fair.

DR. HILLNER: Before -- I want to take a couple of minutes of background because what's different about the treatment monitoring data is that the currently covered cancers, for example, colorectal, lymphoma, all are part of the treatment monitoring cohort that, with the exception of breast cancer, treatment monitoring -- the use of PET associated with treatment monitoring is not a covered indication.

So the treatment monitoring universe of data includes, outside of breast cancer, that's it. That's the whole world. Okay? And this is truly unique. This is 10,000 cases of treatment monitoring over the first 18 months.

Now, for what I'm to show you are the nine question cancers of interest. But we have all data on lymphoma, colorectal, yada, yada, yada. So what is the spectrum of when clinicians are using this?

Approximately 60 percent of the patients are in active treatment for metastatic disease, 60, 62 percent.

The 38 percent are for local regionally advanced disease.

Approximately 80 percent of the cases are only getting chemotherapy at this time. Approximately 12 percent are getting chemotherapy plus radiation. And the final 8 percent are only getting radiation.

We do not have information in any of the settings about the prior treatments. We can't tell you is this first-line therapy in metastatic disease, second-line. If it was second-line, what the first-line therapy was. We don't know what other prior imaging that they had.

We just know that it is being used during the course of a treatment plan. If it's after two cycles of chemotherapy, four cycles of chemotherapy, six cycles of therapy, we could figure that out from claims analysis. But we do not have that information from the referring physician. So from the get-go, I can't answer those questions.

Again, I find the results strikingly consistent. At the far left is adjustments in the dose or duration of the therapy that they're currently giving. And 30, 30, 32, 24, 30, 28, 22, 28, shockingly, I view as being highly consistent.

Switching to another therapy is the vast majority of the time to another form of chemotherapy. Highly strong pattern of switching to another therapy is more frequent than moving to best supportive care. That is either a reflection of earlier in the process of care or consistent with the American high-intensity pattern of multiple courses of therapy. But it's 14.9, 14.6, 13, 15, 15.

DR. JUHN: Can I ask a question as to in your survey, was it clear? I mean, were the questions as you have them posed here kind of very general questions, or were they very specific as to was it a two-day adjustment in dose, or was it a five week addition adjustment in dose?

I guess I'm getting to the issue of, if you're asking -- again, this is all intention to

treat. I mean, this is a survey.

DR. HILLNER: Correct.

DR. JUHN: It is not the actual practice that actually occurred. So I guess one of the questions I have is, in a survey setting, if you have the survey question in such a way that it's so general, you know, can you actually get down to the more specific --

DR. HILLNER: I think our results -- a point of pride, I think you can legitimately critique the uncertainty around adjustment or the duration of therapy. I think that there's a substantial amount of gray zone in that.

I really think that it would be very surprising that switching therapy did not occur. I think potentially a switch away from the current therapy, possibly instead of the patient refusing therapy, that it went to then subsequently to best supportive care, but the continuation of therapy as is when they're indicating they were going to change. I really believe that an audit of the claims data will

show there is a change.

DR. JUHN: Actually, that raises a very important question that I was going to ask earlier. Are there specific plans to do some type of auditing or some type of follow-on to see what kind of --

DR. PHURROUGH: Yes. We have their database, and we are going to

DR. JUHN: Going to follow up?

DR. PHURROUGH: We're sort of developing the kinds of questions we think will be --

DR. JUHN: I think that will add some credibility to this. I mean, I think it's important. I think the results are very, very good. But I think there is that kind of question mark as to did the reality follow the intention?

DR. SLOAN: Related to that, I'm wondering, is there a possibility of double counting? In other words, could you say I saw such and such on the PET. And therefore, I'm going to adjust my duration of therapy and switch to another therapy. Are you counting -- are you sure that you're not counting that

in both columns?

In other words, one decision goes into two columns. Is that a possibility? Could that be potentially --

DR. HILLNER: I want to look to my data managers. And I believe that's not true. But Barry will comment, please.

DR. SIEGEL: Single best answer.

DR. SLOAN: Single best answer. I wonder if you can give me an example. So if you see a response on PET like response, you probably have a certain finite time you can give a drug. And if you don't, you'd switch to another therapy. So can you give me an example of something that would be that?

DR. SHIELDS: I can give you examples 'cause I use this quite regularly.

DR. SLOAN: Okay.

DR. SHIELDS: And quite honestly, I think when we say adjust dose or duration of therapy, as a medical oncologist, skipping drugs for two days doesn't count. I mean, that's what we do for

Thanksgiving. I mean, that's certainly not a routine issue.

I mean, the way we use it oftentimes is trying to evaluate the magnitude of the response. It used to be when I had one drug for one disease, and it didn't work very well, is patients were treated until they progressed and died. And that was the choice.

Now, the average survival of people with metastatic colon cancer is a couple years. With breast, it's three to four years. And these people are on and off therapy regularly.

And we treat them to best response quite frequently, and then stop and watch and wait. And when they progress, we put them back on. And that's true with ovarian. It's true with breast. It's with all these.

And quite honestly, the CT scans, you know, there are still lumps there. And I'm like, okay, is this active or inactive? And what's the value of a break, and can I stop therapy entirely? Can I adjust the dose? Can I cut back from two drugs to one drug?

That's a daily issue in my practice. And PET helps on that.

And those are real important questions for my patients. When I can say, you're on two toxic drugs now, and I'm going to cut you back to one because your response is so good, or I'm going to cut you off completely, and watch and wait, those are the questions my patients live by because they really want those breaks. And I'm sure Dr. Henderson sees that.

DR. HENDERSON: I think you answered that beautifully. But one other point, too. And that is that after you have treated a cancer for a period of time, you've biopsied the cancer. And, for example, local recurrences in skin after breast cancer is a good example. Those studies were done extensively.

And you often find that you have 99 percent fibrosis. You don't actually -- it's rare that you have 100 percent. But when you're looking at a CT or MRI, you're seeing a mass effect. But that doesn't mean that mass effect is cancer. And this is another potential application.

But I think you've covered a lot of the others. But there's probably some overlap between that first and second column that could go -- you could probably move some of the patients in the first column in to the second and vice versa.

DR. SATYA-MURTI: Any other questions? Dr. Wahl?

DR. WAHL: Maybe I missed it from Dr. Hillner. But out of the whole NOPR, what fraction again of the NOPR registrants were in these nine categories?

DR. HILLNER: Excluding treatment monitoring, in the low teens, 12, 14 percent is our working estimate. Possibly -- right?

DR. WAHL: No. I was just saying, if you look at the nine indications versus the total number of cases. I mean, out of the nine indications, I got the sense there were like 70 or 80 percent of the NOPR fell into this nine.

DR. HILLNER: Oh, I'm sorry. I'm sorry. These 9 cancers -- of the universe of all cancers that

are getting PET scanning is 10, 12, 14 percent. These 9 cancers represent probably 80 percent. Do you have it, Barry?

DR. SIEGEL: So this is the one year cohort -- Dr. Siegel again -- the one year cohort for the JCO paper. So treatment monitoring is not included. The break down was 11.7 percent prostate, 9 percent ovary, 9 percent pancreas, 7 bladder, 7 kidney, small cell was 6 percent, cervix was 2 percent, and brain was 1.2 percent.

What did I forget from the list?

UNKNOWN MALE VOICE: (Unintelligible.)

DR. SIEGEL: Okay. I'm glad someone was doing the addition. Did I forget one that's on the list? I think that's it. Okay. So about half.

DR. SATYA-MURTI: Dr. Henderson?

DR. HENDERSON: I'm still struggling with the -- with actually addressing these questions, filling in these boxes, since that seems to me to be our major exercise here. So I've actually filled out based on the NOPR and the responses from Dr. Hillner.

And then I have the responses, which I've been changing it a little bit. Not a lot, but somewhat during the course of the day, that I filled out before coming here.

DR. SATYA-MURTI: I noticed that.

DR. HENDERSON: Well, I thought that was what we were supposed to do, until we came to a final number.

DR. SATYA-MURTI: Yes. Yes.

DR. HENDERSON: But there's still an awful lot of difference. So if you'd just take question number one -- question number two is even more of a headache. But question number one, just take those boxes. There are 36 boxes there. Okay.

So if I take the NOPR data and I'm making some adjustment, I didn't make them all 5's. I just sort of said, okay, well let's make a 4 if it's under 30 and a 5 if it's above. And 1 if it's blank.

So it works out that the -- from the NOPR data I have the 1's and 2's equal 6 out of 36. Well, when I look at the tech assessment I come to 26 out of

36 are either 1's or 2's. In other words, that's where I come out after reading all of that. It means that I have no confidence, little confidence based on that data set that would have had an impact on therapy.

And then if I look at the 4 and 5's, there are 30 out of 36 using the NOPR data, but only 6 out of 36 that I feel either moderate confidence or high confidence.

Now, I've read through this. I've read sections of it several times now. And you know, maybe Dr. Gulenchyn wants to address this.

You referred earlier to conclusions. But I can't really find conclusions in there. And I find in the end, what I have to do is sort of read over. I haven't been able to find any criteria that I could apply systematically. I tried as I read through this. And I went back and forth. I didn't put them on the sheet here first. I just wrote them in the book because I figured, well, I'll go back and try to be consistent.

But ultimately, for this question, I find that the only thing that really applies, particularly if I use the NOPR kind of thing, is that last segment, which is just text, a textural description, and then you have one kind of a table that sort of summarizes the treatment decisions in almost every section, but no real statistics, no real cut points.

So I mean, is this -- would you come up with only 6 out of 36 that you would say would be in the four or fives if you were to fill this out? Have I interpreted this the way you meant it to be interpreted, or did I miss it?

You clearly stated what your conclusions were, and I missed them. Is that a problem? 'Cause you did refer to conclusions earlier in one of your answers to questions. But I just can't find the conclusions in there that allow me to fill out this table or to check whether I agree with your conclusions.

DR. GULENCHYN: The technology assessment that you have in hand at the moment, the bundle,

rather large pile of paper that was published and that you were provided with has a blank section in it under each one of them for discussion of the data and some conclusions if we are able to draw them.

That has -- the paper that you have at the current time is not a complete paper. The time period that we were given for a complete technology assessment was compressed. And we took it as far as we were able to in terms of doing a data summary and a data critique.

We have not drawn conclusions at this point in time beyond any that I perhaps expressed when I was doing the presentation. So there is data there on sensitivity and specificity. There is data there -- there is ROC data. And there's the likelihood ratios that were presented in the data this morning. And those all pertained to question one.

On questions two, three, and four, four as I said, was easy. That's the pancreas. And, yes. There is some limited cost effectiveness data for the pancreas.

On question two and question three which is how do these technologies impact on management, the majority of data is there on Q2. And for each of the nine cancers that we were asked to address, certainly the management changes that are being quoted are in large part, in fact, in line with where NOPR has come in with data indicating management changes.

The level of confidence in those management changes is reflected by the grade that the papers have been given in terms of the A, B, C, D grade. And that will be reflected in a full -- a fuller conclusion under each of the forms of -- under each of the forms of cancer. But it's not complete at this point in time.

DR. HENDERSON: So you're eyeballing it right now, which is what we've had to do, that if you take these 36 boxes, that you'll have 6 of them with 4 or 5 or quite a few of them with that when you reach your conclusion.

DR. GULENCHYN: Maria's going to go now.

MS. OSPINA: Sorry. I guess when we

presented this point at our technology reports, our facts, we are -- we haven't provided yet any discussion -- any interpretation of the data. And I guess that's what you will have to do with the facts and the data that we provided.

DR. JUHN: I have a related question.

DR. SATYA-MURTI: Go ahead Dr. Juhn. CMS format -- and please correct me -- is such that after you are done with the question, there should be within panel discussion. So we'll move onto that after your question.

DR. JUHN: In your preparation for the technology assessment, did you also look outside North America to see if anyone else had done technology assessments in this -- in this area? And, if so, what were those findings? I mean just in general terms.

MS. OSPINA: It's something that -- sorry, we forgot -- in the slides -- if you look at the set of the slides, there is a summary of the findings. That is the best that we can do for today's meeting, trying to provide a summary of the evidence. But

there's a lot -- too many other things that we still need to do for those final sections that -- I guess the final deadline for this report is September 15th.

But part of the things that we are planning to consider is to put our findings in context with other reviews and other reports that have been produced including, of course, the Duke report, but also other meta-analysis that have been published in the medical literature.

DR. JUHN: Right. So can you comment on any -- has NICE looked at this or has any of the other --

DR. GULENCHYN: Yeah. There was a slide in the presentation that we gave that listed the technology assessments that we have considered. And certainly, we have considered ISIS out of Ontario, NICE, the Quebec group, several other American, the VA and the Duke TA. There's one out of Australia and one out of Scotland as well. So, yes. It has been looked at internationally.

DR. JUHN: But if you looked -- I mean, again, just very qualitatively at a macro level, are

your findings consistent with what many of these other HDAs have found?

DR. GULENCHYN: In terms of the quality of the data, yes.

DR. JUHN: And also kind of directionally in terms of --

DR. GULENCHYN: And directionally. Yes.

MS. OSPINA: I guess what is interesting -- and this is something that agrees with what the Duke report states -- is most of the studies have been addressing the -- how useful PET is for staging and recurrences. Those two areas seems to be the most widely studied in the medical literature.

DR. JUHN: Okay. Good.

DR. SLOAN: One quick question.

DR. SATYA-MURTI: Your last question. Were you waiting to make a point?

MS. PATHEL: I was waiting to --

DR. SATYA-MURTI: Please go ahead. We should go on to within panel discussions.

MS. PATHEL: My name is Carrie Pathel

(phonetic). I am a registered nurse and have been working within the PET community for about the last ten years.

A couple things is, I've been at multiple of these CMS Advisory Committee meetings. And it appears that sometimes we're getting side-barred by small issues with diagnosis versus initial staging versus restaging.

As far as the technical assessment versus the NOPR data, the NOPR exists because we all knew that there was not technical data sufficient to make the decision based on solely. So I ask that -- you know, I just want to keep that in the forefront of our mind, that the NOPR exists. Medicare sponsored it. And we all in the community have educated our physicians, our clinicians to the point that we could obtain data, real data from real patients to help make this decision.

So I hope I'm not preaching. But again, you know, those little instances make a big difference.

DR. SATYA-MURTI: Thank you very much. The

very last one.

DR. SLOAN: Yes. Dr. Hillner, you just made the point about the data being consistent with all the nine cancers. And the slides sort of demonstrate that. But your own conclusion, you left out three different cancers. You didn't -- you asked for approval for brain, cervix, ovary, pancreas, small cell. You didn't ask for bladder, kidney, and prostate, for example. I mean, is there something that -- are you sort of --

MR. CAPLAN: (Unintelligible.)

DR. SLOAN: Sure.

MR. CAPLAN: The request was for six cancers.

DR. SLOAN: Right.

MR. CAPLAN: And CMS has decided to address the additional three, which would have been the top nine in the registry.

DR. SLOAN: Okay.

MR. CAPLAN: Okay.

DR. SIEGEL: I just want to make sure this

is clear. The petition, if that's the correct term --

PANEL MEMBER: Right.

DR. SIEGEL: -- that was submitted by the NOPR working group asks for coverage for diagnosis, staging, restaging, and suspected recurrence of all cancers, not six, not nine.

The six came from the title of the extant NCD. And the three were added for the reason we just heard, to what was done in the technology assessment. But our request stands as all cancers. And the recommendation from the intersocietal representative was for all cancers.

DR. SATYA-MURTI: Thank you very much. Panel, I think this time period is intended for discussion within the panel, maybe some type of summarization. If so, I'll go down left to right. And my own --

DR. LICHTENFELD: Can I ask a methodology question?

DR. SATYA-MURTI: The very last one?

DR. LICHTENFELD: Well, yeah. Dr.

Phurrough, I just want to make sure I didn't hear you say -- how are we to consider the NOPR data in our discussion and deliberation?

DR. PHURROUGH: That's up to you.

DR. LICHTENFELD: Okay.

DR. PHURROUGH: And what we'd like to hear from you is whether you think the data is sufficient to answer these questions. It should or should not be used to answer the questions as part of the discussion here.

DR. SATYA-MURTI: I enjoyed these deliberations. Science on one hand and clinician's dilemma. I'm a clinician myself.

The two outstanding things -- aspects that occurred to me is that some of these tests -- the index test was assessed without blinding to the -- to the comparator. So I think your -- a TA brought this up. And that would be a little difficult for me because many a time, to interpret the index list in void is difficult.

But nonetheless, during actual initial

studies, it needs to be done that way. I think it's one of the earliest tenants of EBM for diagnostic tests is to separate the two.

And I know that TA also narrowed the interval because with a diagnosis like a malignancy, if there is too much of a separation, one test might come out better than the other if there is more than, say, three months. So that is a bit of an obstacle why total blinding was not possible.

On the Registry -- again, we brought this up this morning -- when the registrant or the participant is aware that the future of continued wide-spread application of a technology is dependent on the pre- and post-test answers, even unintentionally it is very hard to maintain neutrality or equipoise to know that something is going to follow which hinges on coverage if management was deferred.

I know most -- all of the clinicians will not do it intentionally. But nonetheless, I wonder how much of an operative influence that had.

So those were the only two. It's not

questions, but those were my concerns.

DR. JUHN: Can I just raise a question about the grids that we're filling out? I think the first one is difficult.

DR. SATYA-MURTI: Can we go this way and then come back, if you don't mind? Thank you. Thank you, Dr. Juhn.

DR. HENDERSON: So as I look at it, where we stand right now, if I look at the -- the NOPR analysis is easy to understand. It's very straightforward and very attractive. And so it would be -- and they've very nicely filled in the boxes for us for question number one, which makes it very attractive.

But I must say, I can't think of many situations where using just the tendency of physicians to decide to do something has ever been sufficient. I've been on a lot of tech panels, mostly FDA for Blue Cross/Blue Shield. And that decision by itself has never been sufficient to say this is something that is worthwhile and gives me confidence that the patient is well served, which isn't exactly the way the question

is answered, number one.

But that is the way number two is addressed really. So I'm not comfortable with just saying, let's take anything above 15 percent of the lower bounds or let's take everything of 30 percent and above and say, well, 30 percent of the patients have decision change, that's sufficient.

And likewise, we have this question, can we apply this to all tumors? And again, there is consistency of the data, the NOPR data, and I would say even in most of the data as is presented in the tech assessment as well.

But we do have -- there are some sections in here, clearly the avidity for FDG does vary somewhat from one tumor to another. And there's also some practical questions and issues related to excretion and how that may make assessment of some tumors, particularly in the areas of kidneys and bladder particularly difficult.

And we also know that breast was certainly an example where it was concluded that there -- I

think probably because they focused on only one tumor, I haven't actually read over that. But it looks to me like that in itself is evidence that we can't just check yes for all of these boxes based on the data we have right now.

It does seem to me, however, that if we were to apply the NOPR criteria, that is, 15 or 30 percent, either one of those numbers, of the physicians change their mind as a result, in fact, using the tech assessment and using the NOPR data, we would actually fill out the boxes of question number one and probably number two the same. 'Cause in number two, if we look at these criteria for patient outcomes, percentage of patients improved with test or without test, morbidity avoided after having image information, change in quality, adjusted life, and so on, if we use those criteria, neither group really presented very much data for any of these tumors. At most, we have an anecdote here and there where there's a study that may provide some compelling evidence that -- that we've affected patient outcomes.

But unfortunately, if we begin to add this business of avoiding toxic treatment, I think that's been pretty much covered in the two presentations under number one, not as a separate spinout, if you will, under number two. And so that wasn't one of the things that I used in my initial criteria for number two. I'd have to go back and redo it.

I think we don't have, you know, a commonality in our definitions of how we're going to fill these out. And I have focused, I guess through my training with other type of panels, I've focused on trying to address the questions that the sponsor, in this case CMS, asks rather than coverage. So that's why I keep coming back and harping on the questions.

And I'm still -- on questions three, four, and five, I feel somewhat comfortable. But on questions number one and two, I'm still having a lot of discomfort here.

DR. JANJAN: I guess the thing that strikes me most is the NOPR data that's actual in-practice data. And it reflects a snapshot of clinical practice

in the United States.

The key thing though -- you know, they always say, don't order a study unless you plan to act on it. And so the key piece is what did you do with the study of the nation? And I think I would strongly recommend from my perspective getting the next piece of information from Dr. Hillner. And that's coordinating outcomes with this data because unless the outcomes reflect something important from the data, then you're really not answering the question.

You can try and sort out sensitivity and specificity in that did you change something? But unless it's making some impact on patient care, then, you know, you wonder, well, that will help, I think, better define the indications for PET. And I guess the piece -- as I say, the piece that I think really is necessary as the next step is getting the patient outcome data to correlate with this information that we've been presented today, which I think is a very important snapshot of clinical care.

DR. SATYA-MURTI: Dr. Lichtenfeld?

DR. LICHTENFELD: Thank you.

I sort of voiced my concerns earlier. I think Dr. Henderson actually is going along the same path.

You know, we were asked to do an exercise before we came here. And that was to fill in the blanks as to what we thought. And then to come here today, listen to the testimony, and then see if that changed our opinion. I, you know, still am waiting to take my final exam. I'm in the process of taking it. And I'm going to admit that I have not the degree of confidence that I'd like to have in giving the answers. And the reason I don't have the confidence, I think, has been already articulated.

If you held to a -- you know, I've seen, and all of us have seen -- Dr. Henderson has alluded to it -- we're trying to move into an evidence-based world in terms of how we practice our medicine. But the quality of evidence isn't what we would like to have.

So we have to rely on surrogates to give us the answer. We hope that more often than not that the

surrogates get us to where we want to go correctly. Unfortunately, we found out, without going into details, that the surrogates aren't always perfect. And that's where the basic dilemma is for me.

So I will say that I probably have moved up the number ladder simply because of the NOPR data. And I believe the NOPR data does provide some additional evidence in this. Is it compelling and conclusive? The answer is no for a variety of reasons.

Nobody's fault. Nobody's saying it was, you know, wrong or ill-intent. But it's not to the statistically compelling evidence level that we'd like to have from the types of studies that were analyzed by the folks who did the assessment.

If it sounds like I'm equivocating, I am. I wish I could hand the sheet in blank, but I won't. But I don't -- whatever I -- whatever I would put down here for many of these boxes, I can't say if you came back to me and said what's my personal confidence interval, the answer is, it would be pretty wide. And

I don't think I have -- can make -- make the adjustment -- make the -- necessarily be as precise as I'd like to be.

Having said that, I think it's also important to bear in mind that this is a technology, at least in the experience in the world that I live in and many others here live in, has been held to an entirely different standard than anything I've seen before.

MRI -- CT scan is the one that I like to talk about a great deal. A CT scan here in Baltimore, we had to beg, borrow, and steal to get a CT scan back in 1977, when I started my practice. We didn't have it, by the way, before I started practice. I was in training here. We didn't have a CT scan. We had a couple. And we had to get consults, we had to get approvals. Well, you know, it was a natural selection process.

And then eventually someone said, you know this thing really does work, it's better than the chest X-ray. It really does show us more. That's the

way -- that's the way it developed. We didn't go through this process.

And now, we've sort of come full-circle. And we're requiring a standard that may be -- may be not achievable for all the valid and legitimate reasons that have been cited, i.e., this isn't the way that technology is introduced to our medical care system. The standards are entirely different for technology such as this technology as they might be for the use of a new -- a new drug that comes to market. So there were entirely different standards.

And here we are as an expert panel trying to make some sense out of this. And by the way, we still don't even have the conclusions from the technical assessment. So you can imagine not only, you know, was I indecisive, I mean, I crossed out 15 different times, and I put different numbers in. I was so thrilled to see that my fellow panel members had done the exact same thing when they came here.

So I know that that really gets to the issue. But I will say that, you know, we're caught --

as I said in the very first comment I made today. We're on the horns of a dilemma in terms of how you apply the standards. We have to make a decision. And I just wish that at some level I was more comfortable with the decisions I had to make.

DR. SLOAN: I agree with everything that's just been said. I mean, I think the methodology of the technology assessment is extremely robust. But I think very few things that we do in medicine would meet that standard. And I don't think it's realistic to expect everything we do to meet that standard.

I think it's been stated a lot of times in the last decade since a human-genome project that the standard -- you know, the future of medicine and the goal, the ultimate goal, I guess you should say, is the concept of personalized medicine.

And people talk about genomics, proteomics, epigenomics as a way to go forward with that. But that's decades and billions of dollars away. And I think one of the most important things that we can do is to try to figure out what works and what doesn't.

And I think PET as a functional study may be the ideal -- one way to do that. So I think the NOPR data is very interesting and exciting in that way.

On the other hand -- and it's also true that the most expensive thing you can do is to give someone a treatment that -- both in terms of personal expense and finances to give someone a treatment that doesn't really work 'cause it doesn't work, and they still have the side effects.

That said, the data is very subjective. And it would be very helpful to get some of that -- that additional data from CMS. Because right now, we're really being asked to interpolate between two different types of data and classifications. And I don't know that anyone can really do that. We'll certainly try as best we can, but it's very hard to be objective.

DR. BERGTHOLD: I want to say something positive about coverage with evidence development. And that is that I think it is a real step forward for CMS to be doing this and to be trying it. And if the

methodology wasn't perfect or the questions weren't perfect and, obviously, we're getting outcome in terms of data that is really hard to interpret, that shouldn't, in my view, be in any way an impediment to doing this better in the future.

To say that because we haven't done it before, it's not fair to do it now seems kind of crazy to me. In fact, I think we -- it's too bad we didn't do it before. Maybe we should reopen CT and MRIs. But since we're not going to do that, going forward, this is the future way to deal with Medicare costs and Medicare quality.

And I'm sorry for the people who do PET scans that they -- it has sort of fallen on them to be the first ones. But I do think we need to be continuing to do this.

I -- I have -- I had a very difficult time as well in evaluating the evidence and thought it was very -- a very complicated thing to do. But I was glad to see some evidence for doing this because we don't make decisions. It's not up to us on this panel

to make the decision. In fact, you know, we should remember that. We are evaluating the evidence that's before us, perfect or imperfect. That's our job.

And so our answers will reveal whether or not we think it's, you know, more or less perfect.

Just one final thing, when we talk about sort of physician practice and guidelines, you only really need to look at the Courage Study to know -- and I don't think there's that much difference between a cardiologist and oncologist and how they apply guidelines -- to know that you could have fantastic guidelines, and it will have little or no impact on practice.

So I think we have to have coverage standards and better evidence along with guidelines. I don't have a lot of trust in those kinds of things to change physician behavior.

DR. JUHN: So first of all before I get into a couple of my comments, I just wanted to clarify. Of the two grids we have, I think the first one is about improving physician decision making, and the second

one is about patient outcomes. Is that correct?

So in our evaluation of the literature, it really is to determine whether the evidence, the literature evidence or some of the other studies that have been presented really do have an impact on patient outcomes. I think the challenge is, there were very few of those -- of those studies.

DR. PHURROUGH: There are probably two ways to look at -- at question two. One, is there direct evidence that says the use of a PET scan results in a particular outcome? The second one being, are we comfortable that -- assuming that there's evidence that there's a change in patient management, are we comfortable that that is an adequate surrogate, and there's additional evidence that would demonstrate that a change in patient outcome -- patient management would result in a change in outcome?

So you could choose either one of those tactics. You would need to be comfortable in explaining your --

DR. JUHN: I'm usually a purist, so I'll go

with the former.

So just a couple comments here. One is, you know, I've asked a few questions about the quality of the evidence. And I know that we've had a little back and forth about, you know, different standards for evidence. I think this is a -- I think an issue that through the, you know, kind of obviously very important for this particular technology. But I think as Linda talked about, this really has much broader ramifications. And I think if we really are going down the path of evidence-base medicine, we have to have a very consistent set of principles and standards that we'll be following.

And if we start identifying certain technologies or certain diagnoses where we're going to have a different standard, I think that's going to create, at minimum, confusion. And at worst, I think it kind of gets us kind of going backwards as opposed to going forwards.

So I think there's something to be said for are we going to establish a set of evidence, a set of

principles by which we're going to evaluate the evidence? And if so, good or bad, what does that evidence say?

And then this question of well, if it's not being funded adequately, I think that's a -- that's a separate discussion. I mean, that's a separate discussion about how do we convince those who can fund these studies that there's value in funding these studies? But to say that we have kind of poorly conducted studies or not well -- well designed studies because there really isn't funds, I think that's a very, very challenging -- challenging argument.

The second thing I wanted to comment on is the NOPR database. I think coverage of evidence development, I think, is a very important conceptual approach. I think -- I think it was mostly gotten right with NOPR. I think we got quite a ways down the line. I think the challenge, though, is the data that we've been presented is early, and it's suggestive, it's not definitive. And I think we really need to go -- and I was glad to hear Steve, you say you're

actually planning the follow-up studies to really link real patient outcomes with the intentions that were collected in the -- in the survey data.

So again, I think as a number of folks have said here, we'll comment on the evidence. I think the decisions are really going to be not only CMS's decisions, but it's also -- I think that the broader community to again define what is it that we mean by evidence-base medicine. Are we going to hold ourselves to it, or are we not?

DR. WAHL: Well, I've been working with PET in humans for about 20 years, so I have a bias coming into this based on experience in a fair number of cases. And I do think one of the things that I brought up early on that I worry about in a strict evidence-base analysis is not looking -- not having a good way to incorporate the biological underpinnings of the process.

Because a lot of work has been done looking at the molecular characteristics of glucose utilization in tumors, and it's a rather consistent

alteration. It's a survival advantage that cancers can live in conditions of low glucose -- excuse me -- low oxygen through the use of glucose. And this is a huge advantage, and it's almost a general phenomenon and it was described in the thirties.

So ignoring that and treating each of these as independent entities I think ignores everything -- you know, all my basic science classes in medical school and every -- you know, it just becomes a statistical exercise. And it doesn't incorporate the biology.

And I do think we have to -- in my mind, the behavior of most tumors is quite comparable. And this underlying biology gives me more confidence that if I apply it to a less common cancer, that it's going to work. And I think that's been borne out by the NOPR data.

I feel since we have as a, you know, country moved to make PET available in some of the more common cancers, like lung and colon, lymphoma, and these are being used very widely, denying it to patients who

have the misfortune of having an infrequent tumor is to me tragic because it's just very difficult to meet the evidence thresholds.

Having -- having -- you know, you can say multi-center studies and so on, but we see in the NOPR data, you know, under 200 testicular cancers in two years. You're just never going to get there. So we eventually have to make a leap that the combination of biology and evidence is sufficient to extend our views, be they positive or negative, to other cancers.

The other thing I'd like to say -- and you can probably tell by my leaning a bit and since I do PET and get paid for it -- that I think that it's generally quite valuable. Though the evidence is remarkably unsatisfying sometimes, and having been in the position of putting together prospective multi-center trials, I should at least tell you one anecdote.

When you first develop an imaging test, the IRBs say, well, you can't use the imaging results to make decisions because it's an experimental test.

Then once you prove it's better, you're quickly told that you can't compare it to the other test because now you know it's better. This is where you have the accuracy data.

And this is one of the challenges we face. There's only a certain window of opportunity to do these prospective trials. I think we're a little bit past that point now because we know that the positive likelihood ratios in all these tumors is so high, the positive scan really means something.

I'd like to just come back to one other thing. These patients are getting imaging now. They're getting CT scans. They're getting MRI scans. And by and large, although the analysis didn't look at it specifically, but it certainly referred to the performance of PET is typically better than that of CT -- not always.

So one thing we have to ask for getting image with an inferior test, what happens if you image with an inferior test? You have management decisions that are based on incorrect data. Those management

decisions can be rather bad excursions and expenditures of resources due to the bad diagnostic information. So although not strictly studied, the mini study, the best diagnostic information, is actually the cheapest because the therapies are more appropriately applied.

So those underlie some of my thoughts. And I've been changing my numbers as well. And I'll stop there.

DR. SATYA-MURTI: Dr. Phurrough?

DR. PHURROUGH: I'm not going to comment on what I think about the evidence. You all get to see mine in about three months, I guess. You'll get to see the Agency's. Sometimes those are mine, and sometimes not.

For those of you who are concerned that CMS has selected PET as the entity that we're going to restrict and let everything else run wild, you probably would like to talk to the cardiovascular interventionist who would like to do carotid stenting where we, in fact, don't cover the FDA-approved

indications for carotid stinting as we think FDA was a bit lax -- I'm sorry -- a bit liberal in their indications.

I'd also like to comment just briefly on our sort of our decision-making process, which attempts to say that we're looking for sufficient, adequate information that a particular technology improves outcomes. And if we can answer that question yes, and we find it to be reasonable and necessary, reasonable and necessary is the CMS mantra as safe and effective is for FDA. And we cover it. Coverage can be broad. It can be narrow. There can be restrictions. CED is one of those restrictions.

It is extremely common for us to find that there isn't sufficient data to answer the question as to whether it improves health outcomes or not. And unfortunately, that is typically taken as a disparaging comment on the technology when, in fact, it's a comment on the amount of evidence for that technology.

Unfortunately, that's somewhat semantic in

that it ends up resulting -- if we follow our evidence-based principles -- in non-coverage. Whether the technology works or not, if there's not evidence to show that it works, we're not going to pay for it doesn't help those that are not getting the technology because it follows some separate principles.

But it is clearly a different question -- a different answer than if there's sufficient data to show that it doesn't work. And that's always a much more challenging issue for us. If there's sufficient data that it works or sufficient data that it doesn't work is certainly an easier coverage decision than one in which we don't have a sufficient amount of data to know one way or the other, which is why we have these meetings so that we can get the expert opinion that, in fact, Congress asks us to get on a regular basis around these issues.

One final comment on outcomes data. Just as the NOPR registry is a -- is an experiment in and of itself in the ability to collect data in this manner and use it for the purposes of making coverage

decisions, the use of outcomes data falls into that same category. We have begun in some isolated instances to attempt to marry outcomes claims data with other data that we've collected. As many of you may recall back in 2004, sometime previously, we required registries for implantable defibrillators, ICDs, and were collecting data at the point of implantation around patients to look at whether appropriate criteria were being followed.

Well, we're now attempting to track those patients through our claims databases to see what's their survival. Does a survival in our claims database match those in the trials? Are there increased number of procedures? Are there more MIs? Are there more sudden deaths? What's sort of the outcomes that we can track in our claims database?

And we can track those outcomes fairly well. We pretty much know when our patients live and die. Not always, we occasionally pay claims on dead people, but not very often.

So this is another one of those where we

want to take some fairly rigorously-collected data around what happened at a point in time and then track those patients to see what occurred. There actually is some improvement in the data we can track now that we are able to track part D data. Challenging even for those who work here, your part D data, but that's a technical issue we're working on. But we know what chemotherapy people are getting because we're paying for it. So we can know whether the doses were increased or not increased or when the drugs were changed or not changed. So we're hoping that we can use this to sort of marry these two kinds of data.

This still comes down to the methodological question of, is that sufficient? If it's not -- if it's not the purest randomized trial prospectively collected data, does even marrying those two pieces of data make them sufficient to answer the questions? Not just what we're trying to answer, but the questions that patients and physicians are trying to answer as they see patients.

I'm not asking you to necessarily answer

those questions today. But I do want to just focus on, this is an ongoing maturation process of doing evidence-base coverage decisions to learn what it is that you can use and should use in making evidentiary decisions.

And the questions aren't all answered. We're going to answer them as they go along. And new technologies are going to, in our view, benefit from this new environment, in your view, may be guinea pigs. It depends on your point of view. But as new technologies come along, we are going to apply what we have learned in the methodological world around how we can evaluate these and not evaluate these.

And just as a final comment, regardless of the decisions we make around this, we think this has been -- CMS thinks this has been a tremendous cooperative work with the folks who have been involved in collecting the data in the NOPR registry. It is a tremendous amount of work on their part. And we've been -- we've been, we think, privileged to be involved in this -- in this work. Recognizing from

the beginning that we weren't real clear what we'd get at the end and how well we would use it, but we certainly think this has been a superb effort and a significant advancement in how we do collect data in our patient population.

DR. JANJAN: Steve, if I might?

I think the issue you make about what happens between a clinical trial and general practice is an important one because in phase four -- you know, in a clinical trial, you've got very strict eligibility criteria. Once you get it out into general practice, you don't have those same strict eligibility criteria.

And what may be true in a clinical trial in terms of results may not always be reproduced in the phase four experience. And I think the NOPR data is important from that respect as well because you're getting a more generalized outcomes data than you would from a clinical trial.

And as we try and compare the data from clinical trials and apply that to NOPR and some of the

other things that have been presented here today, I think is an important step because the phase four data, I think, is actually much more important to refine the indications for treatment.

DR. SATYA-MURTI: The fact of having older tests slipped under without the EBM knife, I am sympathetic and empathetic. There are many in neurology -- many tests, such as the EEG, which wouldn't qualify for use with current technology.

And somebody was talking about CTs. We had ten CTs assigned in a large Navy base during my two year draft stint. And, yes. A retired line officer got that always. So that you would relate to that. Okay.

So nonetheless, I think as part of our both evolution and enlightenment, it is unfortunate we are subjected to this. There is one obscure neurologic tests called McNeeder (phonetic) encephalogram, MEG, that is having similar difficulties now. So I -- I suppose it's part of evolution in medical thinking is how I need to look at it.

We could move on to the voting question, unless we have really a very tumescent comment we need to release.

Do I read the questions, as is the custom, or --

DR. PHURROUGH: I think --

DR. SATYA-MURTI: You go ahead.

DR. PHURROUGH: I think for these first two questions what we want you to do is just we'll take a few minutes for you to finish your charts. We'll collect your charts, and then we'll let you comment on your charts -- on your -- why you filled out the charts the way you did while our recorder here is trying to get them computerized.

MS. ELLIS: Excuse me. Panel members, there are new blank questions inside your folder, if your name is already on them.

DR. SATYA-MURTI: And then go back to the voting process. I am not a voting member. So we'll start with question three. Start with three. How confident are you that these conclusions are

generalized to other cancers? We have this voting flag.

DR. BERGTHOLD: Before we do it, I think we need to -- Peter, why don't you explain?

DR. JUHN: Yeah. Just for clarification. So this question is about the conclusions of the studies of the nine cancers, whether those study conclusions from the nine cancers can be generalized to cancers that were not part of that original nine. Is that the intent of the question?

MR. SHIELDS: That was my interpretation of the question.

DR. JACQUES: Well, what we were -- Louis Jacques. What we were thinking at the time was whatever personal conclusions you made in the table about the sufficiency of the evidence, did you, based on however you determined to fill out that table, do you feel that those conclusions are generally applicable?

Now, if your table is entirely heterogenous, then I suppose the conclusion you would say is that

the generalizability of this is similarly limited. And there may be ups and there may be downs.

On the other hand, if your roundtable was largely positive or largely negative, you would essentially be saying yes. And I believe the same thing about the others.

DR. BERGTHOLD: Okay. That helps.

DR. PHURROUGH: And specifically outside the nine cancers to other cancers.

DR. BERGTHOLD: Got it.

DR. SATYA-MURTI: We have a voting system that the Olympics have borrowed from us. So maybe hold up the numbers.

DR. PHURROUGH: Hang on a minute. Hang on a minute. Hang on a minute. Our recorder is otherwise occupied. Someone's got to record over here. Let her finish what she's doing. Hold up your numbers again.

PANEL RESULTS: (Wahl, five; Juhn, one; Bergthold, two; Sloan, two; Lichtenfeld, three; Janjan, four; Henderson, three.)

DR. SATYA-MURTI: I believe one of the

things I'm tasked with is to find out why you chose what you chose. Dr. Henderson, starting with you, why did you go for three? Very briefly.

DR. HENDERSON: Well, actually, in my earlier marks, I think I made that clear. Number one is that the difference, for example, between breast cancer and these other cancers has led me to believe that probably we can't just routinely and knee-jerk say, without looking at some data, that these cancers are all exactly the same.

And then I commented also on the avidity for, you know, for the glucose with the FDG. And then the location of the tumor could make a difference also, also related to the FDG and interference. So those are the three reasons why I thought you have to be thoughtful about each one.

DR. JANJAN: I selected four because of the homogeneity of the NOPR data.

DR. LICHTENFELD: I selected three because of all the reasons I mentioned earlier. I felt it was -- my table was all over the place, so I didn't think

it was fair to apply it to all other cancers.

DR. SLOAN: I selected two more or less for the same reason, the NOPR data was fairly consistent. The previous data was very inconsistent. So I sort of went right between. And my personal conclusions were somewhat varied by cancer.

DR. BERGTHOLD: I thought the evidence showed that it is not as effective, for example, for brain as it is for cervical. So I selected two.

DR. JUHN: I selected number one because I had no confidence that you could extrapolate from one cancer to other cancers when there is such heterogeneity of responses from the nine that we reviewed.

DR. WAHL: And I was an optimistic outlier at five. And the reasons for that are one, the NOPR data were very consistent. Two, having practiced PET, the generalizability I saw is that many of these cancers -- most cancers have high uptake of FDG, and that the detectability issues are typically related to the size of the lesion and the location of the lesion

near FDG excretion sites.

So like in the brain, the brain has a lot of glucose uptake. And you may not detect lesions as well because of that -- because of the background, similarly in the bladder and prostate. So I thought generalizable -- certainly not fully generalizable. But since it didn't say fully generalizable, I had high confidence that the biology carried and that the usual detection issues were related to location and background and lesion size.

DR. SATYA-MURTI: Okay. Moving on to four. How confident are you that these conclusions are generalizable to non-research PET facilities in the general community?

PANEL RESULTS: (Wahl, five; Juhn, four; Bergthold, four; Sloan, three; Lichtenfeld, three; Janjan, three; Henderson, four.)

DR. SATYA-MURTI: Go ahead.

DR. HENDERSON: Well, my reason for the --

DR. SATYA-MURTI: Go through the process of explicatory --

DR. HENDERSON: Well, I suppose in a way it's a lack of any suggestion that there's a lot of variability. There's a lot of PET scan centers now, and they seem to be of high quality. Although we didn't address that directly, I thought that the tech assessment did to some extent. The scientists and radiologist here did to some extent.

And I haven't seen anything to suggest that this is something that can't be done by -- other than a specialized PET unit. But it doesn't have to be a researcher or university-based PET unit.

DR. JANJAN: I selected three I guess because I don't think there's enough phase four data out there to say that four would be appropriate.

DR. LICHTENFELD: And I'll agree with the prior comments.

DR. SLOAN: I selected three because, you know, the NOPR data suggested that this could be done to some extent in the community, not really highly objective data. And I know from having a specialty in a cancer that's fairly rare that certain community

centers that probably do an excellent job on the common things like lung, breast, colon don't -- often do a very poor job of brain and spine. So I selected three.

DR. BERGTHOLD: I selected four for the reasons that Dr. Henderson, I think --

DR. JUHN: Four, same reason.

DR. WAHL: I said five mainly because the NOPR data included so many enrollees from community settings. Had we had the true Olympic scale, I would have probably used 4.6 recognizing that there probably is variability from center to center.

DR. SATYA-MURTI: The last question, how confident are you that these conclusions are generalized to the Medicare beneficiary population?

PANEL RESULTS: (Wahl, five; Juhn, four; Bergthold, four; Sloan, four; Lichtenfeld, four; Janjan, four; Henderson, four.)

DR. SATYA-MURTI: Yeah. I thought that was the easiest question of all, at least for me.

DR. HENDERSON: Well, in a way what you just

said poses a very good argument. And that is that if we accept the NOPR data -- I certainly do. I didn't accept it exactly in answers to questions number one and two. But it was very credible to me. And if we're going to do that since that's Medicare patients, I think, obviously, the two are inextricably linked.

DR. SATYA-MURTI: And many of them were Medicare age anyway, both in TA and the NOPR.

DR. JANJAN: No further comment. Just I agree.

DR. LICHTENFELD: And I agree. I think the comment about the NOPR data was really instrumental. In fact, I changed my number here a couple times. But that was the one that made me vote four.

DR. SLOAN: The same argument.

DR. BERGTHOLD: Actually, I was very pleased to see the range of ages, you know, the range of -- or the number of Medicare-aged people that were in the studies that were done and the technology assessments. Because often what we see in these panels is that there were relatively few people of Medicare age in

the population studied. So I thought that it was very impressive.

DR. JUHN: I gave it a four for a very similar reason. The tech assessment did not give any indication that the Medicare population wouldn't have the same set of outcomes that the entire population had.

DR. WAHL: I gave it a five because mainly the NOPR data being basically all Medicare beneficiaries in the largest set of data.

DR. SATYA-MURTI: They're coming out with the table numbers.

DR. PHURROUGH: While they're doing that, let me throw out sort of the broader questions around methodology.

I know we ask you to answer specific questions, but I'd like to have a bit of discussion around where should CMS go. We think this CED process is an appropriate process. As Dr. Tunis mentioned this morning though, we are limited in what we can do because we have a defined time interval.

And if we require particular studies in general, we can only -- that's the only coverage you get is in those studies. And you can't have coverage outside the studies, which then limits -- we're either limiting access, if we're looking for some fairly rigorous study or to have broad access, we're doing what we did with this particular issue. We're doing a -- sort of a prospective observational study.

Registries sometimes have bad connotations. I think this was a fairly decently designed observational prospective study that used a registry to collect the data.

So give me your comments. Is this a viable way for us to collect information that's going to be beneficial? Is this kind of observational prospective study of benefit? Not just to us in making coverage decisions, but to treating physicians and patients. Interesting to hear your comments on it.

DR. SATYA-MURTI: I'd like to see what the outcome would be. I know you mentioned that a change in management is tantamount to one of the outcome

changes. So perhaps a little better -- maybe a structure for what outcome is expected.

DR. HENDERSON: I think this NOPR is a spectacular experiment. I congratulate you for participating in it and moving in that direction.

Sure. I agree with you that the trial could be improved. The answer that you are trying to keep the barriers low is a reasonable answer. But I think that one of the things that we want to do in our society if we're going to have evidence-base medicine is get to the point where we -- you know, where every physician and hopefully every patient feels comfortable contributing to that evidence.

I think oncology being a very new specialty compared to the others and one in which we've had to kind of fight to get a toe-hold or certainly at the beginning of my career, people would spit on you as you walked down the hallway. I mean, it was very difficult. So I think most medical oncologists have sort of lived their whole life with the idea that they had to get better therapies. What we had was really

awful with all that nausea, vomiting, hair loss.

I think medical oncology is a field where almost every practicing oncologist has had some training in doing clinical trials. I'd like to get to the point where all of the medical profession is that way. They all feel that obligation. And I think this is an excellent first step. I'm hoping that maybe we could pay $75 dollars and get a little bit more data next time.

DR. JANJAN: I think NOPR has been a tremendous success. From a clinician's point of view, it's very important to understand what's being done and why it's being done. And the data -- and if you go the step further on outcomes, you can produce the data that's not possible for clinical trials because of the shear volume of patients that you've put into the effort and the outcome data that you have available to you.

It takes individual institutions years to accomplish a very small patient number. And I think that's part of the problem with the initial study

presented by the Canadian group. Because so many of these clinical trials have such limited patient numbers, and it's just not possible to get those data. I think this is a very important step forward for all of medicine.

And as a clinician, I think it's extremely important data in helping me go to the next step in my practice of care. So I would very much encourage its use in almost every new technology or question that you have coming down the line.

DR. LICHTENFELD: Well, I'm going to basically echo the comments that have already been made.

I think from the professional organizational side of things, I think it's terrific that the organization stepped up. My compliments to ACRIN for participating. I know that they are participating in other trials. And I think that that's incredibly positive.

I think the difficult question that we were faced with is the perfect versus the good. We do not

have perfect data. We will likely never have perfect data. And you're probably very familiar and many of the people in the room are probably familiar, we've had other instances in cancer medicine in the past where we tried to get some -- through CMS, through Medicare reimbursement -- some handle on various treatments. And we were unable to do that.

Well, here we were -- you know, this was successful this go around. And I think that I would consider the data, you know -- I hate to say the word. I wish it were more academic, it couldn't be. That's what I mean the perfect and the good.

Was it compelling? Yes. It influenced my decisions that I made, clearly impacted my assessment. And it was -- I wish it were complementary, c-o-m-p-l-e-m -- a complementary to the technical assessment.

In fact, it created a great deal of -- as I mentioned earlier, equivocation for me because I had two different data sets showing two different things. And I know that both of them were done well, but how do you make a decision on that? Well, maybe as I've

commented earlier, maybe that is -- not being able to make a decision, in fact, may be a decision. And just looking at some of the numbers up there, they're probably a lot more conservative in some respects than I would have anticipated, given the NOPR data.

But I certainly think that everyone -- I can't speak for everyone. But I know I appreciate having that. It did influence my decision. And I certainly would encourage CMS to make those investments going forward so we can get real life, real-time information from quality centers to show how does this stuff really impact what we do in medical practice.

DR. SLOAN: I agree with the previous speakers, Dr. Janjan and Dr. Lichtenfeld. I think the organizers are to be complimented on the NOPR approach. I think it's really refreshing and novel and may provide -- you know, it's certainly a new approach and may provide a new methodology to address these questions.

At this point, however, it's a little

unclear how to interpret the data because of the subjectivity of the data. And I think I would look forward to seeing the results of CMS's additional studies. And this may be something that can be done for other technologies as has previously been pointed out.

DR. BERGTHOLD: Well, I'm going to mention the patient here because from the point of view of the Medicare beneficiary, this is a huge step forward, I think. And going forward in terms of having access to a treatment, but then being able to sort of contribute your data to knowing how it works or how it doesn't for not only you, but everyone else. So I think this is a big plus for Medicare beneficiaries.

I would say a better construction of the questions to make -- it's all about the question, and it always is. You know, you ask a good question, you get a decent answer. And I realize that the construction of the question was a complex political policy compromise. But I think that's sort of one thing to look into the future.

The second thing for the future would be to try to get ahead of the curve a little bit, hopefully before you're being hit on the head with a hammer, you know, to have a little bit more time to develop the process of -- the CED process.

DR. JUHN: Yeah. I want to add my congratulations to CMS for embarking on this. I think this is very much kind of the next step in what we do with evidence and the decisions that surround interpretation of evidence.

The two comments that I would have -- actually recommendations -- one is I'm not sure that showing the interim results, i.e., the results of a survey, on intention is really -- you know, can sometimes cloud the picture, if you will. So I guess I would suggest if you're going to do this again in the future to really think about, can you actually get to a set of outcomes. And have that presented as opposed to the interim data.

The second suggestion is one way to kind of marry the technology assessment and information we get

from registries like this is to actually have -- in this case the NOPR data and the results from it actually subject to an evidence grading exercise by the folks at the technology assessment. I think it would be kind of interesting to see where they might come out on that.

DR. WAHL: I sort of re-second the congratulations on the NOPR approach. I think it's provided really unique information that had never been available in diagnostic imaging before. The large fraction of data from outside of academic centers I think is a remarkable strength.

And I do think that this whole approach could be explored in other therapeutic and diagnostic approaches. So I think that's attractive.

I would say that the data on treatment monitoring I think if that continues to be collected would be very interesting because that is more limited in the literature. But I have been encouraged that the data show the broad applicability seemingly in decision making and across cancer types.

Can I say one other thing while I'm finishing up?

To conclude my comment, having been involved in working through PET from early stages to now, it is somewhat -- it represents an opportunity for our government to invest in well designed and funded prospective imaging trials before we get so far along. It's been a deficit in not getting those done.

Those have been done in many of our therapeutics, but they have been done much less in diagnostics. And I think that this represents a, you know, future opportunity in investment and to integrate imaging studies more -- clinical imaging studies carefully evaluating technologies a little earlier in their evolution to help maybe make better decisions in the future.

DR. PHURROUGH: From a policy viewpoint, coverage of diagnostics is a mine field, much more so than therapeutics. It's sort of the same in the payment for diagnostics, where we don't have the same type of payment scales as we do in other issues. So

it's -- we're struggling to meander our way through how we manage diagnostics 'cause it's challenging.

PANEL MEMBER: That's because radiology fees are so much higher.

DR. PHURROUGH: It's not just imaging. It's --

PANEL MEMBER: I think I'm the only radiologist here.

DR. PHURROUGH: -- simple CBCs cost lots of money to the Medicare program. You can't go see your physician without having some blood drawn.

DR. SATYA-MURTI: Is the overall -- all panelists versus selected panelists -- can you explain the table a little bit more?

PANEL MEMBER: It's probably an average of everybody in question one and two.

PANEL MEMBER: But everything has overall.

DR. SATYA-MURTI: Pardon me?

RECORDER: The (unintelligible), that's just the board voting members. And then overall is the entire --

DR. SATYA-MURTI: Right. Yeah. You want to go through the same? We don't have to explain why we voted that way. Any comments?

PANEL MEMBER: We brought down the average every time.

DR. PHURROUGH: Is anyone surprised at averages?

PANEL MEMBERS: Yes.

DR. PHURROUGH: Do you want to make comments about any of it?

DR. SATYA-MURTI: The overall tends to be overall the same as the panel members except in the diagnosis column. Two versus one point -- I guess even that is not -- with such small samples, that may not be that different.

Thank you very much.

DR. LICHTENFELD: You said if anyone wants to make a comment. And since I have no fear, you know, I will say I am surprised about how low the scores are. And maybe I did pass the test after all.

DR. PHURROUGH: So your scores were in the

same range as these?

DR. LICHTENFELD: Yeah.

DR. PHURROUGH: So that's why you're surprised.

DR. LICHTENFELD: No. No. And I think that, again, that you know, when you look at it overall and you're just trying to process it, you don't find a whole lot of fours. You find an awful lot of threes, and you find twos and ones. I mean, that's where the trend is.

And given the NOPR data, I would have expected it to be perhaps a bit higher. But I think everybody is caught in the same bind. That's the way I read it.

You know, ultimately, it was tough to be emphatically and enthusiastically in favor of even what appeared to be some of the most positive indications that we had. So I think that it just reflects that overall -- I mean, that's not a statistical glance. It's just having -- you look at a lot of numbers and try to get a sense of opinions.

The panel was much more conservative than I would have anticipated.

DR. JANJAN: I would agree with that. And I think that maybe having a range or having a frequency might also help because it -- with only four even among the total overall score, it still is a small end. But having a frequency may help us understand this better.

DR. HENDERSON: I'm just curious whether -- because I altered my scores today particularly as I heard the OPRR data. So in the end, I went through, and I tried to mentally remember everything that the tech panel had integrated. Did others do that too, or did you --

PANEL MEMBER: Yes.

DR. HENDERSON: So everybody kind of changed their scores during the course of the day 'cause I kind of thought that's what our instructions were.

DR. SATYA-MURTI: Yes.

DR. HENDERSON: Put all the data together, and we were free to do that. Okay.

DR. PHURROUGH: For those of you in the audience, this particular panel was somewhat different than many others in that they had significantly more homework than we typically give MedCAC panels, which may be the reason a few didn't show up and decided to drop out. They decided that homework was past their --

This was a tremendous amount of work. And each of you, I think, have demonstrated that you took the task to hand well, and you focused on this. And it's impressive the amount of the work and thought that you put in. And the Agency is really better off for your having spent this time. And so we thank you for all this work.

For those speakers, we appreciate the work that you put into that. For the NOPR folks who have spent lots of hours not just in preparing for this, but in collecting all the information, thanks to you. And for those others who took their time to present and be present and make their comments known, we thank you for that, too.

And the results of this, the actual voting questions, will be on our website by close of business tomorrow -- hopefully by close of business tomorrow. The transcript will be available in a couple of weeks. Hopefully sooner than that, but I tell you a couple of weeks so that you don't expect something in a couple days. And then they'll be a --

DR. WAHL: I would just say that it would be important in this calculation to determine how blank fields were handled in terms of the average because I'm a little puzzled by some of the low numbers. And if it was left blank, I think it should be clear how -- how the average -- a blank could be counted as a zero. Potentially that may not be appropriate.

DR. PHURROUGH: Blanks were removed, and so they weren't counted as zero. But as you'll see across the bottom, each person's votes will be available. And so we'll post the spreadsheet, and so each person's votes will be available for people to look at.

So thank you very much. You'll see our next

comments and our proposed decision in several weeks. Thanks.

(Whereupon, the proceedings were concluded.)

C E R T I F I C A T E

I hereby certify that the recording represented by the foregoing pages was transcribed by me, or under my direction; that this transcript is a true and accurate record to the best of my ability.

COMPOFELICE REPORTING SERVICES, INC.

Teresa E. Compofelice, CVR