Transcript of November 15, 1999 Morning Session

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

00001

VOLUME I

(Morning Session - November 15, 1999)

HUMAN TUMOR ASSAY SYSTEMS

HEALTH CARE FINANCING ADMINISTRATION
Medicare Coverage Advisory Committee
Laboratory & Diagnostic Services Panel

November 15 and 16, 1999

Sheraton Inner Harbor Hotel
Baltimore, Maryland

00002

Panelists

Chairperson
John H. Ferguson, M.D.

Vice-Chairperson
Robert L. Murray, M.D.

Voting Members
David N. Sundwall, M.D.
George G. Klee, M.D., Ph.D.
Paul D. Mintz, M.D.

Richard J. Hausner, M.D.
Mary E. Kass, M.D.
Cheryl J. Kraft, M.S.
Neysa R. Simmers, M.B.A.
John J.S. Brooks, M.D.
Paul M. Fischer, M.D.

Temporary Voting Member
Kathy Helzlsouer, M.D.

Consumer Representative
Kathryn A. Snow, M.H.A.

Industry Representative
James (Rod) Barnes, M.B.A.

Carrier Medical Director
Bryan Loy, M.D., M.B.A.

Director of Coverage, HCFA
Grant Bagley, M.D.

Executive Secretary
Katherine Tillman, R.N., M.S.

00003

TABLE OF CONTENTS

00004

TABLE OF CONTENTS (Continued)

VOLUME III

PANEL PROCEEDINGS

(The meeting was called to order at 8:00 a.m., Monday, November 15, 1999.)

MS. TILLMAN: Good Morning, and welcome. Dr. Ferguson, Dr. Bagley, members and guests, I'm Kate Tillman, Executive Secretary of the Laboratory and Diagnostic Services Panel of the Medicare Coverage Advisory Committee. The committee is here today to provide advice and recommendations to the Agency regarding formal requests pertaining to human tumor assay systems. This is the first meeting of the laboratory panel.

We are happy to have such a distinguished panel. Thank you all for coming.

Today I would like to welcome Dr. Bryan Loy, carrier medical director, from Administar, who is our guest.

We have one member of the panel who has received an appointment to temporary voting status, and that is Dr. Kathy Helzlsouer.

We have a couple of pieces of business to take care of here. The appointment to temporary voting status. This is signed by Michael Hash, Deputy Administrator for Health

Care Financing Administration. Pursuant to the authority granted under the Medicare Coverage Advisory Committee charter, dated November 24th, 1998, I appoint the following person as voting member of the laboratory and diagnostic services panel for the duration of this panel meeting on November 15th and 16th, 1999: Kathy Helzlsouer, M.D. For the record, this individual is a special government employee and is a voting member of the panel under Medicare Coverage Advisory Committee. We have undergone the customary conflict of interest review and have reviewed the material to be considered in this meeting. Signed, Michael M. Hash, Deputy Administrator.

The conflict of interest statement: Conflict of interest for the laboratory and diagnostic services panel meeting, November 15th and 16th, 1999. The following announcement addresses conflict of interest issues associated with this meeting and is made part of the record to preclude even the appearance of impropriety. To determine if any conflict existed, the Agency reviewed the submitted agenda and all financial interests reported by the committee

00007

participants. The conflict of interest statutes prohibit special government employees from participating in matters that could affect their or their employers' financial interests. The Agency has determined that all members and consultants may participate in the matters before the committee today.

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

Now I am going to turn the meeting over to our chairman, Dr. John Ferguson, who will introduce the panel.

DR. FERGUSON: Good morning, and

welcome to everybody here. I would like to have the panel members introduce themselves, starting from over here on my far left.

MS. SIMMERS: I'm Neysa Simmers.

DR. FERGUSON: Could you also say where you're from and what you're doing in life?

MS. SIMMERS: I am Lisa Simmers. I'm from Bridgewater, Virginia. I am currently a

00008

health care administrator, and am here in the interest of the laboratory community, I guess.

DR. SUNDWALL: I'm David Sundwall. I'm a physician and I'm president of the American Clinical Laboratory Association, in Washington, D.C.

DR. FERGUSON: You have to speak into these microphones like a rock singer, I think.

DR. KLEE: I am George Klee. I am from Rochester, Minnesota, and I'm a clinical pathologist.

DR. FISCHER: Paul Fischer. I'm a family physician from Augusta, Georgia.

DR. BROOKS: John Brooks. I am chairman of pathology and laboratory medicine at Roswell Park Cancer Institute.

MR. BARNES: Rod Barnes. I am the industry rep on the panel. I work for AlCon Labs in Fort Worth, Texas.

DR. BAGLEY: I'm Grant Bagley. I'm the Federal representative on the panel, and director of coverage in HCFA.

DR. FERGUSON: I am John Ferguson. I am a practicing neurologist, and I have just retired from the NIH, where I directed the

00009

consensus development program for the last 11 years.

DR. MURRAY: I'm Robert Murray, a clinical biochemist in practice in Chicago, Illinois.

DR. LOY: I'm Bryan Loy. I am with the Kentucky Medicare carrier. I represent the Medicare system at the state carrier level.

MS. SNOW: I am Kate Snow. I am the consumer rep on this panel, and I am the director of senior services for Northern Michigan Regional Health Service, and I am an advanced practice nurse in gerontology.

DR. KASS: I am Mary Kass. I am chairman of pathology at Washington Hospital Center, and director of integrated laboratory services for MedStar Health.

DR. HAUSNER: I am Richard Hausner. I am a pathologist practicing in Houston, Texas.

MS. KRAFT: I am Cheryl Kraft, administrative director of laboratory services, Minneapolis.

DR. HELZLSOUER: I'm Kathy Helzlsouer, medical oncologist and professor of epidemiology at Johns Hopkins School of Public Health.

00010

DR. MINTZ: I am Paul Mintz. I direct the clinical laboratories and blood bank at the University of Virginia Health System, where I'm a professor of pathology and medicine.

DR. FERGUSON: I would like to now turn this over to Grant Bagley. Grant?

DR. BAGLEY: I'll just make a couple introductory remarks and sort of bring everyone up to speed about what we're doing and how the process works.

The coverage process for Medicare is one which from the very inception of the Medicare program has been marked by local diversity and at the same time, the ability to have national conformity when the science and practice so dictates. It has always been that way and it continues to be that way today.

What we're about here is considering issues for national coverage decisions. Very much like the federalism model for everything else, states or in this case Medicare carriers, can have variable policies, but when the science, when the issue is sufficiently justified, we can develop a national coverage policy. That national coverage policy then takes precedence,

00011

and all Medicare carriers in every state and every area follow that same process.

So we are going to talk a little bit about how Medicare coverage works and how we are going to deal with it specifically in this issue.

What we're talking about is the Medicare statute, and the Medicare statute has one overarching principle, which is in the terms of a bureaucrat, 1862.A.1(a) of the Social Security Act, and this is what it says: That no payment shall be made under Medicare for a service which is not reasonable and necessary for the diagnosis or treatment of an illness or injury. Those are very important words. Reasonable and necessary, diagnosis or treatment, and a disease or illness.

Now, reasonable and necessary has never been defined. We've never defined it explicitly and said, this is what it takes to prove reasonable and necessary. But over the years we have articulated principles by which we say, reasonable and necessary means the following things: It doesn't mean safe and effective. It has to be safe and effective to be reasonable and necessary, to be sure, but it has to be a bit

00012

more.

So in terms of what we have required to show something is reasonable and necessary, first of all, if it requires a safe and effective determination, and clearance for marketing by the FDA, we've always considered that to be a first step. And second, if it doesn't require clearance for marketing by the FDA, we still make an inquiry that it must be safe and effective. But demonstrated effectiveness is one in which we have said the benefits have to outweigh the anticipated risks. It has to be FDA approved, if required. And there has to be authoritative evidence that it improves outcomes, because after all, that's really what we're talking about.

So really, the difference between a

threshold issue of is it safe and effective and can it be marketed is somewhat different, you know, and we have to look at it a little bit more. So it has to be safe, to be sure. Any product, even a diagnostic test, has to be safe. It has to be effective, that's clear. But it also has to have benefit which is outweighed, or at least outweighs the risk involved in even the procedure or even a diagnostic test, because it's

00013

going to guide therapy.

But not only do the benefits have to outweigh the risks, but there have to be some kind of outcomes, there has to be some improvement in clinical care. Is it an improved outcome, does it give better treatment, does it give better results, or in terms of the diagnostic tests, does it give information which can guide or improve therapy.

And finally, does it have value? Is there any value to this procedure? For diagnostic tests, it's an issue; for anything else it can be an issue of does it improve therapy, do we get not necessarily better survival, but do we get improved quality of life.

Well, how do we determine this? And again, we've articulated these over the years and said, you know, we have to look at clinical studies and from these clinical studies, we have to be able to make determinations. And so we have to be able to look to available evidence and say, are there fundamental safety questions? Does a product live up to its claims? Does it provide the clinical utility that we can use in

00014

practice, because after all it has to be, remember the statute, reasonable and necessary. And we look at the outcomes and do the clinical studies to provide evidence that there is an improved value from the service.

Of course we can look at it in a number of ways. We can look at outcome measures in

terms of simply survival, that certainly is the crudest measure we can use for an outcome. But we can look at process changes, which may be indirect, and we can say, how does it influence the disease process, and can we make inferences about value from that. And we can observe just simply effects in terms of does it change a measured process, does it change a physiologic process. Is blood pressure improved? Do we have a metabolic process change? Is cholesterol lowered? And then can we relate those to an outcome.

So even when we look at secondary end points, when we are looking at a physiologic measurement or metabolic change, can we make the direct link to an improved outcome. And I think it's going to be important to keep that in mind as we look at intermediate end points.

00015

And in terms of looking at the science, this has always been what we've used to determine what's reasonable and necessary. You know, if we look at information, we look at collections of data, and we look at studies, is to keep in mind that we have to consider the bias that can be introduced, are patients selected in less than a random fashion so that the outcomes might be, you know, influenced by the way patients are selected. Do we select patients for one group and then do they become evaluated by another method in terms of trials with more than one arm. Do patients disappear after being entered into the study, and if so, for what reason? Is this going to affect it? Do we have some way of having people evaluate the results of the study without knowing what the outcomes should be or are going to be? And is there an adequate way to control for the information?

These are all things to keep in mind when we're considering clinical data and clinical study. Are they big enough? Are we measuring something which is large enough that we can make a determination? Whatever we've measure, have we

measured enough to say this is truly an effect?

Do we have enough subjects in here? Do we have enough patients to make that determination?

And within Medicare, we always consider what we're dealing with. I mean certainly, some diseases have a very high prevalence, they have a large impact on the Medicare population, and in those situations we need to have a great deal of evidence to make a change. On the other hand, some diseases are not highly prevalent, they deal with just a smaller population of people, and in those cases we have to look at the degree of precision in the clinical studies in a somewhat different way.

To consider the natural history of diseases and the issues we're talking about today, we have to consider what the uninfluenced outcome would be in terms of what kind of a difference does it make when we start to alter things. And in looking at clinical studies, we have to consider both the issue as presented and we have to look at the source they came from.

I think I was on a panel with some folks from Australia where they do a much -- they have a much different process than we do in terms of deciding their coverage in terms of looking at

not only the science, but once they've looked at the science, they then look second, and they say now that we've decided we're going to cover something based on the science, let's decide if it's worth it, let's look at what it costs and make that determination. And in doing that we made an interesting point, which I think is worthwhile to relate here, and that is that if you're going to look at a survey, you know, if you're out to buy a car and you're looking at a survey, and you want to look at the report of all the new options that are available in new cars, that you're probably going to say it makes a difference to me whether this is from an independent consumer agency or whether this is a

report produced by the auto manufacturer. And the same thing ought to be true when we look at studies, when we look at clinical information. We need to look at the source and say, not necessarily that there is bias introduced, but we need to look at full disclosure of the source of information. We need to look at whether there is real bias or whether there's apparent bias, and that it's not really there, so we need to look at the source of information.

We also need to look at the credibility of information. There is a hierarchy of evidence which we need to consider, and that published peer review literature is considered evaluated by the larger community. It's considered to be of a higher validity and we need to consider that. We aren't always going to have that and we will have to look at things which are not peer reviewed but have been presented and published without peer review. We're going to look at things that are not even published in full form, so-called abstract form, where they have been presented at a meeting and an abstract is simply printed, and we are going to look at unpublished data which has been subjected to considerably less scrutiny.

So those are the kinds of things we're going to look at in our evaluation of evidence presentations today. You need to look at what the information is showing you. You need to look at what's happening, where it came from, and then evaluate it based on its quality, the hierarchy of the evidence and the source.

Now the process we're using, and it's the process Medicare is now using for national decisions, is very different from in the past.

This is a part of that, this open public committee meeting, and we put together an entire new process which is now open. It's open because this is a public meeting, and we're having a full and frank public discussion on an issue. It's quite defined in terms of its process. We bring

together technology issues and you'll find that we aren't bringing together specific products or specific tests, but we are bringing an area of technology together and we're looking at a whole area of technology, because we don't cover products, we cover areas of technology. It's one in which there's going to be full public participation here, and we're going to make an explicit decision based on the recommendations of the panel. And I want to make it very clear: This panel does not make coverage decisions. This panel is for the purpose of giving us technical advice so that HCFA itself can then make an explicit decision. And not only will we make a decision in a fair and proper fashion after final panel recommendations, but that that decision is subject to challenge, if we significantly misinterpret the evidence or if we fail to consider all the evidence. Of course

00020

this is a public process, so we expect all the evidence to be here.

So in looking at the clinical trials we're going to look at, we want you to focus on looking for definitive answers, clinical utility, does it improve outcome, is it appropriate, and can we determine for which patients it should be appropriate, so we can administer the Medicare benefit. Is the source of the information unbiased, is it free of conflict, and can we make noncontroversial decisions? Because remember, when I described the process, the very last step of this whole process is subject to challenge when we make a decision. And by this open committee, this advisory process in which there is full participation with the public, we would expect to address all of these issues so that there would be little basis for challenging any decision.

Now, getting to the issue at hand, of looking at sensitivity and resistance tests in terms of oncology, just a little bit of history. Medicare has looked at these in the past. We've

looked at them for quite a while. As long ago as 1980, the technology was around, we looked at these technologies, and at that time what HCFA used to get internal advice was the physician panel. The physician panel was a group of physicians who worked for the Health Care Financing Administration, and also physicians who were in the Public Health Service, with other agencies, that were brought together to look at scientific issues and give the Agency scientific advice. It was an internal predeliverative advisory panel, something which I should say also, is perfectly acceptable, even under the Advisory Committee Act, because it predelivered consideration by internal government employees. This physician panel got together, looked at the issue and requested a technology assessment at that time. And that has been a number of years ago.

It was then, as late as 1987, that same physician panel met again. Based on the technology assessment that was available at that time that they reviewed, and they considered the issue and felt that the use of tumor cells for sensitivity determinations was still experimental and that there was not enough information to provide coverage at that time.

In 1991, the issues were looked at again, and at that time we were using, it was the physician panel but it was now called the technology advisory group, still composed of internal government physicians. They were from a little bit wider scope in terms of bringing folks from the FDA, from AHCPR, from NIH, and they discussed this at the same time, and they discussed the issues around assays at the same time, and agreed that the existing language which we put in the coverage issues manual, which said that this technology was that this technology was at that time experimental, should be retained.

It was then in '97 that the technology

advisory committee, which was an outgrowth of the same internal deliberative body, looked specifically at extreme drug resistance testing and considered whether or not extreme drug resistance testing was in fact the same kind of technology that was looked at before in terms of sensitivity. Was it really the same thing, was the technology the same, and was the utility the same. And at that time the technology advisory committee came to the conclusion that perhaps extreme drug resistance testing was enough

00023

different of a methodology from sensitivity testing, and that the clinical utility was enough different that the coverage issues manual exclusion of this technology, saying it was experimental, that was put in over ten years previously, perhaps didn't apply and that drug resistance testing was enough of a different technology that it could be left to carriers to have discretion to cover that technology.

The current policy, then, is that human tumor drug sensitivity assays are considered experimental and therefore, not covered under Medicare. That is a statement which leaves no discretion for Medicare contractors, and that statement is in force today.

Now it's interesting in that we made the interpretation through the technology advisory committee that drug resistance testing was enough different from sensitivity assays that it was not covered by this prohibition, and there has continued to be confusion over that issue over the past, you know, several years since this was done. So that's what's currently in place.

In order to reevaluate our position in that coverage issues manual statement, we have

00024

convened this panel and we are going to present the following questions, and I will quickly go over them, because these questions are going to be focused on tomorrow and are going to be ones that we are going to ask the panel to answer.

First, is the scientific evidence that is amassed thus far, presented to the panel and that's going to be discussed here today and tomorrow morning sufficient that we can make the appropriateness determinations about a coverage utility and about what should happen in terms of clinical care in using these tests?

Are the assay techniques described in the literature for single drugs sufficiently transportable to multidrug therapy, and there's going to be a question presented about the appropriateness of using single drug information in terms of testing in multidrug regimens in terms of treatment.

Does the scientific evidence demonstrate the clinical benefit? This can be very important, because there's going to be information presented about different kinds of tumors, hematologic tumors, solid tumors, and being somewhat different in character, and should

00025

we be able to make determinations of clinical care based on testing which are going to guide therapy, because after all, remember, we talked about clinical utility and value as being very important things that we can draw from these tests. So if we can't make the clinical utility argument or if the value isn't there to be able to directly influence therapy, then that's going to be important to consider in terms of is it reasonable and necessary.

If test results in terms of sensitivity or resistance give us predictions about a tumor response, should in fact those predictions guide what happens in terms of direct therapy? Because after all, one of the things we need to know, which as I stated, we need to know not only clinical utility but appropriateness, and we need to be able to say when is it appropriate to do these tests and what is the appropriate clinical action once the test is done, because that's the kind of information we need to be able to put together a coverage policy.

Is there sufficient scientific evidence to demonstrate the clinical utility in selecting appropriate chemotherapy?

And finally, the committee will be given the opportunity to raise any additional concerns.

So basically, these are the questions that we're going to present to the panel. What should we be looking at in terms of measuring this technology? Should we be looking at survival or should we be looking at intermediate responses? Are there appropriate measures that we can look at in terms of response to the tumor, in terms of quality of life, in terms of other intermediate outcomes, or should we be looking at survival, and is one an appropriate surrogate measure for the other. Is information on single versus combination drug regimens relevant in terms of using the results of the test in clinical care?

Does the evidence that's presented that we consider here over the next day and a half demonstrate that there is in fact a clinical benefit, not just interesting information, but is there a clinical benefit which we can derive from the use of this methodology? And if there is a clinical benefit we can derive from this methodology, should it in fact determine what the treatment should be in terms of particular patient care.

And then finally, are there additional concerns that the panel after a day and a half of considering this technology wishes to bring to the forefront?

So those are the issues we're going to be talking about and that's a general overview of the process HCFA uses and the kinds of information and the level and hierarchy of evidence that we wish to have considered over the next few days, or day and a half. We're going to present those questions tomorrow. There will be

a discussion of those questions, and we will be asking the panel to vote specifically on those answers and give us determinations which we can then use in the form of recommendations to either clarify, to ratify or to change the existing policy which we have, which is noncoverage of human tumor assay systems, and I think a somewhat confused approach to drug resistance testing in terms of that coverage issues manual application.

So that's the charge to the committee. There will be public presentations. There will

00028

be presentations from HCFA and from other sources. I think we will all hear different interpretations of information. There will be full discussion, and we look forward to this process giving us recommendations which we can then use to modify or ratify our existing policy.

MS. TILLMAN: Now Dr. Ferguson has a few remarks to make.

DR. FERGUSON: Thank you. It's the job of this panel to review the evidence and its quality for this group of in vitro drug assays, and arrive at some conclusions regarding the appropriateness of these tests in treating cancer patients. In an ideal world, the evidence would dictate yes or no. Unfortunately in the real world, we are likely to find something less, certain conditions for select patients, et cetera. Asking the research community to consider patient outcomes in evaluating new diagnostic tests seems to be setting the bar higher for the quality of evidence than previously. However, I believe that all of us want the best possible outcomes for all the patients we see, no matter where we sit.

00029

As we spend proportionally more money on health care, we should try to achieve the best possible outcomes for our patients, and this may require setting the quality of evidence bar

higher than 10 or 20 years ago. The job of this panel is to evaluate the data we are presented with for these assays and to use this evidence to answer the questions that HCFA has posed. It's a bit of a conundrum for society, I think, only paying for what works and yet not stifling innovation in the process. Our job, this panel's job is not easy, and we recognize that the presenters don't have an easy task either, especially given the short time to present their work which has occurred over a number of years.

In the interest of time, I would like to try to encourage all of the presenters to stick as closely as possible to the outlines we have, and I would like to get started with our FDA. Kate, do you want to introduce Dr. Harvey?

MS. TILLMAN: Sure. Our first speaker is Dr. Brian Harvey, who is the associate director of the Division of Clinical Laboratory Devices for the Food and Drug Administration. Dr. Harvey?

00030

DR. HARVEY: Good morning. First of all, I would like to thank the Health Care Finance Administration for the invitation for us to speak this morning, and I would like to commend HCFA for moving towards an advisory panel process, and we are glad at FDA to be a participant in that process. What I would like to do this morning, I am Brian Harvey, a senior medical officer at Center for Devices and Office of Device Evaluation, and currently acting associate division director in clinical labs. And what I wanted to do this morning is actually talk about the FDA process.

Often when we hear about HCFA's role in the evaluation of new technologies, we hear well, if the advice is FDA approved then it can go on to the HCFA process. And what I -- the major points I really want to get across today is that there are many roads to the U.S. market that medical devices can go through. One size does not fit all. And by actually going over the

various methods that medical devices can get to the U.S. market, give a better understanding of sort of the terms approve versus clear, exempt, et cetera.

As most of you know, the regulation of medical devices in the United States really didn't start until May 28th, 1976. There were some medical devices that were regulated under the drug law before that time, but the vast majority of medical devices began to be regulated with the medical device amendments to the Pure Food and Drug Act, May 28th, 1976. The law itself was sort of a hodgepodge of many different concerns, which sort of reflect the great variety which are medical devices, and as we go through some of the aspects of the law, you'll see how the fact that the majority of devices were not regulated has really fed into the whole construct of medical device regulation. I will touch upon the Safe Medical Device Act in 1990 as well as the more recent FDA Modernization Act of 1997, which we all call FDAMA.

So once again, medical device amendments, 1976, it was the outline which we still use today stratifying medical devices in classes. It's a risk based classification, class one being those devices which are very low risk, class two devices an intermediate or moderate risk, and class three devices being the highest

risk devices. There is actually a very long definition of what a medical device is, I'm not a lawyer, but I won't even spend the time reading that. It's a full page long and in the interest of time, the point being it is defined in law as well as defined in terms such as safe and effective.

The vast majority of medical devices in the U.S. do go through something that's called a 510(k), which I will explain in a minute. That aspect of the law was strengthened with the 1990 SMDA law. It required an indication for use

statement, so therefore in a specific part of the application, the specific indication for use for which the company, the sponsor wishes to get FDA clearance was clearly stated. There was a summary of safety and effectiveness in each application, and FDA through the Freedom of Information Act, was able to make that available to the public to give an insight into what led to different medical device decisions.

With the FDA Modernization Act, there were actually several other aspects that were clarified. The Center for Devices, a few years before this act, actually instituted a

00033

reengineering effort and many of the aspects of the FDA Modernization Act became codified in the law through the FDA Modernization Act, trying to increase the emphasis on post-market evaluation of devices, but keeping an adequate premarket evaluation, the whole concept of interactive reviews, trying to increase communication between the industry and the FDA. Greater inclusion, not only in the public advisory panel but internal meetings. Greater outreach to academic societies. And there is also a section of the FDA Modernization Act, Section 205, which many of you have been hearing about in the news, which is the least burdensome method to get to the U.S. market.

And actually, I recommend that you all go to the FDA website, which I will give later on, to look at the draft guidance document on least burdensome, because we still are in a public comment period and we welcome your comments. One of the things you will note is that the current document says it does not apply to IVDs, in vitro diagnostic devices, and one of the clear aspects that we are getting in the public comments is the importance of including

00034

IVDs in the process. And as part of my efforts in the clinical laboratory division is to incorporate a section for in vitro diagnostics in

the least burdensome framework since it's a very important aspect of medical devices.

So what are the different roads to the U.S. market? Well, starting in the beginning, under the IDE, or investigational device exemption, if something is considered a significant risk through the local IRB, it then comes to FDA for a review of the protocol. In 1995 an agreement was worked out between Health Care Finance Administration and the FDA to try to designate what was a truly experimental investigational device and what was a more run of the mill or traditional device that just was a newer version. One of the aspects of medical devices for those who are involved with drugs, sometimes catches people off guard, is how medical devices really are just sort of a technology creep.

And let's say in pacemakers, the older version through the newer version, very, very minor changes require a new application. So you may have a device that's very similar to the

00035

traditional device that's now in an investigational device exemption study, and it might not have gotten covered because due to a new bell or whistle, it was not yet on the market. So through the wisdom of an agreement between HCFA and FDA, the decision was made, there really should be a designation where FDA says this is a way out very experimental device, or this really is just a minor modification, in order to meet the FDA requirements they are going through an investigational device exemption.

But our recommendation is based on our evaluation and it is a nonbinding recommendation, that this should be considered for HCFA reimbursement. So the A versus B designation, B being that FDA feels that it should be considered for reimbursement, and 80 to 90 percent of IDEs actually have that B designation. So if something is an established IDE, it gets this B designation as something that could be considered

for reimbursement. So, the original idea in this 1976 law, the whole concept being is that there was a number of medical devices that were on the market; through the use in the market, they were found to be safe and effective, and if that

00036

device was on the market before May 28th, 1976, and a sponsor or company could come in and show that their new device was substantially equivalent to that old device, they could submit a premarket notification, designate it 510(k) based upon the law, that line of the law, and they were able to get to market.

So they did not to establish de novo safety and effectiveness, but through a substantial equivalence flow chart, they are able to show by direct comparisons, both clinically, engineering, bench testing, et cetera, that these devices are substantially equivalent. And what we have actually found in a very positive way is that there has been a technology creep and improving of devices, although the older devices, the newer devices were found to be substantially equivalent to the older devices, when you actually look over time, there is a gradual improvement.

So it's been a way for the companies to innovate. It is actually in the spirit of least burdensome, long before that provision was written, a way to get to market in a smaller package not requiring advisory panel review, but

00037

with internal review for these devices to get to market. And as part of the broader market, traditionally class one 510(k)s were very small; class two 510(k)s, depending on the type of device, were either smaller or larger, depending on whether or not there was needs for clinical data. And then there were some class three devices that were deemed to be 510(k)s. As part of the more recent reengineering efforts and recent laws, those have actually either been converted to class three PMAs, which I'll talk

about in a minute, or have been down classified to class two.

So in the broad scheme of things, the way to think of it is, the vast majority of devices on the U.S. market are class two 510(k)s, and if you look at the numbers from fiscal year 1998, there are about 4600 class two 510(k)s cleared for market, and the term is cleared for 510(k)s, versus approved for PMAs. 4600 510(k)s compared to about 50 PMA applications that were approved, and about 250 to 300 PMA supplements. So you can see, the vast majority of medical devices in the United States have actually been cleared through the 510(k) process.

00038

So the PMA is the premarket approval application. That's the one that people traditionally think about when they think of an FDA approval. It's based on valid scientific evidence. Often an original PMA has to go to the public advisory panel for their recommendation. The valid scientific evidence is actually defined in the law as well controlled investigation, partially controlled studies, studies without matched controls, well documented case histories, reports of significant human experience. So you can see some parallels between the FDA law and the HCFA law that Dr. Bagley alluded to earlier.

So you can see, it's the whole gamut of different sorts of both clinical and evidence. Now in the original 1976 amendments there was another route for class three devices to come to market, and that was the PDP, or product development protocol. And the thought was that if there was something that required a clinical trial, that the companies may want to have public advisory input long before they got to the final presentation that normally happens in the PMAs. So what happens is that a company submits a PDP protocol, in the PDP it must have the animal

00039

testing, the bench testing, as well as a proposal for a clinical trial. That is then reviewed by

the FDA and taken to a closed session of an advisory panel, since it's still proprietary information, confidential information. The advisory panel comments on the protocol design and has input into that. As part of the protocol, there are actually set end points that have been designated for success criteria. So obviously, it's to be used with those medical devices that are well know as far as what to look for as far as a success criteria. Then if it's deemed approved by the panel and the FDA, the sponsor or the company goes out and does the protocol, and if they actually meet those success criteria, the PDP is deemed approved and you do not need to go back to the advisory panel for a final approval. So once again, another path to market, it's not a PMA, not a 510(k), but it's equivalent to a PMA for a class three device.

Another aspect, another way to market is the HUD or humanitarian use device. This is the, HUD is analogous to the orphan drug part of the drug law and actually, the initial application to FDA for designation goes through

00040

orphan drugs at FDA in the Center for Drugs. And the concept is that if there is a disease which affects less than 4,000 people per year in the United States and is not being adequately being treated by any current medical device, then a company can come in, and if they have been given that designation by the orphan drug people at FDA, then they can submit an HDE, as opposed to a PMA or a PDP, for their class three device. Safety definition in the law and in practicality is the same as a PMA or PDP. However, instead of establishing effectiveness, they only have to show probable benefit. And the concept being, is that there are fewer patients to study, the benefit to this patient group far outweighs the risks based upon the safety analysis, and the review time is shorter, and these devices are able to get out to the public.

So one of the things that the HCFA

advisory panel may be asked to comment on, not only this panel but all of the panels, is this whole area of HDE. So it is an FDA approval just like a PMA or PDP for a class three device, but the criteria are different. And once again, the point being that there are these many ways to get

00041

to the U.S. market. And perhaps the best way to describe it is not so much an FDA approval, but has a certain medical device met the FDA threshold?

So it gets us into specifically in vitro diagnostics and there are sections of the law and the regulations that deal specifically with in vitro diagnostics, and I can see I'm running late on time. Many of you are familiar with all these labeling requirements, the whole concept of reagents and instruments, how these are all integral parts of in vitro diagnostics. Laboratory tests, if something is done at a specific laboratory, it's not exported anywhere else, it's sort of the concept of a home brew, the FDA has chosen not to regulate at this time home brew assays. These are considered class one exempt medical devices. Now if you had a home brew which then was being exported, then there may be parts of that which would be subject to some of the various aspects of FDA regulation.

Just on a side note, the whole CLIA effort, which is currently being run by CDC, the decision has been made to transfer that to the FDA, so this will therefore be the same

00042

regulation, and the FDA will also, for in vitro devices, will also be doing a parallel clear review. At this time there are no planned changes in the criteria that CDC has been using, but you will be hearing more of a clarification on that aspect of the law.

So now the issue that was in the news this past week, the analyte specific reagent, it was an area that was sort of an internal SOP, a standard operating procedure in the in vitro

diagnostic group, but the rule was formalized on November 24th, 1998. The concept being is, although you have these home brew assays, they are only being done at one site, you wanted to make sure that the various components of those home brews met certain FDA criteria, the whole concept being is that if you had an analyte specific reagent, you wanted to make sure it met certain good manufacturing levels. And as you can see in the actual regulation, they talk about antibodies, specific receptor proteins, nucleic acid sequences. So you see a heavy emphasis here on biological agents which for other, in other contexts may actually be regulated at FDA through the Center for Biologics. And there are various

00043

impacts on manufacturers on labeling through the analyte specific reagent.

So to get to today's issue, initially FDA was sent a letter from HCFA, and the inquiry was, are there any medical devices that had been FDA approved that fell under the scope of the types of medical devices that we're going to be discussing at today's meeting. And at a branch level, the reviewers who were involved in this area went through the database, and their initial review was that there was nothing in the FDA database. Because of that review at the branch level, a letter was issued, from which many of you have seen the letter, which actually generated quite an industry response and actually is part of the whole concept of least burdensome and FDAMA and interactive process, this has actually turned out to be a good thing.

From my point of view actually, this is when I was brought into the process. I was not directly involved with that initial review. But what we did, based upon the overwhelming input from the industry, it triggered an internal FDA review of the issue, and it actually went up to the level of the new center director, Dr. David

00044

Feigel, who took over for Bruce Burlington, and

it actually turned out to be a good thing, because Dr. Feigel previously was at the Center for Biologics, was very familiar with the various aspects of what is covered under the analyte specific reagent through his work in biologics, and before that he was a division director in the Center for Drugs. So we were very, very lucky to have sort of a broad perspective of Dr. David Feigel.

In addition, Linda Kahn, who was one of her deputies at the center level, was involved, and she was a lawyer by training, had spent a lot of time up in chief counsel's office at FDA, and her input in reviewing the actual regulation and the spirit of the regulation came into play.

And I was also actively involved, and my role was, I am board certified in internal medicine, I still practice on evenings and weekends, and before that I was a research biochemist, so I sort of brought both a practical clinical approach to the problem as well as a traditional Ph.D. biochemistry approach. And that with Dr. Gutman, who is the division director, in looking at what was the spirit of

00045

the analyte specific reagent statute, and the spirit was that although with home brews, they are used at one site, we want to make sure that good manufacturing practices have been used for all the various components.

And in those home brews that use FDA approved drugs, that really is not an issue. If there is a drug which is a chemotherapeutic agent that has been through the FDA approval process, although not at the Center of Devices but the Center for Drugs, they have met all the strict criteria in manufacturing that are really necessary. So really, when you look at the spirit of the regulation, anything that contains an FDA approved drug really does meet that spirit of that.

So the official -- the follow-up letter dated November 9th, did go through to say that

based upon further evaluation, the FDA not believes that the drugs being used in these assays fall outside the scope of the analyte specific reagent rule and because these products have been approved and are regulated by the Center for Drug Evaluation and Research, there is assurance that they have been produced in

00046

compliance with good manufacturing practices. We have concluded that in-house home brew assays prepared using these reagents do not need to meet the requirement of the rule. And then it goes on to say, however, we recommend that certain labeling requirements be considered when these are done.

Now as a caveat to that though, however, if these are ever used in kit form, or that kit could be sold and exported to may different laboratories, then they may fall inside the scope of a class three PMA or PDP, or depending on the claims, a 510(k). But if it's at a specific site and falls under the home brew concept, then that's not something that requires FDA direct review. So that's -- I just wanted to go into those details.

So to summarize, and to get additional information on all the different areas I talked about today, there is a group called the small manufacturers assistance, and you can be a large or a small, you don't have to be small by definition. They are a group of people who have access to all sorts of information at FDA, and now with the worldwide web, there are various

00047

parts of the FDA web site for the Center for Devices that have guidance documents in all the various aspects of which we talked about today. We encourage you to go to that. If you have specific questions, you can talk to the small manufacturing people, and you're certainly always welcome to call us at the clinical labs.

But to summarize, I think the best way to consider the FDA process is that there are

various ways for devices to get to market. Just to review, there's class one exempt, so therefore, we never see them, but when we say exempt from 510(k), we don't mean exempt from good manufacturing processes. There are those class one devices that have been reserved, and they still do have to come to the FDA. Class two 510(k)s, we spent time talking about. And finally, for class three devices, PMAs, PDPs, HDEs.

So therefore, perhaps the best way to talk about it is has the FDA threshold been met? And then it's ultimately up to you all to look at the evidence from there. Thank you again for your invitation.

DR. FERGUSON: Thanks, Dr. Harvey. I'd like to go right ahead now with Mr. Kiesner, from Oncotech. Are you ready?

MR. KIESNER: Yeah.

DR. FERGUSON: Since we are 15 minutes later than the schedule says, I'm just going to put things 15 minutes ahead, and take it out of the lunch period at this point.

MR. KIESNER: Thank you very much. My name is Frank Kiesner. I am president and CEO of Oncotech, one of the companies that are in this industry. I am here today to give more of an overview of the industry and set the stage for subsequent discussions which will focus on the clinical utility and the clinical application of these technologies.

Before I begin, I think it's important to recognize that we are sharing in an historic moment here. That this type of panel, this type of open discussion of medical and patient issues is just starting and that from the outside, we in the industry have been able to witness the gestation of this process, and I can honestly see that what we are involved with today is a major step forward and I think that the coverage and analysis group should take credit for that.

As Dr. Bagley was talking, I recalled a town hall meeting that I attended about a year and a half ago, where Dick Coyne and Dr. Bagley proposed some structures. There were about 600 of us in the audience, and based on that meeting if there is one thing I am absolutely certain of, is the HCFA staff went through the legal, political, the administrative issues relating to this process. They definitely were not short of free advice.

Secondly, I would like to comment just briefly about the FDA issue. And you have all read the letters going back and forth. We are very pleased that this issue was resolved, and with Oncotech, we live in a glass house. By that I mean we every day have to deal with our own issues and our own problems, and I would only hope as we deal with these, that we have ourselves the same sense of urgency, the same decisiveness, and the same unfiltered honesty that we have witnessed within the FDA over the last three weeks, and I think it's a real credit to their organization and to their management. We are very pleased that the issue was resolved.

I have tremendous respect for the

00050

people that participate in this industry. They are motivated by doing what is good for cancer patients. I am going to share some numbers in relation to the industry to try to get things into a setting. The problem is while everybody is willing to contribute their numbers to industry numbers, there are antitrust issues and problems with duplication of numbers, so what we've chosen to do is just look at the Oncotech numbers, but recognize that the work of others in the industry would probably increase the numbers I am going to show about 25 or 30 percent.

In terms of drug resistance testing over the last several years, over 55,000 cancer patients have been tested. If you look just during the last year, or year and a half, we have received tissue samples for testing from over a

thousand hospitals throughout the United States, we have reported results to over 2600 physicians, and we have tested 60 different tumor types. The technology is being used in the medical community.

Where are we in terms of payor acceptance? The story really began in 1994 when Blue Shield of California had a panel meeting

00051

just very similar to this, open discussion, presentations from those in the industry, and a good solid dialog of the science. What they concluded in 1994 was that drug resistance testing in oncology is accurate and reliable and there is sufficient data to determine their safety, clinical utility and impact on clinical decision making.

Where have we gone from there? If you look at current payor acceptance, in terms of payor contracts, we have with different managed care entities, 31 million lives under contract as far as payment for drug resistance testing. We have a contract relating to the pricing that involves 2300 hospitals around the country. In terms of not the contract, but in terms of what our payment experience has been for this type of service, in terms of non-Medicare carriers, the managed care and the third party or the indemnities, in the last year and a half, we have probably billed about 17 to 1800 different entities. 1600 of those have paid for the EDR. And I don't want to imply that they've paid everything that we've billed, but they have paid for, they have paid some amount for EDR.

00052

The second thing is that in relation to the question of medical necessity or investigational denials, less than one percent have been written off for this reason. Now what I mean by written off is very important, and it's not that questions haven't been raised. At any given point in time, our finance department would be dealing with 25 to 50 different carriers, and

we would have to deal with the question of medical necessity or investigational status. What that number indicates is that after we go through that process, that less than one percent are actually written off on that basis. So I want to be very clear on that.

How does that contrast with the Medicare experience? Basically, all of our claims have been denied on a local coverage basis, as the technology being investigational. But that's the purpose of this meeting; we are looking at developing a national policy that will be able to integrate all of the information that is current, into a rational approach to this group of technology.

Dr. Bagley alluded to the carrier issues manual, the national coverage policy,

00053

5041. It was originally enacted in relation to a human tumor stem cell assay. It was 1970s technology. There were technical problems with it. Basically, it was used only in a research setting and for the last 15 years has not been used clinically. It was a very important technology though, and it was important because it highlighted some of the issues involved with the testing of cancer on an in vitro basis. And it was a major step forward because the people that were involved in that technology learned from it and went into a second generation technology, and ultimately to the technology we're using today, which is a third generation technology. So the point is that there has been an evolution in technology, there's been a learning process and a growth, and I would just urge that we look at what is available today and what is the science to support what's available today. It's not something that was in existence in 1982.

A recommendation is that this provision 5041 should be removed. It was the right thing to do at that time, there is no doubt about that, but it's outdated, it doesn't apply to what's

00054
being done today, and in the kindest terms, we feel that it may be confusing to local carriers.

In terms of standards, the point that we would like to make is that drug resistance, in vitro drug response testing is a laboratory test. We are not marketing a product like a drug that goes into the human body and affects both normal cells and malignant cells. We are dealing with information, and the criteria by which in vitro drug response tests should be measured are the same criteria against which other diagnostic tests should be measured.

There is a question four, which relates to, should payment be dictated by the results of drug resistance testing? How we answer this is to look at what happens in the real world. And we're dealing with information with a diagnostic test. The fact is that this information is only one of many pieces of information which a physician at the bedside has to integrate together to determine what is in the best interest of this patient. And it's laboratory information, it's clinical information, and it's a multitude of human factors, all determined at the site, that should determine the applicability
00055
of this technology. In that case, we don't think that drug resistance information should replace clinical trials, it should only supplement it. We don't think that it should dictate treatment, but it should be one of several factors that are integrated into the treatment decision.

And finally, we feel that it should not be used to dictate payment. I can't think of any single issue that would arouse or marshal together the opposition of the oncology community than the thought that a test is going to determine what they have to do at the bedside, singly and in and of itself.

So that brings us to the main focus of the meeting today, and that is the fundamental question: Can you take malignant cells from a

patient into an in vitro environment, test them in a controlled laboratory assay, identify either resistance or sensitivity, and then translate that into usable information that can be helpful to the clinician when he is at the bedside. That's the fundamental question. And in order to help answer that question, you will find today that there are a number of leading physicians and scientists here to give you their thoughts, their

00056

views and their interpretation. They will focus on evidence, clinical application and they will focus on the patient benefit. When you're listening to these individuals, recognize that without exception, they have spent 20 to 25 years of their lives dealing with these technologies. They bring a unique perspective.

They just don't know a technology; they know an evolution of multiple technologies. In terms of the literature, they don't know an article; they have read and studied and been able to integrate all of the articles together and created a body of knowledge. And finally, the one thing that should be evident is that the people that are involved in this industry are not just scientists developing a laboratory test; they are clinicians. And they have a perspective to see how you can take laboratory data, input clinical decisions, and over a long period of time they have witnessed the patient benefit.

Thank you very much.

DR. FERGUSON: Thank you. I guess you have organized this session, so the next speaker?

MR. KIESNER: Dr. Weisenthal.

00057

DR. WEISENTHAL: Before I get started --

MS. TILLMAN: Dr. Weisenthal, excuse me just a moment. We request that all the speakers that are going to come up just make a statement as to whether you're here on your own behalf or who is sponsoring your trip.

DR. WEISENTHAL: I am here on my own behalf, I bought my own plane ticket and am paying for my own plane ticket. Can I ask, is Mr. Randy Stein here? Mr. Randy Stein? I didn't see Randy. He was a patient who was going to follow me. Is Dr. William Grace here? Dr. Grace, hi.

DR. GRACE: Good to see you.

DR. WEISENTHAL: Frank mentioned some of us having 25 years experience in this field. My experience began in the year 1969 when I started doing cell culture drug resistance testing on human tumor specimens while a graduate student at the University of Michigan. My career really began in earnest when I started doing this in the fall of 1978 while I was a clinical associate in the medicine branch at the National Cancer Institute. And ever since July 1st, 1979, this has really been my full-time job.

For the first eight years, between 1979 and 1987, I did this on a research basis as an associate professor at the University of California, Irvine. Since 1987 100 percent of my time, full time has been spent providing this as a service to patients and physicians in the community. I have about 25 minutes to discuss my life's work. That's not a lot of time, and there's so much, you know, that could be said, and should be said. I will just have to try to do the best I can, I guess.

In the beginning, though, I wanted to put everything in context, and you're going to hear from the following speakers admonitions about scientific rigor and levels of evidence and things like this. I think that you have to put this in context. Mr. Kiesner mentioned that we are talking about a laboratory test. We're not talking about a treatment, we're talking about a laboratory test. If you look at analogous laboratory tests such as bacterial culture and sensitivity testing, there is much less direct data indicating correlations between the

laboratory tests and the clinical response, and there certainly is a lack of data indicating that it makes an impact on patient care, whether you use the test or not.

After 20 years in medical oncology, there's still a debate, should you treat with empiric antibiotic therapy or should you really go to great lengths to try to identify the organism and do sensitivity studies. More than half of the chemotherapy that's given in this country is given for non-FDA approved indications. These are off label indications. In many cases of situations in which Medicare routinely pays for therapy, all that can be pointed to is one or two small pilot studies. In many cases, many oncologists choose drugs on the basis of an abstract that they heard at the American Society of Clinical Oncology.

Two weeks ago I talked to Dr. Robert Livingston, who is a professor at the University of Washington, and probably one of the top five experts in the world on chemotherapy and lung cancer. He's very active in the Southwest Oncology Group. We have an explosion of cancer drugs that have been approved in the last five to ten years. We've got Docetaxel, vinorelbine, Gemcitabine, Irinotecan, et cetera. I dare say that the most common regimens used to treat patients today are regimens made up of newer drugs. Carboplatin plus Taxol; Docetaxel plus Carboplatin; Gemcitabine; vinorelbine platin, and so forth. According to Dr. Livingston, firstly, there's no data that none of these are any better than platinum Etoposide, two drugs which are both off patent, much cheaper, outpatient therapy, and personally in his own opinion, there isn't. You know, he doesn't believe that regimens like platinum Taxol are superior to platinum etoposide.

And yet, this is the reality today, and that is that the vast majority of individual

patients are being treated with treatments that have never been approved by the FDA and are based on levels of evidence that are very preliminary, and that is a fact. And I would like you to keep that in mind when you are looking at the levels of evidence that I am going to be presenting here in the following speakers. I am going to turn this on; okay. I suppose, if I talk loudly, can I go off? I've got to talk on the mike? Too bad.

I want to tell you a little bit about the technologies. What we have done here is, this is a Petri dish with some liquid media and this was a patient's stomach cancer. And this has been chopped with scissors into small little pieces about a half a millimeter to a millimeter. Now there are lots of different technologies, but I'm going to try to show you that the technologies have a lot more in common than they have that separate them.

One of the differences between the technologies is that some investigators will stop at this point. They will cute the specimen into pieces a half millimeter or so, and they will plate that in plastic dishes with liquid media, and expose them to drugs and then determine drug effect. In other cases, patients will take -- investigators or laboratories will take this and pass it through wire mesh screens to give you smaller pieces. And finally, what most laboratories do, is they take the fine pieces and they further digest them with collagenase to break down the tissue matrix to liberate small clusters of tumors.

Now when you do that, and what we've done here is that we've taken the stomach cancer, the same one I showed you on the slide previously, and it's been ingested with collagenase, and we've spun this down on a cytospin slide, and we've stained it with a stain that stains dead tissue and dead cells green, and

living tissue pink. And you can see here, clusters of pink tumor cells amid a background debris of dead tissue, and there will be single inflammatory cells such as macrophages. Well, applying various methods, you can get very nice enrichment of, you can get rid of all the chaff and get down to the wheat, and what you're left with is microclusters. So one difference between technologies is that some use what I call macroclusters, that is, visible tumor pieces, others digest them down to smaller quantities to give you microclusters.

This explains why sometimes the drug concentrations used in the assays are a little bit different. If you've got a large piece of tumor, the drug doesn't penetrate into it very well. And assays that use little pieces of tumors tend to use higher drug concentrations than if you break it down to the smaller cluster

00063

level. So -- but in both cases, you're dealing with a similar situation; you're dealing with a tumor and you're testing it in a three dimensional form. And this is very important. So we're testing three dimensional microclusters of cells, other laboratories might test three dimensional macroclusters.

This is the same tumor now, the stomach cancer, and it has been cultured for four days in the absence of any drugs. This would be a normal saline control. And this is a drug which was only partially effective, so you've got some reduction in the number of cells. Again, some of the dead cells stained green rather than staining pink.

A somewhat more effective drug is this one, and now it has mostly been killed, and you only have a few small clusters of viable tumors left. And a drug that killed everything would give you this, so you'd get the absence of the pink clusters. And the way that this particular end point is scored is manually. Yesterday -- I had to take the red eye last night, because

yesterday I spent ten hours counting laboratory assays. It takes me about three hours of my own

00064

time to do each and every one of our assays.

It's -- I consider the morphologic end point that I just showed you really the gold standard. I like the standard. I like to see the tumor cells on the slide. I like to know whether the drug has worked or not. This technique, though, has some drawbacks. First of all, it takes a lot of time. There aren't very many board certified medical oncologists that are willing to spend three hours looking through a microscope on an individual assay. It's also subjective.

Now, that led investigators to try to come up with easier end points, end points that were not subjective and were automated. So what we're talking about here, if you go back two slides, here we are looking at living cell and then with drugs, either cell death assays, and the main assays that I'm going to be talking about here are cell death assays. Later on, Dr. Fruehauf and Dr. Kern are going to talk about cell proliferation assays. But I think you can take all assays and kind of divide them down the middle, it's kind of like the animal kingdom and the plant kingdom, but you've got assays based on

00065

the cell proliferation end point, assays based on the cell death end point, and I am going to talk about the cell death assays.

Now there's many ways of detecting the death of a tumor cell. This should not be of concern to you. As a clinician, there are many ways of detecting the death of a patient. You can go up and feel for the carotid pulse or the radial pulse. You can put your stethoscope on the chest and osculate for heart sounds, you can observe for spontaneous respirations. You can see if the pupils fixed and dilated. You can take an electroencephalogram, you can measure core body temperature. All of these are methods

for determining, is the patient living or dead.
Likewise, at the cellular level, there are many ways of determining is the cell living or dead. There is more than one way to skin a cat. So for example, you can look at the morphology of the cell and say has it been killed, has it undergone apoptotic death. You can say, has it lost its ATP. When cells lose their viability, they lose their ATP very rapidly. When cells die, they lose their Krebs cycle reductase activity. So there's one of

00066

these assays, the MTT assay, that measures the Krebs cycle enzyme, so when the cell dies, it loses that enzyme activity. And then there's another assay called the fluorescein microculture assay or the fluorescent cytoprin assay, both are really the same thing, and what they're doing there is measuring the membrane integrity with a dye called fluorescein, which is cleaved by membrane esterase and gets trapped in the cell if it has an intact membrane. But the point is here, there are many different ways of determining cell death, just -- there are other ways of determining other things too.

Estrogen receptor. Most of the literature which validates the estrogen receptor was based on wet lab assay procedures, but that's been replaced as you know with immunohistochemistry, and at the beginning there really weren't any clinical correlations, but they showed that basically the immunohistochemistry correlated with the wet lab procedures. But these are all methods for detecting cell death.

Now, these assays -- that's important, because the assays are very difficult to do in

00067

the sense that, I mean, actually generating the data that's going to be presented is an enormous amount of work, and I would love it if I had myself done large numbers of prospective randomized trials in huge numbers of patients, to

show beyond the shadow of a doubt that patients did better when treated on assay results. Goodness knows, I tried, and I and several other people made major efforts. I won't give you the anecdotes of the various trials that never got underway, or got underway and were well funded but didn't accrue patients and so forth.

But suffice it to say, this is difficult work; if it wasn't difficult work, after 20 years of full time in it with people like me and Dr. Bosanquet and Dr. Kern, and others, Dr. Salmon, Dr. Von Hoff, who is certainly one of the most energetic organizers of clinical trials, even he was unable to successfully complete a single study. So this is very difficult, so it's important to look at all of the evidence. So that's why I am going to try to make a point that you need to lump together these various cell death end points.

Basically the assays are done in the

00068

same fashion. You take the tumor, you culture it for four to five days; you expose it to drug, and then you determine, are the cells living or dead. And the fact that there is different ways of determining is the cell living or dead is not of importance.

This is another assay here. This is the MTT assay, and this is based on mitochondrial succinate dehydrogenase activity. Living tumor cells will produce a lot of pink reagent and if they have been killed they don't produce that reagent, so this is a positive control. These are ineffective drugs, this is a single effective drug.

And what we do is since there's advantages -- the advantages of the DiSC assay, which is the microscope assay is that to me, it's the gold standard. You're actually looking at the tumor, you're seeing whether the drugs really work. The disadvantage is that it's a subjective test and it's labor intensive.

The advantage to the MTT assay is that

it's objective, you get a nice machine readout, but it's not specific for tumor cells. If you have some normal cells in there, it can skew the

00069

results, so it's very important that you take a lot of efforts to make sure that you've got a population of cells.

In practice, we do both end points. These cells here, we've had some fast green dye added to them, they're going to be spun down on cytospin slides. These are the same drugs tested in the MTT assay. So we run all these assays in parallel; we always do an MTT and a DiSC, microscope assay, and by doing that, I think I've got a good handle on what's happening.

Now, these end points correlate very well together. This allows us to lump together the results for analysis. This shows 775 solid tumor specimens tested to Cisplatin, and on the Y axis is the MTT assay result, on the X axis is the DiSC assay result, and you can see that in cases where we've got pure tumor preparations, there's a very good correlation between the two end points.

There have been many papers published in the literature. These are just -- I know it's difficult to read, but these are papers comparing the two end points, DiSC and MTT, fluorescein diacetate and DiSC, MTT and fluorescein

00070

diacetate, DiSC and ATP, and all of these end points for cell death, not surprisingly, correlate with each other very well.

How is this information used in the real world? Well, what is done is this, and that is that in the beginning, 20 years ago people had the idea that what they were trying to create was a scale model of chemotherapy in the laboratory. And so they tried to use what are known as clinically achievable drug concentrations. And Dr. Alberts, who's a speaker here, is a real pioneer there, and Dave did a lot of work in the late '70s figuring out exactly what the

clinically achievable levels of different drugs were, and he created some tables that I and other investigators used initially.

I will tell you, though, that if you read the literature today, that's not what people do. Here's what they do, and that is that you get a drug and you do some training set studies, but you try to find the concentration that gives you the widest scatter of results. So on this slide here, what I'm showing is a thousand randomly selected fresh tumor MTT assays for Cisplatin. And this is percent of control cell

00071

survival. 100 percent cell survival means the drug didn't work, the cells are all alive; zero percent cell survival means the drug did work, the cells are all dead. And what you can see is that in a thousand randomly selected solid tumor assays, there is a widespread scatter of results.

So in fact, you try to choose the drug concentration which gives you the greatest standard deviation. You choose a concentration with an index concentration which gives you the greatest scatter. You can then draw the line down the middle for analysis. And operationally you say if the cells are killed in the culture dish, that's resist -- they're sensitive to the drug. If they are not killed, they are resistant to the drug.

Now in practice, you can see that there's a lot of grouping around the middle, so obviously what we do if they are around the middle, that is, if they are plus or minus a half standard deviation from the median, we just say it's in the median and we really can't tell you anything about it, about it. But if it's down here, it's clearly sensitive; if it's up here,

00072

it's clearly resistant.

Now 20 years ago when we first started doing this, we formulated a hypothesis, and our hypothesis was that if you used this method and

you obtained a broad scatter of results, that on average, patients with resistant assays would do worse than patients with sensitive assays. That was the hypothesis. And that is really what I call the central hypothesis to all of this testing, and that is, the central hypothesis is the drugs testing in the sensitive range will be more likely to work than drugs testing in the resistant range. 20 years ago, just a hypothesis. What did the data show?

Well, in the 20 years since then, there have been many papers published, now in excess of 40 papers showing correlations with cell death assays and results of chemotherapy in the patient. For purposes of this slide, I have arranged them in order of increasing response rates in the overall patient population, so this white dashed line shows the response rates in a given study for all the patients in the study. Each of the vertical lines represents a different study. So in this slide, I think I'm showing, if

00073

I can read it, 36 studies, or 35 studies, totaling 1603 patients. But what you can see is that rating from low response rate tumors to high response rate tumors, and this would be something like previously treated phalangeal carcinoma, and this would be acute lymphoblastic leukemia, but in 35 out of 35 studies, in every single case the hypothesis has been confirmed. In fact, patients who are sensitive in the assay do better than the group as a whole. Patients that are resistant in the assay do worse than the group as a whole. And patients that are sensitive in the assay do dramatically better than patients that are resistant in the assay.

So in other words, this assay is an excellent prognostic factor for prognosis if treated with chemotherapy. If you are treated with chemotherapy and the test is in the sensitive range, you do better than average. If you're treated with the drugs in the resistant range, you're worse than average. In solid

tumors, the advantage to getting an assay sensitive drug over an assay resistant drug is a nine to one advantage, patients are nine times more likely to benefit if they're sensitive in

00074

the assay than if they're resistant.

People say that this field is controversial. This is not controversial. These data are unchallenged. There has never been a single study of these technologies in modern history which has failed to show this. If you break it down by tumor types, and here I'm sorry, I can't read it, my contacts -- I took the red eye last night and my contacts are a little bit not clean, but what I have done here is broken this down by disease type and it includes things, stomach cancer, breast cancer, ovarian cancer, non-small cell lung cancer, multiple myeloma, chronic lymphocytic leukemia, acute nonlymphocytic leukemia, acute lymphoblastic leukemia, and so forth and so on. But again, if you break this down by tumor types, patients that have sensitive assays do better, patients that have resistant assays do worse.

So hypothesis I would put to you is not an extraordinary hypothesis, it's a very ordinary hypothesis, and yet, this is an extraordinary level of proof. The hypothesis holds. These data are unchallenged. No one has ever shown anything to the contrary.

00075

You can use the technique that was described in the New England Journal of Medicine a few years back, of cumulative Meta analysis, and I don't have time to explain this, it's in my handout, but basically when you do this, these are 95 percent confidence limits, and what you see is that if you had a P of 10 to the minus eighth, patients that are sensitive in the assay do better that the group as a whole. At P 10 to the minus eighth patients that are resistant in the assay do worse than the group as a whole, and these are again, thoroughly consistent.

Now receiver operator curve plots and Bayes' Theorem. Receiver operator curve plots are, receiver operator plots are used as an assessment of laboratory tests. To generate receiver operator plots, one needs to know how changing the cutoff lines affects sensitivity and specificity. However, with the literature validating cell culture drug resistance testing, what are available instead is the sensitivity and specificity of a single cutoff line, which is around the median. And also, the test accuracy at different pretest response probabilities. The broad applicability of test results to different

00076

disease states may be evaluated by comparing calculated base predictions to actual observations. This is described in detail in my handout, if I go a little rapidly.

Most studies do not show this type of data. This is a single study by Wilbur in non-small cell lung cancers published some years back, but basically it was showing that when you change the cutoff of the assay from 90 percent survival to 80 percent, 70 percent, 60 percent, what happens is the actual sensitivity and specificity of the assay changes, as you would expect, but in all cases, people with sensitive assays are more likely to respond than patients with resistant assays. So this is -- these are not an artifact of just drawing, you know, picking a cutoff.

By applying Bayes' Theorem, you can generate the following theoretical curve. These tests in aggregate, if you add up all the 2,000 or so clinical correlations that have been published, they have an overall specificity for drug resistance of .92, an overall sensitivity for drug resistance of .72. And if you do that, you get these sorts of predictions, and this

00077

shows the relationship between pretest response probability, expected response probability, and then response probability given a different test

result. So the blue line shows what it would be predicted for patients with a sensitive assay.

So in other words, let's take colon cancer as an example. Untreated colon cancer's got a 20 percent chance of responding to 5 FU. If it's assay sensitive, the prediction says that the patient has a 40 percent chance of responding. If it's resistant, it goes down to about 2 percent. Contrary-wise, if you're dealing with untreated ovarian cancer, which has a 75 percent response rate, if you're sensitive in the assay, it goes up close to 90 percent and if you're resistant, it falls down to about 15 to 18 percent. So those are just the theoretical predictions.

How about if you break this down by individual types of tumors? And I think that these data compelling showed that these assays are broadly applicable for really all types of tumors in which they've been studied, both solid tumors and hematologics, ranging from stomach cancer, colon cancer, non-small cell lung cancer,

00078

ovarian cancer, breast cancer, chronic lymphocytic leukemia, acute lymphoblastic leukemia, acute nonlymphocytic leukemia, in all cases it holds exactly according to base predictions.

There are many correlations published in the literature about patient survival, and these are the patients that, survival of patients sensitive in the assay, survival resistant in the assay. I give references in my handout, and these will be discussed by other speakers.

Now I have to, I'm already bumping up against my time limit, but I have to discuss a group of papers which are very, very important, because you as panelists have probably spent the most attention to these, because these were studies done at the National Cancer Institute, published in prestigious journals, and so naturally you think that these are really quite important papers. There was a review by Cortazar

and Johnson in The Journal of Clinical Oncology. What this review showed was that there were three non-randomized small studies which showed nonsignificant inferior survival with assay directed therapy, compared to control therapy.

Again, these were non-randomized studies and the results were nonsignificant but still, three of them showed, suggested slightly inferior survival with assay directed therapy compared to control therapy. What's important for you as panelists to realize is that one none of these studies, not a single one, evaluated the fresh tumor assays which are used in the real world and have been used for the past 12 years, and which are now being considered for reimbursement. This whole paper is utterly irrelevant because it does not review the technologies that are under consideration here.

Specifically, I want to take you through the three NCI studies. The NCI did a study in non-small cell lung cancer, they did a study in extensive disease small cell lung cancer and limited disease small cell lung cancer. In general, the non-small cell study was highly negative, highly negative. There's no one that could read that paper that could possibly conclude that this particular assay was of any utility whatsoever. It's a totally negative study. The extensive disease small cell study was modestly positive. The limited disease study

was highly positive, and I'd like to tell you why that is.

First of all, the limited non-small cell study. This is probably the most important one for you to consider; this gets quoted the most. In 1994, when I presented this at California Blue Shield, I had to spend half of my time debunking this one paper, because somebody at the University of California San Francisco brought it up. For the past seven years I've had people come up to me over and over and they say,

well, they tried that at the NCI, it didn't work, they're the mecca of meccas, if they couldn't get it to work, what makes you think you can get it to work? Well, I've been doing this full time for 20 years and if you work at something very hard, you can actually get it to work.

But let's talk about this. Non-small cell lung cancer study from the NCI. Firstly, they used passage cells. These were not fresh tumor assays. It said in the study methods that they were fresh tumor assays, they were not. That's an incorrect statement. This paper was not written by an investigator associated with the study. This was written by an investigator

00081

named Gail Shaw, who at the time was an oncology fellow. She rotated through the NCI Navy branch. This was after the investigators, after Dr. Meadows' group already left. She went and did these chart reviews, she wrote the paper, and she incorrectly stated that these were done on fresh tumors.

In fact, I called Audie Gasner at the University of Texas, and he confirmed that every single one of these studies were done on passage cells. That's what they were trying to do, they were trying to see, could they use cell lines to do assays. So these are not fresh tumor assays, these are on cell lines.

Why is that important? Well, because papers have shown if you generate cell lines, that with subsequent passages, that the drug resistance changes. And this has been well shown in the literature.

Secondly, these were monolayer cultures. These were not -- they were not testing three dimensional cultures of clumps of cells, clusters of cells, they were testing monolayers. And in a study, seminal study in PNAS, 1993, Tyker and Kerbil, they showed that if

00082

you do monolayer cultures, that that doesn't correlate, but that when you do three dimensional

cultures, it does. These were monolayer cultures. No one does monolayer cultures. In this study they did.

They had a 22 percent overall evaluability rate, and 7 percent with lung primaries. In 1985 I did a study in conjunction with the Loma Linda VA, Dr. Dave Wilbur, in which they just sent us by regular mail specimens of non-small cell lung cancer. So this would take two to three days to arrive in the mail, and this was with technology in 1985. We had an overall evaluability rate of 75 percent. Today -- I reviewed my data last night before coming here, and in the past five years, we have received 347 non-small lung cancer specimens, and 326 of those assays were evaluable, which is a 93 percent overall evaluability rate. Twenty of those had a negative histology, and that's a reason for inevaluability, so if you only look at cancers that actually had cancer when it made it to our lab, we had a 97 percent evaluability rate, including a 96 evaluability rate with 124 primary lung tumors.

00083

So these guys are testing a subset of patients, 22 percent, and 7 percent with lung primaries. So what do we know about that subset? Well, it turns out that they had previously shown in Annals of Internal Medicine that when they got a tumor that they were able to subculture, that just the fact that the cells could be subcultured was a powerful negative prognostic factor. And they said that this is a marker, in the Annals of Internal Medicine paper, for biologic aggressiveness. So think about it. The only people getting assay directed therapy are the people with the worse prognostic group, with the biologic aggressive group, and they're being compared with a group of patients that you can't subculture, and they have the biologically indolent group.

And finally, and last but not least, they did not give assay directed therapy until

the fifth treatment cycle. They biopsied the patient, it took them four treatment cycles to get these cell lines going to test them, and so they didn't actually get the assay directed treatment until five treatment cycles.

This paper has been thrown up in my face again and again and again at the best universities in the country, and this paper is a bunch of rubbish. It should never have been published. It is misleading, and it's terrible that it keeps resurfacing. And I hope that the previous speakers, or speakers that follow me just don't -- this paper is irrelevant, let's not waste any more of our time about it.

Now, the small cell lung cancer study in extensive disease, this was modestly positive. Why was that? Well, they still used passage cells; that was bad. This paper in the International Journal of Cancer again showed that if you used passage cells in small cell lung cancer, that doesn't correlate. However, small cell growth is three dimensional spheroid cultures, unlike non-small. That's good. They had a 55 percent assay evaluability rate. That's good too. They're not dealing with a selected population. Assay directed patients were a similar prognostic group relative to control patients. So all these were good. And the only thing that was bad is that they weren't giving assay directed therapy until the fifth cycle.

Now I mentioned that the results of this study in limited disease were positive. In fact, the patients who got assay directed therapy lived a median of 38 months and those who got standard therapy lived a median of 16 months. This was statistically significant. It's a small study, but it was statistically significant.

So why was this study positive and the other one wasn't positive? For perfectly explainable reasons. They're not treating a bad prognostic group. They're using three

dimensional cultures. I would say they would have had even better results if they had used the assay chosen drug up front, but they didn't.

And in the extensive stage non-small cell study, this was less positive than in the limited stage study; why was that? Well, it turns out that the assay directed were also a worse prognostic group. In the extensive disease study, patients could only get assayed if they had peripheral lesions for biopsy under local anesthesia. They did not do general anesthesia for this. And they showed that just, when they analyzed the patients that had biopsiable tumors versus patients that didn't have biopsiable tumors, there was a significantly shorter

00086

survival if you had a biopsiable tumor. That makes sense. They've got more extensive disease; of course they're going to die faster. So if the only people getting assay directed therapy are people that have a higher tumor burden, that's really biasing it against it. And also, they're not getting treated until the fifth cycle.

In summary, these NCI studies have nothing to do with the real world. They don't apply to the technologies that you're evaluating. They are utterly irrelevant.

There have been many studies showing correlations with patient survival. I wish that I had time to take you through these and show you the survival curves. In particular, one group of studies is not going to get presented here, and I just want to tell you very briefly about it. And that is studies by Vierman's group in acute lymphoblastic leukemia. There is some controversy. Should you include pediatric leukemia in a discussion about assays applicable to Medicare patients. I think that you have to, because it makes a consistent story.

If you look at the data that validate these technologies, just to pick one type of

00087

disease, human lymphatic neoplasms, AOL and COL.

In 1962 an investigator named Schreck showed in Annals of Clinical -- Journal of Clinical Investigation, that if you did an apoptosis assay on fresh cultures of COL, that radiation response in that assay correlated with patient survival. That work was lost. Nobody knew about apoptosis in the '60s, nobody cared about it. In the early '60s -- in the early '70s, people that were working on assays and leukemia had the idea that you had to do clonogenic assays. In fact, those of you who are familiar with the literature, they would look at clonogenic assays, actually do clones of clones. They would plate single cells, let them grow two to three weeks until you had a clone, remove the clone, desegregate it and reclone it, so you had this very cumbersome assay that would take about six weeks, and people thought that you had to do that, because there was this phenomenon of the stem cell, and the only thing that's relevant for chemosensitivity is the stem cell.

I came up with this really radical idea, based on Dr. Schreck's work, which I was familiar with, that if you just expose the cells

00088

to the drug and divide them into groups, and one's above average and one's below average, that that will be a strong correlation with clinical response. I published several papers on these in the early 1980s. And what I showed in one of the papers, for example, is that if you looked at assays on both COL and AOL, pediatric AOL, adult COL, that there was strong correlations with clinical response. Furthermore, if you looked at previously treated patients, they were much more resistant, significantly more resistant than untreated patients. And finally, if you I followed individual cases of patients over time, if they were assayed multiple times with no intervening chemotherapy, there was no change in the assay results, but if they had intervening chemotherapy, they became demonstrably more resistant in the assay. As a result of these

papers, there were other investigators that got into the field.

DR. FERGUSON: Dr. Weisenthal, there are three other people apparently that are supposed to talk, if I extend the time to 10:15, which I said I would, and so perhaps you could --

DR. WEISENTHAL: Well, Mr. Stein is not

00089

here today.

MR. STEIN: I'm here.

DR. FERGUSON: So that's four other people. So maybe, if you could wind up?

DR. WEISENTHAL: Okay. It's very frustrating. You know, I've got some really good stuff to tell you.

DR. GRACE: I'll donate my time to Dr. Weisenthal. Will that help?

DR. FERGUSON: That will help some.

DR. WEISENTHAL: I think I can finish up in five minutes, can I -- okay.

I was going to summarize the data in human lymphatic neoplasms. So basically, you know, that's what I showed. Dr. Bosanquet has for the last 18 years been studying these assay systems in chronic lymphocytic leukemia, and I'll let his work speak for itself.

The work that can't speak for itself is a parallel work that was done in acute lymphoblastic leukemia, and this work with the MTTS, and what these investigators did at the Free University of Amsterdam, first they started using the DiSC assays, just as I described. They said it's a lot of work, you've got to count with

00090

the microscope, so they preferred using the MTT assay, but they used exactly the same culture conditions, 96 hour culture, same exact identical conditions. What they showed in a series of very rigorous trials, published in excellent journals, several publications in Blood, publications in the Lancet, these are superb studies, and I am -- they should not be excluded from this consideration. They showed in very rigorous

studies strong correlations between the assay result and patient survival.

In fact, the assay results were the strongest predictive factor, and it turns out they were the only independent predictive factor.

And all these other cell marker studies that people do on pediatric AOL were not significant once you consider the cell culture results.

So if you take and look at historically, the correlations with response, the correlations with treatment status, and then if you look at Dr. Bosanquet's excellent studies in chronic lymphocytic leukemia, and you also consider very identical studies in acute lymphoblastic leukemia, it is a continuous consistent whole.

Now the last thing I want to briefly address is the issue of drug synergy. Should you test single agents, combinations? What this data are showing is that most drug combinations in human solid tumors are not synergistic. There is very little, if any, evidence of clinical synergy in clinical data treating human solid tumors. Combinations, citoxin plus adreomyecin is never synergistic. Taxol and platinum is not synergistic, it's additive. These data show, it says platin and etoposide -- these are results where it says platin alone, etoposide alone, this is what you would expect if they were additive, and in fact they are additive. Now there are occasional combinations which are uniquely synergistic. One of them is gemcitabine plus cisplatin, which is one of the most exciting new combinations to come along in a long time. In contrast, this is a highly synergistic combination, and so when you've got a synergistic combination, it make sense to test the drugs in combination.

I am going to show -- this is a patient that Dr. Nalick is going to be presenting. This is an ovarian cancer patient, control culture.

Carboplatin alone had a minimal effect. Gemcitabine alone had a minimal effect. Combination of the two wiped everything out. Highly synergistic.

The next speaker you're going to hear is Mr. Stein, and he was the subject of this paper in Scientific America last February, and he's going to tell you his story, but I'm going to show you his assay. These are his control cultures, this is pancreatic cancer. Control culture. Platinum alone, modest effect; Gemcitabine alone, modest effect; platinum Gemcitabine, wiped out.

I'll stop here. Anyway, and other speakers I'm sure will amplify the remarks that I said.

I'd now like to introduce Mr. Randy Stein.

MR. STEIN: I want to thank all the distinguished members of this advisory committee for listening to my testimony. I also want to publicly state that I have no financial interests or involvement with any manufacturers of any products being discussed or with their competitors. I would also like to inform you

00093

that I feel the importance of my being here to testify is of such magnitude that I flew here from Southern California and postponed a trip to Acapulco. I plan on joining my wife, who went as previously scheduled and is awaiting my arrival directly after my testimony.

I truly feel I had no choice but to testify though, because I realize that without the cell culture drug resistance testing, I would be dead. And to think that other people will die if drug resistance testing is not approved, while I sit and bask in the sun is totally unacceptable to me. You see, I was blessed with an incredible gift, the gift of life, and now everything I do is about giving back, touching as many lives as possible and trying to make a difference in the cancer community.

I was diagnosed with four stage non-operable pancreatic cancer that had metastasized to my spleen and kidneys on January 22nd of 1997. My CA-19 tumor markers, a blood test used to determine the severity of the disease, were at 12,930, with normal being 0 to 37. My gastroendocrinologist sent me to a local private practice oncologist named Dr. Stuart

00094

Nagasawa. Dr. Nagasawa explained to my wife and I what I had, the grim statistics associated with four stage pancreatic cancer, and then told us if he were to treat me, he would like to get aggressive. He went on to explain to us that getting aggressive meant having a laparoscopy done, taking a tissue sample from one of my many tumors, and sending it to the Weisenthal Cancer Group.

He then told us the principles behind sensitivity testing and why that was so aggressive. He also explained that with such a fast growing cancer, my chances for recovery would be better by knowing which chemotherapy was the most effective, and even more importantly, which was the least effective on my tumor. He also explained that with conventional treatment, we would not know the effectiveness of any chemotherapy on my tumor for three months, and at this time that was my expected life span. This seemed like a no brainer to me, no now or maybe blow my chances on a chemotherapy that wasn't effective.

Being a little on the snobbish side and realizing the probable outcome of my disease, we

00095

wanted to talk to some of the top doctors in California. Second, third, fourth and fifth opinions. To tell the truth, I was sorry we didn't stop at the first. We went to the most famous and most prestigious facilities available, UCLA, USC, City of Hope, and the John Wayne Cancer Center. We were told go fishing if that's what you like. We were told, maybe three

months. We were told, I don't know why you're still alive. We were kept waiting by one of the grand gurus of cancer for over an hour, told yep, it's pancreatic cancer, you have three months to live, and I have to leave; I'm late for a root canal appointment.

We discussed sensitivity testing with each and every one of them and their collective reactions were all the same. They were against relying on the cell culture drug resistance test. Their reasoning was simple, the test tube doesn't exactly duplicate conditions in the body, and the testing may prejudice the doctor's choices. These same doctors were also sure I would be dead two and a half years ago.

After much deliberation with friends and family, we decided that although testing in

00096

the test tube may be different than the human body, it did give us a better idea of what the tumor was resistant to and what might work. And let me tell you, when you have three months to live, that sounds a lot better than just guessing what chemotherapy to use, or doing what everyone else is doing, with little or no help.

Conventional for four stage inoperable pancreatic cancer patients has a 3 percent chance of prolonging life for more than three months, and 0 percent past one year. And as far as the doctors being prejudiced, I'm a child of the '60s and I hate prejudice, but this type of prejudice seems very reasonable. We made the decision to use Dr. Nagasawa. We felt getting aggressive made more sense than waiting to die.

The current FDA approved treatment for pancreatic cancer with metastases is Gemzar, and had I treated within the FDA guideline, I would be dead. When we received the results of the drug resistance testing, Gemzar alone scored very poorly. But the combination of Gemzar, when combined with Cisplatin reacted very favorably on my tumor sample. Dr. Nagasawa explained to us that although by themselves, the Gemzar and

00097

Cisplatin scored poorly, together there was an incredible synergy, meaning one plus one did not equal two, it equaled ten. And without the testing, we would have never known it.

He then told us I would be on this combination of chemotherapies until further notice. Three months later he took another CA-19 tumor marker test. The results came back at 8,970, down by approximately 30 percent. That gave us hope, and hope is crucial to anyone's survival. After nine months, in September of 1997, my tumor markers came down to 6,300, and my chemotherapy was changed to every two weeks. The following year showed a continuous decline in my CA-19 tumor markers. I started to gain back the 50 pounds I had lost, and I was able to discontinue the use of all pain medications.

Eighteen months after diagnosis, on June 17th, 1998, my doctor called and said I've got good news. I just received your latest CA-19 tumor markers and they came back at 31.4, well within the normal range, congratulations. I looked at my wife and she looked at me. The tears started rolling off both of our faces. I then replied, congratulations to you, Doctor.

00098

The rest of that night was spent celebrating. We picked up chili cheese dogs and Dom Perignon, called all the wonderful friends and family that were so incredibly supportive during this time, and had one incredible evening. The joy of that night will live with me forever.

August of 1998, my chemo was reduced to once every three weeks, giving my wife and I back a life and allowing us the ability to travel, as well as spend time with our loved ones. July of 1999, my doctor arranges a PET scan, and the results come back: No caner anywhere in my body. August of 1999, with our healing professionals, the decision is made: No more chemo.

The NCI pamphlet on pancreatic cancer

on page 10, states that cancer of the pancreas is very hard to control, that the disease can only be cured when it is found in its early stage, before it has spread. Last week we had a huge fund raising function for the pancreatic cancer action network. I had the opportunity to tell my story to the 800 people in attendance. There were doctors in the audience from the top facilities in the country, Johns Hopkins, M.D.

Anderson, Sloan Kettering, and many others. Needless to say, they were all in awe of my recovery and were very anxious to speak personally to my doctor and myself regarding my recovery.

I was lucky. Although my insurance now covers this test, at the time it didn't. Dr. Weisenthal allowed me to pay him over time, and by the grace of God we could afford to do that. Most people on Medicare can't, and they don't need or deserve to die. It is up to us who survive to become activists and to do our best to see that proper actions are taken and that the effective treatment and diagnostic aids are researched and made available in the future. All of this is why I am standing here today, cancer free and begging you to approve this type of testing.

Thank you for your consideration of this very important coverage.

DR. FERGUSON: Thank you very much, Mr. Stein.

Richard Nalick, or are you --

DR. NALICK: I am.

Good morning. My name is Richard Nalick. I'm a gynecologic oncologist at USC School of Medicine, professor there, clinical professor, and for about the last 15 years in private practice in gynecologic oncology in Los Angeles. I have no involvement with any company or any individual or manufacturer of any of these products being discussed today.

I am here today because of my interest and passion in this particular form of testing of chemotherapeutic agents. I finished my training in gynecologic oncology in about 1974, went to Texas for three years at Parkland Hospital, and then back to USC as a professor. And at the very beginning, I was interested in this form of testing. It made the same sense to me as testing a urine for culture insensitivities. If you can do that for bacteria, why not do it for cancer cells? Of course we had less drugs at that time but it was still interesting.

We tried to set up an assay at USC and it was a clonogenic type of assay, and it was fairly good, but we didn't have the finances to really carry that through. I then started sending the assays of tumor to Dr. Von Hoff in San Antonio. The problem there was that most

00101

often, because they had to be sent in dry ice, the tumor tissue usually didn't make it to San Antonio, at least not in good shape, and I had several communications with Dr. Von Hoff and we had some data, but it was somewhat difficult to interpret, and I stopped using that assay.

About that same time a man by the name of John Daniels, who's a medical oncologist and Ph.D. at USC School of Medicine developed his own clonogenic type assay, only this he did at USC, and I started sending tissue to him. I eventually collected at least data on 200 patients of my own, that I obtained initially for use later, if the patient did not respond to primary treatment. However, as time went by I saw more often and more often that the findings in the test correlated with my findings in the patients, so I started using the assay, certainly in patients who failed the treatment, and towards the end started using the assay up front, because I had such confidence in those findings.

However, after accumulating about 200 patients of experience, Dr. Daniels went on to other things and started another company, and his

physicians were referred to Oncotech, which had just started. So I have experience with about 200 patients with the clonogenic assay, and then at least a hundred if not more patients experience with proliferative assays at Oncotech, and that test proved very effective in my practice.

And at this point now, I started using this assay up front. Some people thought that wasn't ethical, it wasn't correct, but with my experience I had at that point was such that I knew that if the test showed extreme drug resistance for that drug, that drug did not work in my experience, so I stopped using it. And I would then pick the drugs left that looked the best, keeping in mind toxicity, and their track record in oncology.

Well, after a period of time, I sent assays to Dr. Weisenthal's group, and I actually compared them with Oncotech and Weisenthal, same patient, two assays, and they correlated fairly well. But in my opinion, the Oncotech assay was certainly excellent for finding drugs that the patient would not respond to very actively. The Weisenthal assay, however, allowed me to test more drugs.

And I also felt relative to what you have seen already and will hear later on, that testing combinations is important also, and actually with Oncotech we tested a few combinations. But since with Dr. Weisenthal's group being utilized, I've tested all patients with single agents and combinations. So I always test for carboplatin, cisplatin, one of the platinums in Taxol, one of the platinums in Topotecan, one with Gemcitabine. Other combinations that aren't usually used, but that have shown benefits in other studies, such as Navelbine and Thiotepa. I will test Doxorubicin and Doxil.

And I can tell you that up front, long

before the papers came out, we knew per that assay that there was synergy between Gemcitabine and platinum, that we didn't know before. And that it had become, as far as I'm concerned, that should be the gold standard today, not Taxol and platinum, in my opinion. We also saw that combinations worked in other situations too; very often, carboplatin and Taxol would be equal to Gemcitabine and Taxol.

But my feeling was that if they were

00104

equal on the assay, and knowing that really platinum was probably the most important drug, and now knowing that there is synergy between Gemcitabine and platinum, and knowing that there is no hair loss with platinum and Gemcitabine, but there is with Taxol, and knowing that there is there is no significant neurotoxicity with carboplatin and Gemcitabine, there is with Cisplatin, but when you combine platinum and Taxol, you have very significant neurotoxicity, and this leads to very important toxicity for the patient in terms of quality of life.

So since the onset of my work in gynecological oncology, probably the most common cancer I deal with is ovarian cancer, it is sensitive to these drugs, but the overall prognosis is still poor, and the five-year survival is till around 38 percent for all stages. And since 70 percent of the cases we see are advanced disease to start with, it's extremely important to be very aggressive surgically and to be very aggressive with chemotherapy up front, at the beginning.

I think it's totally wrong to treat a patient and hold chemotherapy until the patient

00105

metastasizes. I mean, I feel just like Shakespeare said in Hamlet: Diseases desperate grown are by desperate appliance relieved, or not at all. And I think you have to be aggressive from the beginning. So my aggressiveness is based on extensive radical surgery, tumor

reductive surgery down to an optimal level if possible, and then treating the patient with the drugs that have been found on the assay to be the ideal combination, keeping in mind the goal of curing the patient when possible, palliating them always, and minimizing toxicity, which is extremely important, because that's basically the quality of life problems with hair loss, neurotoxicity where they can't walk or pick anything up, and not to mention bone marrow toxicity and so on.

So I pick the safest combination that looks most effective on the assay. I will continue to do that until I quit practice. And I now have, I think probably the largest series that any one physician has in the country. It's between 450 and 500 patients, and I am trying to write that data up. And I know what it's going to show, because I've already seen it in my

00106

patients. So, this is what I feel, and I think this will be proven by further speakers.

Thank you very much.

DR. FERGUSON: Thank you very much.

MS. TILLMAN: Dr. Nalick, did you state whether you were here on your own behalf?

DR. NALICK: Yes. I am. May I show three slides?

DR. FERGUSON: Well, we have William Grace and John Fruehauf in the next five minutes.

DR. NALICK: Okay.

DR. FERGUSON: Or we can just forego the break.

DR. NALICK: I'll just tell you the case without the slides. It will take me two minutes. As an example, I had a patient who was a gynecologic nurse, oncology nurse. She had early ovarian cancer, treated at UCLA, had a hysterectomy, had fairly decent tumor reductive surgery. She received platinum and Taxol. She had persistent disease. She had a second look that showed persistent disease. After a long period of time, she was finally accepted for a

bone marrow transplant. She received that at UCLA, was in the hospital for almost two months

00107

with a bill of over $200,000. Her disease still recurred.

I eventually saw her in 1997, opened her abdomen. She has unresectable disease. She was tested on the assay. She had a bowel obstruction, had a colon resection. She was found to be resistant to every single drug of 27 drugs, except synergy with Gemcitabine and platinum. Dr. Weisenthal showed her slides. She was treated with that combination, later had a third operation, had only microscopic disease. Had radioactive P-32 placed in the abdomen, and then followed with six more courses of Gemcitabine and platinum. She's now totally free of disease. This is a woman who had stage four disease, positive pleural effusions and unresectable abdominal carcinomatosis. She's free of disease, off chemotherapy for a year and a half, and she's still working full time as an oncology nurse.

DR. FERGUSON: Thank you. Ladies and gentlemen, we have approximately 50 minutes for six speakers. And I'm going to take the chair's prerogative at this point to say okay, I would like to just do them serially, and you can take a

00108

break as you need it. But then I'm going to limit everybody to seven minutes apiece, unless this group of people decides otherwise. So, I would like to call William Grace.

DR. GRACE: I'm William Grace. For 24 years I was chief of cancer research and chief of medical oncology at St. Vincent's. I am now full time private practice. I have no financial interest in any of the companies that are represented here, but I have an enormous conflict of interest, because Larry Weisenthal makes me look good. As you know, I am in New York City, where we have the world's greatest cancer centers. And of course as you know, most people

don't fit into clinical trials, and in New York City, probably only one in 20 gets into them. So I see a lot of patients who have a lot of very advanced disease, and one of the things that I have done with my patients, and indeed in my family, who have had cancer, I have used Larry Weisenthal and other members here in order to help manage their care.

What I can tell you is I could have paraded in here a roomful of patients with stories such as we've heard today from patients.

00109

And I am a strong believer in this technology, and it's amazing. I have found that whenever I present this technology to my patient, almost none of them ever resent the cost involved; they find the money somehow. It would be good for some of my Medicaid patients and some of my Medicare patients on fixed incomes if they could have this technology available to them. And I will tell you one thing; it likely wouldn't reduce the cost, because I know many of my colleagues who don't believe in this technology, essentially chemo patients to death with one chemotherapeutic combination after another, not using this, but just going from one ASCO Journal to another ASCO abstract to another, and boy, that adds up to an awful lot of money. And if you could predict that these patients would not respond, you'd save patients a lot of grief and you'd save patients a lot of money.

So, the only thing I'm going to do is turn my remaining four minutes to somebody else, because I am here on my own, I'm here because the assay, one, makes me look good in a very competitive market, and I believe that this technology should be brought, and indeed the

00110

combinations that Larry has given us have been applied to my pancreatic patients. And I have become now a guru, along with Howard Bruckner, in New York for pancreatic cancer, and we are all using Larry's stuff to tell us what combinations

to use, and we're doing some exciting work in pancreatic cancer. You just will not believe the results when you finally see them.

Thank you.

DR. FERGUSON: Thank you very much. Let's see. John Fruehauf?

DR. FRUEHAUF: It's a pleasure to be here today. I am John Fruehauf. I'm a medical oncologist. I have a conflict of interest. I am the medical director for Oncotech.

And I began my career in oncology really as a an M.D. Ph.D. student, and my Ph.D. was in pharmacology, in Chicago. And I studied the effects of BCNU on leukemia, using tritiated thymidine as a measure for measuring cell proliferation. I had learned that technique as a co-step commissioned officer in a program at the NCI where I spent three months in Dr. Herberman's laboratory. And then went on to do my fellowship in medical oncology after residency at the

00111

University of California, and my fellowship was at the NCI. And as a first year fellow, I participated in some of the small cell lung cancer studies that were presented, where I was treating patients based on these assays. So for about the last 19 years, I have been involved in this field. And as a practitioner at UC Irvine, where I treat patients, I can see the value of these technologies.

And I wanted to talk a little bit, briefly go over the historical perspective of where we came from to get to where we are today. And I want to talk a little bit about the clinical guidelines. The speaker for the FDA talked a little bit about how we judge laboratory testing, and I wanted to talk about the levels of evidence that we use in decision making, and summarize clinical data, and then present an algorithm for how we can use this technology clinically.

Now the principles that we employ in cancer treatment are basically that most patient

are not likely to be cured. And so, we focus our efforts on palliation, to prolong life rather than cure patients. And we know there is

00112

morbidity related to our chemotherapy treatment. People have talked about alopecia, neurotoxicities, gastroenteritis, all sorts of risks, and we have to balance these risks of therapy with the benefits of modest life prolongation for most patients.

So when we look at outcomes, we look at survival, we look at improvement, in disease free survival, complete response rates, we look at cost effectiveness, but quality of life is really one of the critical end points as an oncologist who sees the patient sitting down in front of me, how are you feeling today? Do you have neuropathy? Should we change your chemotherapy? Because if you aren't going to cure somebody, your goal is to help them live a quality life until they die. So this testing can be used to do that, by eliminating ineffective therapy.

So, how successful are we in treating medical oncology patients today? There's about a million patients diagnosed a year. 64 percent of these patients present with localized disease. So there's a large bulk of patients, 44 percent can be cured with surgery, 18 percent can be cured with radiation therapy, but a dismal 2.4

00113

percent are cured with chemotherapy. Chemotherapy, unfortunately, has not reached the level of great success at this point in time. And so we're struggling in research to find new drugs and to do a better job.

If people present with metastatic disease, again, 3 percent are cured. 5 to 6 percent will have remissions for two years; 15 percent can have a remission for a year; 76 percent have no, or minimal life prolongation. So the vast bulk of patients we're treating with chemotherapy are not particularly benefitting from this. And so with that backdrop, we can

look at the value of tests to help avoid ineffective therapy, where this is so common.

Now we all take an oath as physicians, the Hippocratic oath, to keep patients from harm and injustice. And I think it is an injustice to treat people with drugs if you can figure out ahead of time, they're really not going to benefit the patient. And I tell my patients, each patient I treat, what are the ground rules that we want to do, use as a team in approaching your disease? And we all agree, we never want to make the treatment worse than the disease. And

00114

of course, giving ineffective therapy that's toxic breaks this rule.

And Dr. DeVita, in the third edition in his textbook, Principles and Practices of Oncology, has stated that the most important reason that people do fail in treatment is drug resistance. So this is why in our laboratory, we focused on drug resistance as an end point.

Now, this all started back in the '70s, the 1870s, when Pasteur looked at chemotherapy, and bacterial cultures. So this has a long history in terms of in vitro testing. But more -- as the technology evolved, the first studies on cancer cultures were published in '54. In '56 it was found, very importantly, that agar could selectively allow you to measure drug effects on cancer. And this was really a breakthrough in terms of then developing the clonogenic assay. I think other speakers have pointed out, and will point out that these basic clonogenic stem cell assays were not effective. There were many problems. Clump artifacts; it would take three weeks to get an answer; you only got an answer half of the time. This was not a good technology.

00115

Advances were made, introduced, in second generation, and then finally, third generation technologies. And as Dr. Weisenthal pointed out now today, we get answers 85 to 90

percent of the time within seven days, which is relevant to having an impact in clinical utility. So I think how long it takes, how often you get an answer, was a great advance over the original clonogenic assays.

So we want to talk about a little bit today about how assays can predict response, which is one end point, but also importantly, do they relate to patient survival? Now, what are the statistical requirements? And I'm just going to say we will show about, talk about sensitivity and specificity, predictive accuracy, reproducibility, cost effectiveness, and that these are the kinds of means you use to validate a home brew test in a laboratory following CLIA guidelines.

Now there are levels of evidence we follow. Level one is metanalysis for prospective studies, and Dr. Weisenthal has shown you a metanalysis of sorts. Although most of the studies were not necessarily randomized or

00116

controlled, it doesn't meet level two and level three requirements, which is evidence obtained from at least one well designed experimental study. And I think critically, that these studies are internally consistent. And level three is evidence obtained from well designed quasi-experimental studies such as non-randomized controlled single group studies.

Because tests are not drugs, you don't really compare Test A to Test B in a prospective randomized study. You compare a test to outcomes. Does the test predict an outcome. So I think that actually, this kind of experimental model fits testing, whether or not a laboratory procedure can predict an outcome in the clinic.

And so, Grade B evidence is what we want to look at here, because that's taking levels of evidence type two and three, or four, and showing that the findings are generally consistent. And I think what Dr. Weisenthal has reinforced is that if you look at all the studies

that have been done in the last 20 years, these studies show very consistently that you can identify ineffective agents.

Now, it's also been -- this is very

00117

difficult to read, I apologize, but over here what this says is that class two and three evidence is used to make decisions in the treatment of breast cancer.

DR. FERGUSON: I'm giving you four more minutes, because you were donated.

DR. FRUEHAUF: We've divided our time up with the other people in our group.

DR. FERGUSON: Oh, you have?

DR. FRUEHAUF: Yes.

DR. FERGUSON: So you will be finished by 11?

DR. FRUEHAUF: I was going to take about 20 minutes; so what time did I start?

DR. FERGUSON: Okay. And so that Orr and Hoffman, and Bosanquet, and David Alberts will finish in --

DR. FRUEHAUF: They will take about ten minutes each.

DR. FERGUSON: Well, it was about eight the way I calculated before, so I mean, there are four after you.

DR. FRUEHAUF: Well, I'm going pretty fast here.

So, I want to talk about the data now,

00118

with that background. The cancer is a really big problem. We can't really cure people very often, and if we can identify drugs that don't work, that can have great value. Did the data show this?

Now here are correlations that we published in the Principles and Practice of Oncology. Dr. DeVita asked Dr. Bosanquet and myself to write a review, you have that in your packet, this table comes from that review, from the textbook. And we looked at 4,263 cases tested with a variety of technologies, and what

we found was that the predictive accuracy of these tests was 90 percent or better for predicting drug resistance, and about 72 percent for predicting response. And the sensitivity and specificity for this technology regardless of the end points is very comparable, 85 and 80 percent.

Now how would that compare to tests we use every day? In fact, the predictive accuracy of drug response assays is comparable to hormone receptor assays, and slightly better than bacterial culture and sensitivity assays. And I would submit, who would want to treat a breast

00119

cancer patient with Tamoxifen before getting an ER or PR result? Or if you have a patient with a refractory bacterial infection who's neutropenic and they're not responding to primary empirical therapy, almost every one of us will get cultures to direct our therapy. So these are commonly used tests that have comparable predictive reliability to in vitro drug response.

Now we did a study of 450 cases. Actually UCLA did a study, which is the validation study for the technology we used in our laboratory at Oncotech. And there were 332 colony end point assays, which at that time was a gold standard, and the newly developed technology was thymidine incorporation, and so there were 118 of these assays that were then compared to the gold standard to make a determination of, can the new technology give us similar results.

And without belaboring the point, tumors are cultured in three dimensions, which is important, as Dr. Weisenthal indicated. They are cultured in drugs for five days, which means that their exposure, in vitro drug exposure or concentration times time, is going to be about five to 20 times higher than you can achieve in a

00120

patient. And at the end of a three-day exposure period, for the last two days to make five days, treated thymidine is added and if the cells are

dividing they will incorporate thymidine and you can measure that with a scintillation counter.

Now these tumors were cut out of patients, shipped across town, put into a culture, chopped into pieces, exposed to five times higher exposures than you can give in patients, and if they grew through that drug exposure, what was the outcome clinically? What we could see is that the colony end point was very comparable to the thymidine end point. In your handouts you can see and probably read more legibly, that if you looked at the overall assay predicted response probabilities, zero people were responding in the assay if they had a below, one standard deviation below the median result. And this zero response rate was true whether they got drug combinations containing different single agents that were tested in the assay, and it was true across different tumor types as well. So there was a robust quality to showing people didn't respond, zero response basically, except for one responder with this cut point of one

00121

standard deviation below the median. So this was chosen then, to evaluate patients further.

And this is showing that overall, it didn't matter what drug was chosen or what tumor type was evaluated, that this end point was true. So here if we're looking at different drugs, basically very few people responded. This is the same patient here in the extreme resistance group. Responding didn't matter what the drugs were or the tumor type.

And this is a summary overall, and the black dots are the thymidine end point, the open circles are clonogenic assays, and we can see that they're all clustered together. They were found to correlate directly. So this was a way of validating thymidine compared to what had been an important end point, the colony forming assay. And this is the threshold that was chosen as the resistance end point that shows that only one out of 127 patients responded clinically if

their tumor fell in that category.

So Bayes' Theorem, I won't go over this in detail because Dr. Weisenthal did, but you have a pretest result, you do a test, and you can assign a post-test result now to the patient's probability of response. And this is the prediction based theorem showing that in this study at UCLA, that the patient results fell on those predicted Bayesian lines, and the pretest probability was then altered by the test to give a post-test probability.

And here we can see, if you're an extreme resistance category, your post-test probability is going to be significantly lower than your pretest probability, whereas if you're in the low resistance category, it's the opposite.

So we looked at a number of patients who for Taxol, to compare because of the reasoning that if you're over 65, are you going to be different than you're under 65. And I think Medicare would be concerned that there might be differences in age groups for results in the assay. We found no difference if they were less than 65 or more than 65 in terms of the frequency of extreme resistance to Taxol, which is commonly used in breast cancer.

Now, another way of validating a test is to determine if the end point your test measures can be confirmed by a second end point. So we looked at peglotical protein expression, we published this in Clinical Cancer Research, where the degree of peglotical protein increasing on the tumor correlated directly with decreased response to Taxol in the assay. So we took a known mechanism and correlated it with the assay result, and found a direct correlation. Adreomyecin did not correlate as well, because there were multiple mechanisms of resistance for adreomyecin.

And I think this is a very important

point to emphasize, that we can't go out and measure specific mechanisms, because cells are very complicated. So what an in vitro test does is it takes all the mechanisms working in situant cells, grown as little clumps, to recapitulate the in vivo growth. And it shows that multiple mechanisms can be integrated into a net effect result that correlates with clinical response. And I don't know of any patients who if they don't respond to chemotherapy, are going to do well. The only patients to do well are the ones that do respond.

So this is just briefly then, to close, a study that we have done that was peer reviewed

00124

by the ASCO committee and presented at the ASCO meeting last year, where we looked at breast cancer survival as another end point, and compared survival of 96 patients who were tested in the assay, who received chemotherapy with their EDR scores, nodal status, and clinical stage. And if they were resistant to two drugs, they were given a score of 0. If they had low drug resistance to both drugs, they were given a score of 4, and so forth in intermediate categories. And just briefly, I'll say there were no statistical differences between the groups who were resistant and sensitive in the assay in terms of stage, lymph node status, tumor size and so on.

They were also treated evenly in terms of hormonal therapy, mastectomy versus lumpectomy, and chemotherapy agents that were chosen to treat them. This was a blinded study. The patients were treated with chemotherapy empirically, and the outcome and survival was compared to the assay result. We found in multivaried analysis, progression free survival was significantly worse. The relative risk of progression was 2.9 fold higher for patients who

00125

were resistant versus low resistance, and that was similar to the poor prognosis conferred by

high nodal status greater than 10, or stage four versus stage one. And similarly in overall survival, there was a significantly worse survival if you had any resistance in the assay. And this shows the progression free survival curves, which were significantly different, and the overall survival difference. For patients with low resistance in the assay versus whatsoever to the drugs they received.

Now there are many other survival studies that have been done, smaller studies, but again, the majority are internally consistent for showing significant differences in survival where test resistant people survived for significantly less time than patients who got drugs to which they were sensitive in vitro. This is a more recent summary of survival correlations, where in these studies with over 300 cases, 400 cases, where different people looked at different tumor types, it shows significant survival advantages to receiving drugs that were not resistant compared to drugs that were resistant.

So I think that the end point of

00126

survival has been addressed, correlations between the assay technology and other validating methods have been addressed, so I believe that if these tests are used by normal incorporation into the routine practice, you get a biopsy, you look at the diagnostic information from pathology, you look at prognostic markers, you look at drug resistance information, staging information, and then planning goes on between the physician and the patient, based on integrating this information together.

So in vitro assays do correspond to response with specificity and sensitivity that are adequate and comparable to other clinical tests. Survival is significantly associated with in vitro response. Assay directed therapy has improved outcomes. And other people talked about cost. You've heard about survival. So I believe that levels of evidence two and three have been

met, the standard criteria for covering these kind of things, and this is how we approach tests in medical oncology.

So, because there isn't time for questions, I will conclude at this point. Thank you.

DR. FERGUSON: Thank you. We have actually 25 minutes now, for four presentations, Dr. Orr, Dr. Hoffman, Dr. Bosanquet, and David Alberts. And that's, you know, seven minutes or so apiece, and I guess since the presenters -- I guess I'm going to have to crack the whip more than I have. So, Dr. Orr? I am going to hold people to seven minutes this time.

DR. ORR: I am Jimmy Orr. I am a gynecologic oncologist in private practice in GYN oncology in Fort Meyers, Florida. I currently am a clinical professor at the University of South Florida. And I do have a conflict of interest, in that some of the data that I will present today was supported by Oncotech, some of the peer review data.

If one talks about therapeutics and the roles of therapeutics, we've already alluded to Robert Lowell's rules, who was a barred professor of medicine at Columbia. And all of these apply, I think, to the treatment of patients with cancer. If what you're doing is good, keep doing it. If what you're doing is not good, stop doing it. If you don't know what to do, do nothing. And finally, never make the treatment worse than the disease. And each of these apply, I think, to the treatment of women with gynecologic cancer.

As far as peer review evidence as it relates to the use of drug resistance assays in women with ovarian cancer, and I will remind you that they comprise about 25,000 cases a year in this country, and the median age is about 64. Two years ago Roswell Park presented some data looking at drug resistance assays, and I think

two very important aspects of that paper need to be emphasized.

Number one, in looking at patients with extreme drug resistance assay, if you compare those patients who had no extreme drug resistance to platinum and Taxol, or extreme drug resistance to platinum and Taxol, it took two very important end points, complete surgical response and progressive disease, one can see that the presence of EDR to platinum and Taxol halved the complete surgical response. And that becomes extremely important, because if you look at complete surgical responders, survival is clearly different from those who were found at second look to have persistent disease.

00129

And the incidence of progressive disease during initial treatment was almost doubled. In a current abstract that has been submitted, when one looks at extreme platinum resistance, and looks at progression free survival, one can see that patients who are extreme platinum resistant have roughly half of the progression free survival, and roughly half of the estimated five-year survival, both being statistically and certainly clinically significant.

I would like to address for the remaining three to four moments, is this test cost effective. In a recent peer review journal article we submitted, the Cancer Journal, we evaluated the cost effective treatment of women with advanced ovarian cancer by cytoreductive surgery and chemotherapy directed by an in vitro assay for drug resistance. All patients received cisplatin but the second drug in combination was guided by the results of the extreme drug resistance assay. That is, platinum with Taxol, or platinum with cyclophosphamide, as guided by the assay results. As one knows and understands, there are significant incidents of extreme drug

00130

resistance across the board of patients who have

up front treatment with ovarian cancer, in the neighborhood of 35 percent for platinum, 20 percent for Citoxan, 15 to 20 percent for carboplatin and cisplatin. So the incidents of drug resistance is very common in patients who have not received previous treatment.

Our overall survival was 66 percent at three years. There was no significant difference between patients treated with Taxol carboplatin and platinum in Citoxan. If one looks at the average cost to the bottom line, the average total drug cost to the average drug cost per patient, and then adds in the average chemotherapy cost per patient with or without the assay, and then sorts out treatment related results as far as drug cost per patient, assay cost per patient and treatment cost per patient, we can see that assay directed therapy appeared very favorably in comparison to the standard as most would say today, of Taxol and platinum. And more importantly, the cost effectiveness per patient, that is, taking the cost divided by the overall survival, also appeared very important.

Assay drug related treatment can be

00131

used cost effectively in a significant number of women with gynecological malignancy, and particularly in those women with ovarian cancer. Assay guided therapy can improve their survival and is clearly prognostic.

Thank you very much.

DR. FERGUSON: Thank you very much, Dr. Orr. Next is Robert Hoffman.

DR. HOFFMAN: My name is Robert Hoffman. I am the founder and president of AntiCancer, Inc., who sent me here. I am a graduate of Harvard University, where I did my Ph.D. in cell biology. I trained in tissue culture at Massachusetts General Hospital with John Littlefield, and I have been in this field for approximately 30 years.

We have developed what we call the histoculture drug response assay, which is based

on three dimensional culture of cancer tissues, and I'd like to introduce you to this assay, if I may. The emphasis is on three dimensional culture that preserves the tissue structure of the cancer during the culture. The evaluability rate is approximately 95 percent. Mono and combination chemotherapy can be evaluated.

00132

There's correlation with sensitivity, correlation with resistance, correlation with survival, and increased survival has been found for assay directed therapy.

This is just an example of a stomach tumor. This is what it looked like coming out of the patient, looked like after two weeks in histoculture. So, the key point here is preservation of tumor structure and tumor physiology.

To give you an idea of how the test works, this is just one example. This is human breast cancer. Sensitivity to Doxorubicin measured with the MTT end point. And here we have tissue culture plates; these are sponge gels on which pieces of tumor are cultured, and we have increasing concentrations of Doxorubicin, and you can see just from the no treatment, the dark staining MTT, to treatment at 29 micrograms per ml, where there is no staining at all. And you see a gradation over the increasing concentration. So this gives you an idea of how the assay works.

Here is an example of breast cancer with a front line therapy, adreomyecin or

00133

Doxorubicin, and we were measuring the percent of dividing cells, in this case by a thymidine assay. And the point of this is that there is approximately two to three orders of magnitude in sensitivity over a series of patient tumors. So empirical therapy, I believe, does not give us very valuable information.

This is a study we published a few years ago on the clinical applications of the

HDRA or histoculture drug response assay, published in Clinical Cancer Research, where we emphasized correlation with survival.

And these are studies done on gastric cancer patients treated with mitomycin C and five fluorouracil, using the MTT end point. And we just, this is survival and this is disease free survival, and these are patients that were HDRA or histoculture drug response assay sensitive, surviving a considerable period of time. These were patients that were resistant in the assay to these drugs, mitomycin C and 5-FU. And the recurrence free survival has a very similar curve. In other words, the HDRA sensitivity correlated with increased survival.

We did further studies on the survival

00134

with the HDRA in a 46 center study, also published in Clinical Cancer Research, and I'd like to share these data with you, some examples of survival. These are gastric cancer patients, high stage, three and four, and the patients that were sensitive in the assay are, have a considerably long survival. The patients that were resistant in the assay, in this case to mitomycin C and UFT, which is a 5 fluorouracil derivative. So again, correlating with survival.

In a subsequent study that's not yet published, we correlated survival in the histoculture drug response assay with response to mitomycin C. And all the patients here were treated with mitomycin C. Survival again, here are the patients that are sensitive in the assay, surviving significantly longer than the patients who were resistant. HDRA resistant patients here, HDRA sensitive patients here. So there is a statistically significant difference in survival between HDRA sensitive and HDRA resistant, as there have been in all the studies I've shown you thus far.

Here I would like to show you a correlation between assay directed therapy and

00135

clinician's choice, using the histoculture drug response assay for survival. And this is with gastric cancer patients, and here are the patients with, that are directed by the assay, assay directed therapy, and there are a series of 12 patients here, and this is their survival for a period of 18 months or so, and here are the patients that were resistant in the histoculture, in the HDRA, and treated by clinician's choice, and you see their survival here. And the average survival for the HDRA or the assay directed patients who were HDRA sensitive was 9.8 months, compared to 4.7 months for those who were HDRA resistant and treated by clinician's choice.

So what I -- we've published over 40 papers on the HDRA and what I've summarized for you here are studies with survival, which we consider to be the ultimate end point. And we have shown that the HDRA not only correlates with survival, but even in a prospective study, assay directed therapy can seemingly increase survival.

Thank you very much.

DR. FERGUSON: Thank you, also for staying on time. Next, Andrew Bosanquet?

00136

DR. BOSANQUET: Thank you very much. I am Dr. Andrew Bosanquet, from the Bath Cancer Research Unit, in England. I've come here by the invitation of Larry Weisenthal. Because I run a small charity, my fare is being paid for by Oncotech. At Bath Cancer Research, I want to show you some of the work that we published in the last year with the DiSC assay that was proposed by, sort of invented by Larry Weisenthal. In 1991 we published this survival curve in chronic lymphocytic leukemia patients. We have worked very much with chronic lymphocytic leukemia and most of the results that we've got are with this disease, and all the results that I am presenting now are with chronic lymphocytic leukemia.

So here you see in the 1991 work,

patients who were sensitive to the drugs they received survived longer than those who were resistant to the drugs that they received. In this latest paper that we published in the British Journal of Hematology this summer, we are looking at fludarabine, and we used the DiSC assay to determine drug sensitivity to fludarabine. Patients were treated independently

00137

and then we compared the results. The treatment of patients was either labeled with fludarabine, i.e., fludarabine was given within the first year of the test being done, after the first test being done, or with any other chemotherapy. And the point was that no fludarabine should be given.

We have 243 patients who came into this study. Those who received fludarabine versus those who received other chemotherapy, these are the numbers, very similar age and stage, sex ratio, relatively similar previous chemotherapy. The response to chemotherapy, which is one of the end points that we've looked at often, in the test sensitive patients was around 18 percent and in the test resistant patients was a zero, or one patient out of 15 responded, but just for a short time. Very significant difference in response.

But survival is the important thing. Here, of patients who received fludaribine, here is their survival if they were test sensitive, here is their survival is they were test resistant. Now in chronic lymphocytic leukemia, you would expect a survival of four or five years, and with fludarabine, fludarabine is the

00138

modern drug to use for this disease, and so you would expect patients to have this sort of survival. But notice these patients; they have been given the best drug for this disease, they are expected to survive for five years, and they were all dead by 17 months. These patients we looked at in some detail, the 15 of them. They were too sick, having received fludarabine to

which they didn't respond, they were too sick then to receive any other chemotherapy. If they had only received some other chemotherapy first, other than the best drug, what is considered the best drug, they could have survived longer.

Was it that these patients had a poor stage and so on? The answer is no. These are the same two curves divided into those without any previous treatment, and there is still this same difference in survival in those who had received previous treatment. Stage and sex and so on were very similar between those two.

Now this is the same 15 patients who died very soon after receiving fludarabine, but note, these are all fludarabine test resistant patients, and this line is a line of patients who received any other chemotherapy. It didn't

00139

matter what they had, as long as it was chemotherapy within one year of the test. It wasn't assay directed even, it was just they did not receive fludarabine, they did not receive the best drug. And they survived longer, because these patients were test resistant to the best drug.

So here we see patients surviving longer even though they are not chosen by the test result. And if we look at these two groups of patients, again, there's very similar characteristics, the age, stage, and sex, and previous chemotherapy that constitute parameters that are important in the treatment of CLL. But actually you will see that they are actually very sensitive to other CLL drugs. Both groups of patients, 80 percent of them were sensitive to other drugs, whether it be prednisolone, doxorubicin, pentostatin, vincristine.

And so as a result of these and other experiments that we've performed, the DiSC assay is part of the second randomization in the U.K.'s national medical research counsel, CLL 4 trial. And the second randomization is to treatment guided by the DiSC assay, versus treatment guided

00140
by protocol, which is essentially physician choice.

So in five years time, I hope we can have a result from that for you, which will be, this is going to have 500 patients entering into it. This will be a good robust study using randomized control trial.

Very briefly, we did test a very similar drug, calatropin. 34 patients were treated. We did a concurrent DiSC assay, so this was a prospective study looking at -- patient characteristics, I won't go into, and here you say see the test results, raw data on the left-hand side on whether the patients received a complete, partial or no response. And as you see, those who responded had low test results, i.e., high sensitivity, and those who did not respond had a very resistant test, apart from these two who were withdrawn drawn early.

Just briefly on the economics of it, here are a set of CLL patients who were either given drugs to which they were resistant to, and every other test, drug that we tested, they were also resistant to, so we couldn't expect to do very much for these patients. Here is a survival
00141
of patients who were given drugs to which they were sensitive to, and here is the survival of patients who were given drugs to which they were resistant, but they had drugs in the test to which they were sensitive. They should have gotten drugs they were sensitive to, and survived along this line. And if we work out, if we look at the data on this, this was very significant, this unused sensitivity we called it, where they could have had better treatment.

And the cost per life year gained, if we'd used DiSC assay guided treatment there, over all the patients tested, not just that group, but over all the patients tested, was 1,500 pounds, or $2,500, and this value of $2,500 compares with the cost of treating CLL patients per life year

gained, which is enormous, compared to the cost of extending the life for a year in patients by using a drug sensitivity test.

DR. FERGUSON: Thank you very much, Dr. Bosanquet. Do you have handouts for some of this material?

DR. BOSANQUET: We do have copies of the three papers to which I referred, and we can give them to you.

00142

DR. FERGUSON: Now, Dr. Alberts, is Dr. Alberts here?

DR. ALBERTS: Yes.

DR. FERGUSON: Okay.

DR. ALBERTS: I am here on my own recognizance. I am a professor of medicine and pharmacology and public health at the University of Arizona, associate dean for research in the College of Medicine. And in terms of my experience, I am the chair of the gynecologic cancer committee for the Southwest Oncology Group, and have been since 1977. I also chair the cancer prevention and control committee in the Gynecologic Oncology Group.

I came to the University of Arizona in 1975 to help Dr. Sid Salmon develop an assay that could individualize chemotherapy for patients with a broad variety of tumors. I must say that it's a sad note that I'm here today, because Dr. Salmon died October 6th, of pancreatic cancer. But I think his spirit is here and in fact, he pretty much developed this field.

I will point out that since 1975, the options, the possibilities for treating ovarian

00143

cancer are tremendously increased, that's my expertise, ovarian cancer, and in fact, it's a very confusing field. People might want to make you think that it's a simple field in terms of selecting agents. There are 22 drugs that are FDA approved that have activity for ovarian cancer, 11 of them are specifically approved for

ovarian cancer, and it is absolute chaos certainly in the second line treatment of these patients to determine what drug should be used for any one patient. And I can assure you that physicians are not infallible in this situation. On the other hand, I think what you've heard today is that the tests that we have available to us can lead us out of the wilderness in relationship to these problems.

Now I think, I am very impressed with the presentations today. I mean, I know where I would vote on this. There is acceptable quality control and reproducibility, acceptable accuracy, and acceptable clinical utility of these tests, and this has been shown over and over and over again, for a variety of tumor types. We know that drugs that don't work don't help people. And certainly, if we can identify at least those

00144

drugs that are not active, and not give them to patients, we're not going to at least harm those patients. Giving inactive drugs to patients is harmful, and it's cost ineffective.

I think Mr. Stein very eloquently pointed out that if a patient was given the opportunity to really understand what the options were, there is no question that they would want to be treated according to the best knowledge that existed for them on the basis of their tumor. There are always questions about risk benefit concerns, palliation, quality of life, and when you have assays that are 99 percent accurate in identifying inactive drugs, we've got to be serious about taking these results to heart.

In this very same city, just, I think it was just exactly a year ago, actually not even that, nine months ago, at Mercy Hospital, I participated in a symposium, a gynecologic cancer symposium, and I was asked to speak on this: Drug resistance assays, when possible, should be used to guide primary therapy of ovarian cancer, and I was asked to speak on the pro side. It was

sort of a randomized way in which speakers were selected.

I finished this talk asking the audience, which were, there were about a hundred gynecologic oncologists in the audience, if you're sitting in your office and you have specific information on the tumor of the patient that you're treating that shows that nine out of those ten drugs in the second line treatment are associated with extreme drug resistance, and one of these drugs is associated with sensitivity, and you're going to see that patient in five minutes, would you choose to look at that data, would you be interested in that data, or would you like to avoid that data? A hundred percent of the people said of course, they would take into consideration the data that were presented in the laboratory report from a valid lab. I won the debate, by the way.

Well, I just want to sum up. I'm sort of the summary speaker here. Drs. Weisenthal, Hoffman and Bosanquet have all shown survival advantage for test sensitive versus test resistant drugs. Dr. Fruehauf pointed out to the accuracy of the extreme drug resistance assay, 99 percent to identify ineffective drugs, and that test results are valid across a whole variety of tumor types and drug classes, and finished with a discussion of poor breast survival with test resistant drugs. Dr. Hoffman talked about GI cancer survival being poor with inactive agents, and I think that's really impressive, with gastric cancer especially. And finally, Dr. Bosanquet's presentation that you've just heard, showing again, poor survival with chemoresistant disease.

And I just -- I am not going to belabor this any further. I am going to give you my conclusion slide. In vitro drug response assays for cancer specimens have definitely matured with third generation technologies. I was involved

with the first generation and I think it's extremely interesting that Dr. Salmon's own human tumor stem cell lab has been converted completed to using tritiated thymidine end point assays for solid tumors. The accuracy, sensitivity, specificity are excellent, and comparable to conventional testing. Results apply to both first line and salvage settings. Cancer is still primarily incurable but many new agents are available with activity. Their selection must

00147

be, not just should be, guided by data, but not gut feelings. And unfortunately in oncology today, and I think you're all aware of it, gut feelings are too pervasive in our selection of treatment. These tests should be covered by HCFA.

Thank you very much.

DR. FERGUSON: Thank you. I hope the panel and the audience and participants will excuse this sort of marathon session, but I think people need to have their say, to present their information.

We will go right on to Dr. Nagourney, and I think these next four presentations are going to be allotted 15 minutes each, and we will be just a few minutes late for lunch.

DR. NAGOURNEY: First of all, I'd like to introduce myself. I am Robert Nagourney. I am a hematologist oncologist and also founder of Rational Therapeutics, which is a laboratory which applies cell death end points in the study of human tumor biology. I am here by virtue of United Air Lines frequent flier miles, and I paid for my own hotel room.

I'm going to try to put this in a

00148

slightly different context, as I present. I would like to look at some of the scientific issues that have led me to certain conclusions and perhaps will bring you some of the insights I have gained over about 18 or 20 years of work in this field.

I'd like to start off by saying that good medicine always follows good science. When we think, and you have heard a review of the various techniques that have been applied in primary culture studies going back to 1954 through more recent techniques. Various investigators have looked at this area and have tried to cess out mechanisms by which they can assess responsiveness in individual patients based on findings in a laboratory. The underpinnings of this might be described by the equation biomass equals cell growth minus cell death. This equation has been primarily examined when biomass grew in a tumor as a function of cell growth. But I think we are now witnessing a change in that thinking and the focus is now shifted to cell death events.

So I'd like to discuss a little bit about advances in our understanding of tumor

00149

biology, the concept of cell proliferation end points versus cell death or apoptotic end points, and how they may help us to decipher some of the data. It was said by John Reed in a recent editorial in Journal of Clinical Oncology, October '99, that essentially all traditional anticancer drugs use apoptosis pathways to exert their cytotoxic actions. Thus, drugs that were largely developed for cell growth inhibition and other purposes perhaps really act through different mechanisms. In addition, it is difficult to measure these events using cell growth end points, which was pointed out actually earlier by Shakespeare, who said that the absence of proliferation doth not apoptosis make. That's not actually a true quote, although it sounds like it could be. Don't look for it in the Henry cycle.

In any case, what we've really focused on is the concept of an apoptotic event, the induction of cell death in vitro as a predictor of outcome. When Isaac Newton was asked how he discovered gravity, he said by thinking upon it

continuously. And between 1990 and 1995, as an in residence faculty for UC Irvine, I really

00150

thought on continually what it was that constituted cell death events in a test tube. What could you measure that might allow you to predict a patient's outcome based upon apoptosis rather than growth inhibition?

My first stab at this effort was to try to apply the morphologic changes described in 1972 by Kirwelli and Curry in their original paper, in this dosed response curve from cells in culture all the way down to the apoptotic and shrunken cells morphologically characteristic of apoptosis. We tried to apply the DNA degradation profiles known as the 180 KBP DNA degradation ladders. We found that to be a relatively difficult method to use, and in fact was not predictive, because ladder profiles are not actually predictive of human tissue in primary culture, but more of a cell line phenomenon.

We then examined a different end point, which is the inverted field gel electrophoresis which uses a 50 KBP DNA degradation, and we were able to show that this indeed did correlate with responses in some of these profiles where patients had very excellent responses. This is a gel, inverted field gel of patient's tissues

00151

studied following drug exposure, and looking for the 18 hour DNA degradation profile. However, again, this can only be applied in pure cultures.

In addition, a variety of papers were showing that some of the end points using DNA labeling, such as the insulin tunnel end points, did not reliably identify apoptotic cell death. We moved from that area then, and during the same time, into membrane perturbations, alterations of mitochondrial function and membrane potentials that might enable us to predict responses based on what was occurring at the metabolic level. And although again, these were very interesting

findings, they did not apply broadly to human tissues, because you needed pure cultures, and it was only applicable to, in our studies, the chronic lymphocytic leukemia and leukemias in general.

Finally, we moved to some of the DNA markers, some of the mutational events that lead to cancers as modulators of apoptosis, and the morphologic events at the bottom can be characterized by the interplay between positive and negative modulators of apoptosis. We are

00152

currently completing and have submitted a study on the correlation between BCL XL overexpression and drug resistance in human primary cultures. However, again, this is a system that can only be applied in very pure cultures, and did not allow us a practical and high throughput end point.

All of the assays that we have been interested in are what I would describe apoptosis based or cell death based, and these have been described so far for you as DiSC, or the one that we applied which is a modification thereof, the apoptotic, MTT, ATP, FMCA, and others. My principal work has been with the differential staining technique, which takes cells in culture for three days, and then assess the responsiveness of the tissues based on the ability to induce apoptosis morphologically and metabolically.

Now, as a comparator, I thought it might be of use to look at an older technique, the soft agar cloning in a preclinical setting and compare it with a cell death end point for a drug that we now know is used in the treatment of ovarian cancer. In 1992 a study was published using the soft agar assay, where they examined

00153

the ability of a drug, Topotecan, to induce cell inhibition or growth inhibition culture, and they showed under these conditions of culture that 83 percent of renal cell cancers were sensitive, leading to a clinical trial conducted at Sloan

Kettering by Dr. Ilsin, in which 15 patients received Topotecan for renal cell carcinoma, without single response.

At about the same time, we were applying the cell death measures in a laboratory setting to the same drug under similar conditions. And we found quite surprisingly, in 1994 and presented in 1995, that ovarian cancers appeared to be a particularly good target. Two years later, and three years later, several published studies revealed the same, and the observation led to an FDA indication for that drug. So in a growth cell assay, a very erroneous result was predicted, whereas in a cell death assay, a more robust end point, a very accurate prediction.

In point of fact, these observations have continued through the years that I've applied this laboratory test to a variety of observations. Starting with alpha interferon

00154

synergy, which was subsequently proven in 1996, 12 years later, by Wadler, et al., to show that this was occurring by virtue of up regulation of thymidine phosphorylase. The original observation of chloradioxydenasene's activity in hairy cell leukemia was conducted in my laboratory at Scripp's Clinic, subsequently proven by Larry Purot, published in the New England Journal of Medicine, providing a 90 percent complete remission rate in hairy cell leukemia. Our observations of pure synergy using a cell death end point, have led to a point where Howard Hockster has reported with ECOG a 100 percent response rate with fludarabine plus Citoxan in patients who received a combination really which was found years earlier in our laboratory. Our observation that chloradioxydenasene in blastic CML was subsequently confirmed in a clinical trial showing a 47 percent response rate in patients with blastic CML.

An observation which I'm particularly

proud of, and one which you've heard consistently and which I think functions as a perfect example of the discriminating and robust nature of this

00155

assay was my observation in the laboratory between 1992 and 1995, reported originally in 1995, of the true synergy between Gemcitabine and Cisplatin. When we were first provided this drug, LY 1808, that drug didn't even have a name. We began to study it, and found that there was an enormous amount of synergy. That has now been confirmed, as you have heard repeatedly, in a variety of studies, having been approved for FDA indication in non-small cell lung cancer. However, ovarian, breast, and other diseases are rapidly showing the same data that we have generated back as far as four or five years ago.

And finally, as I've mentioned, the Topotecan data. I would like to use the Gemcitabine data which you have seen, and so eloquently provided by Dr. Nalick and also by the gentleman who had presented with pancreatic cancer, Randy Stein, who was the beneficiary through Dr. Weisenthal's laboratory of this observation.

When I describe this interaction, I would like to point out that if you use the laboratory test as the only indicator of where the FDA should find use of this, you would find

00156

that bladder carcinoma, which has now been published, to provide 60 and 70 percent response rates, with substantial numbers of complete remissions. Bladder cancer is one of the best candidates for this combination. Ovarian cancer is the second best when you use a gradation of IC-50, median IC-50 as the determinant, ovarian cancer would be your second choice.

Interestingly, sarcoma, non-small cell lung cancer, for which there is an indication, and interestingly breast cancer, which I will show you some data on, the only data actually in the world on this point. When we worked on this,

we tried to come up with laboratory based therapeutics that might mimic the laboratory testing, and so I'd like to show you a couple of things that have grown directly from the lab testing. And this is ovarian cancer data, some of which you've heard. I just presented this last Thursday in New York, and this is a Phase II trial and completion. The interesting point of this is that of 17 patients who had failed up to six prior chemotherapy regimens largely deemed untreatable, our response rate overall has been 70.6 percent, with now four complete remissions.

Most striking in this has been the observation that two of two complete remissions were obtained in patients who had failed prior bone marrow transplants. In addition, platinum resistant and platinum sensitive patients have been shown to respond.

I think the data indicating that the laboratory assay correlated with response is provided here, showing that those patients who responded to the combination were the most sensitive, versus the non-responders, less sensitive, and a statistically significant difference between the two groups. When we extended this on the basis of the laboratory, in a disease that is not used, in which these drugs are not applied -- in fact, the only data available if from our laboratory, and I can't present it in a formal way because it's still in submission for review for publication, but the only data existing on this comes from our observation in the laboratory, which indicated that a disease for which platinum and gemcitabine are not widely used would be an excellent candidate for this treatment.

In our experience in this Phase II

trial of 30 patients, we have had an overall response rate of 30 percent and again, very interestingly, two of four post-bone marrow transplant patients have shown objective

responses. When we examine survival in this group of patients, you can see that the patients who were found assay sensitive in the green line, versus the patients who were found assay resistant in the pink line, statistically significantly differed, and this cut across other statistical considerations, including HER-2 positivity and number of prior treatments and performance status. The single strongest predictor of these patients treatment response was in fact their sensitivity in vitro.

So when we look toward the hardest evidence that we can provide to a committee like this as to what it is that would provide you evidence to move forward on an approval, I think this is sort of a work in progress, with a lot of very encouraging observations. You've already heard from Dr. Bosanquet some very elegant data that he's generated over the years. This most recent paper, I think the most compelling, published in the British Journal of Hematology in

00159

'99. Dr. Hoffman has presented you some very exciting data in a small study with colorectal cancer. I've given you some data on breast cancer and ovarian cancer, much of which is in progress.

What I'd like to point out, however, is that in the coming years, there will be trials under GOG auspices, a trial in development right now through a group in New York for an assay directed ovarian cancer first line trial, and a meta-analysis, part of which Dr. Weisenthal has provided to you, as an indication of the merit of this and the developing data to support the merit of this in the years to come.

I leave you with a quote from Albert Einstein, who said that in a good mystery story, the most obvious clues often lead to the wrong suspect. In our attempt to understand the basis of nature, we find similarly that the most obvious intuitive explanation is often the wrong one. Thank you.

DR. FERGUSON: Thank you. Okay. Dr. Kern?

DR. KERN: I appreciate the opportunity to address the committee today. I want to first

00160

reveal the financial support received in the early research development of the test, and also mention that I left the academic world in 1988 to join Oncotech as its first director of operations, a commercial firm. I left Oncotech a year ago and now I am a paid consultant for ImPath, who sponsored my trip and who currently markets the test as a drug resistant assay.

We consider in our laboratory, there are two aims of predicting response to chemotherapeutic agents. One of course is to improve response rates and perhaps even survival by selecting active agents. These are the so-called chemosensitive assays that most labs were concentrating on in the 1980s. However, in our laboratory we decided to concentrate on the resistance aspects of the assay, that is, trying to reduce the side effect by ruling out agents that would not work clinically.

One way we tried to achieve this was to use very high drug concentrations in the laboratory. By doing this we used an average exposure of five to ten times what is achievable in the clinic as a maximum tolerated dose. We reasoned this would not only increase the

00161

predictive accuracy for resistance, but also would minimize the likelihood that a clinically active drug would be overlooked. In other words, we wanted to see, look at the aspect of, would the patient be harmed if the test was faulty or gave false results.

Now in a study conducted at UCLA and published in 1990, we looked at correlations from 450 patients. In this slide what I'm showing is the data where each patient is plotted as a single dot. The patients were tested in the laboratory with the same chemotherapy to which

they were treated in the clinic and the results are plotted here, first as responders, as complete or partial responders, and in the column on the left, non-responders. And we looked at the assay results. At the top are those patients that were very sensitive in the laboratory, and in the bottom those patients who had tumors very resistant to the chemotherapy to which they were tested.

And in this bottom group, which we called extreme resistance, a term coined by Dr. Weisenthal and myself in 1990, we found there was only one responder in this group of 127

00162

patients. In other words, the assay was 99 percent accurate in predicting clinical failure. Also, I wanted to point out that the assay did not predict clinical resistant patients, but rather, this data is tumor and drug specific. In other words, the data is for the exact drugs to which the patients were given. They were tested in laboratories with the exact drugs that we used in the clinic, a very important point.

Now there were different levels of evidence to judge laboratory tests as well as therapeutics. It has been proposed that the National Cancer Institute standards of levels of evidence might be applied to laboratory tests. I just want to point out that these levels of evidence weren't designed, however, for therapies, new therapies. And I believe more appropriately would be to look at the levels of evidence proposed by the CDC for evaluating clinical tests. One is to look at the accuracy and precision of the test, its clinical effectiveness, the clinical context in which the test is used, including, do the patients have free access to the test, turnaround time and cost, the practical values of the test, and also,

00163

what is the impact if the test is wrong? What would be the effect of, if we gave medically misleading information?

One of the classical ways of looking at the effectiveness of a laboratory test is to look at its receiver operating characteristics. In this slide I will show the receiver operating characteristics first for the assay's ability to predict sensitivity, that is, predict active drugs. Now, the use of the test can be determined by how far it deviates from this diagonal line. If all the data fell upon this diagonal line, the laboratory test would be totally worthless. By measuring this area under the curve, one can get an estimate of how worthwhile the test is.

Also, the worthlessness or usefulness of a test can be estimated by the prevalence of the marker or the facts that the test was designed to measure. In this case we're trying to measure resistance. So in a clinical setting where you're looking at a high prevalence of resistance, that is, in very refractory cancers, the test is not vary good in identifying active agents. The red line shows the ability of the

00164

test to detect resistance, clinical resistance. When there is a high -- I'm sorry -- when there's a low prevalence of resistance, that is, for those cancers that are extremely sensitive to chemotherapy and highly curable, the assay is not very good at predicting drug resistance or clinical failure.

But in the real world, knowing that most of the tumors fall in a central range here, the test is extremely good for both identifying active drugs and also inactive drugs. In fact, it's almost perfect in this range of identifying drugs that will fail in the clinic.

Another way of looking at the test is to look at its negative and positive predictive values. I have plotted across -- the predictive accuracy, the positive predictive accuracy are a number of data sets, first, from one laboratory shown in green, against a number of different tumor types. Second, I have plotted results from

a number of different laboratories, mostly representing one tumor type. For example, Dr. Albert's in ovarian cancer, and so on.

You see the assay is extremely accurate in predicting clinical resistance. The negative

00165

predicted value here I plotted as its converse. That is, 100 minus the negative predicted value. So the test showed about a 90 to 100 percent accuracy in predicting clinical failure. However, the prevalence of resistance had a dramatic effect on the ability of the assay to predict drugs that would work in the clinic.

I also want to point out that although there is no single gold standard about which test is better, this data indicates an incredible of interlaboratory reproducibility. Using this kind of data analysis, you can compare results from a number of different laboratories and in the hands of experienced investigators like those represented here, the tests are remarkably similar. Also, the test is extremely accurate. All the tests show extreme accuracy in predicting clinical failure.

Looking at standards to judge the medical usefulness of the test, which is of course your job, knowing that the drug resistance assay predicts clinical failure with extremely high accuracy, the test can be used, first of all, to avoid giving worthless treatments to cancer patients, and to help the patients avoid

00166

the terrible side effects of useless chemotherapy. The assay is also cost effective because when you eliminate worthless therapies and useless side effects, it obviously makes economic sense.

I want to just take a minute to indicate what I think are some of the reasonable indications for the use of this committees. Remember, we talked about the test appears most worthwhile in the real world represented by the breast cancers, the ovarian cancers, the lung

cancers. In breast cancer, even if the test was only used to identify adreomyecin resistant patients, and to switch those to, let's say CMF or an alternative therapy, the test would be cost effective if you only identified 10 percent of the adreomyecin resistant patients and switched them to CMF.

Here we talked about Gemcitabine. Certainly useful in lung cancer, but if the test was only able to identify 3 percent of the patients that were Gemcitabine resistant and switched them to platinum etoposide or platinum vinorelbine, the test would pay for itself.

Also, equally important are where the

00167

tests may not be very useful, and two particular examples are in cancers where there is an extremely high prevalence of drug resistance, for example, in renal cell carcinoma, the test would not be very useful. That doesn't mean that it's a bad test. The test still has the same sensitivity and specificity, it just means that there are no worthwhile drugs, no useful drugs for treating, drugs like kidney cancer. On the other end of the spectrum, testicular cancer is clinically very responsive, very low prevalence of resistance, the test is probably not needed in that clinical study.

So I want to stop here and just summarize. First, let's talk about how the patients benefit from drug resistance testing. Chemotherapy causes a lot of suffering for cancer patients. The suffering is tolerable if the treatment leads to prolonged survival. But wouldn't it be beneficial to the patients if they could be spared the needless suffering of worthless chemotherapy?

Let's consider the cost benefits of the test. The Medicare system could benefit immediately if it didn't have to pay for drugs

00168

that are useless in the clinic. So don't you think Medicare would benefit economically by

using the drug resistance tests?

And finally on a, consider it a personal note. Imagine if a wife, spouse, parent or loved one should be unfortunate enough to be diagnosed with cancer. Would you not want that loved one to have the benefit of drug resistance testing? So with that in mind, I respectfully ask the committee, please do not deny America's seniors the access to the proven benefits of the drug resistance test. Thank you.

DR. FERGUSON: Thank you, Dr. Kern. I think our next speaker, rather than Dr. Bailes, will be Dr. Hazie.

DR. HAYES: Hayes.

DR. FERGUSON: Hayes, I'm sorry, from Georgetown.

DR. HAYES: I'll introduce myself. I know who I am. I am Dr. Daniel F. Hayes. I'm a medical oncologist. I'm the clinical director of the breast cancer program at Georgetown. I'm also a member of the American Society of Clinical Oncology.

I have a couple of credentials. One of

00169

those is that I am the chair of the solid tumor and correlative science committee of the Cancer and Leukemia Group B, one of the major multi-institutional groups funded by the Federal Government. Our committee is essentially the tumor marker committee of the CAGLB. And I'm also a member of the American Society of Clinical Oncology tumor expert guidelines panel that was convened roughly five years ago. I am not the chair of that. And I'm also hoping that I will have some slides here, which is why I'm wasting your time mumbling around here until this thing gets up and running.

And finally, I apologize. Dr. Bailes, who is the president of the American Society of Clinical Oncology meant to be here today and was unable to do so. And Dr. Dan Van Hoff also meant to be here and also was unable to be here. And finally, I have no conflicts of interest with any

of the companies that have been presented here, not performed any research with any of them. I guess my only conflict of interest is I am a member of the American Society of Clinical Oncology.

Finally, while this thing is hopefully

00170

waiting to boot up, I will say, because I wrote some notes in case my slides didn't work, that the American Society of Clinical Oncology is neutral on this issue and is not here to serve as either a proponent or opponent of your decision to reimburse for any of these assays in the elderly population. Rather, we are here to make a plea regarding reimbursement for treatment, and in fact this was brought up by the last speaker, and this is one place where I believe we would contend that that would be an error, at least with the currently available data.

Thank you. I apologize for the time it's taking to do this. So as I said, we are actually neutral on this issue. We are actually very interested and believe there is a substantial amount of interesting data that are coming out of the various studies, much of what you have seen today. We appreciate the remarkable progress in technology that has occurred over the last 15 years since the original publications by Salmon, et al., but we remain very concerned about the reliability to exclude therapy, especially in relationship to combination chemotherapy.

00171

We have been very interested in establishing guidelines for the members of our society. There have been at least two guidelines panels related to tumor markers. One of those which was specifically related to tumor markers for breast and colon cancer, the most recent update of that was published in JCO in 1998. A second related to tumor markers, but not specifically excluding, or specifically related -- the specific focus on it was the follow-up of

primary therapy in breast cancer, again, published in 1997.

In both of these, particularly in the first, we spent a great deal of time trying to decide when a tumor marker is truly ready for prime time, something I believe you are being asked to contend, and in fact the discussions by the FDA this morning I found very interesting. We like they found there are very few rules about how to use a tumor marker in clinical practice, unlike how to use therapeutics. And our predecessors 30 to 40 years ago, Fry, Holland, Karnoski, established rules and guidelines about how to talk to each other, what is a Phase I, what is a Phase II, so on and so forth, what's a

00172

complete response. Those sorts of things have not been established well for tumor markers, and we found it very confusing.

Ultimately, one of the things that we suggested was that the results of the tumor marker must be known to influence the decision to result in improvement in overall survival, disease free survival, quality of life or cost, echoing many of the things that have been said here today.

Also echoing many of the things that have already been said, we noticed an astounding amount of heterogeneity among tumor marker studies, and these are for many reasons. Patient selection, different assay issues, the use of different drugs; for example in this case, we've been talking about the difference between the use of single agent therapy and multiple combination therapy. And probably importantly, the difference between the settings, whether or not these sorts of assays should be used to direct the therapy in the adjuvant setting where you don't have an identifiable end point immediately but rather, must wait for progression or survival, or the use of metastatic where you have

00173

things like immediate quality of life and

immediate response rates.

These are some conceptual slides that we developed relative to any tumor marker, and that is a pure prognostic factor, for example, separates two groups of patients in the absence of therapy or in the presence of therapy equally. It does not tell you whether or not that therapy will be helpful but rather, it tells you how the patients will do. Often, the difference between prognostic and predicted factors gets mixed in many papers and also in many discussions. We felt it important to point this out.

So that for example, these curves are clearly separate, but they are equally separate in the absence or presence of therapy. It does not tell you whether these patients should be, or how they should be treated, it just might tell you that they should be treated because their prognosis is worse.

A pure predictive factor on the other hand, the curves are not separated at all in terms of prognosis in the absence of this specific therapy. I put no therapy here, but in

00174

fact, we could be talking about the specific therapy one is concerned about, whereas in the presence of that therapy, the curves are separated. In this case, if it's a predictive factor for sensitivity to therapy, then the curves are separated, with those of a factor positive are doing much better than those who are factor negative.

Indeed, the real issue that we're discussing, I believe here today, is the separation between these curves, not whether or not they are statistically significantly separated. That is what the P value tells you. What really tells you is the magnitude of the difference, and in fact this has already been brought up earlier at least once today. And that is, for example, many of us would be very willing to use this marker to separate patients into two

groups, and treat them differently, especially if the toxicities of the drug are high. Whereas in this group of patients, we would probably all treat this group the same way we treat this group, because the magnitude is not large.

Now in order to really assess the utility of a predictive factor, it requires a

00175

control group that did not receive the therapy. That is of course best done in the presence of a prospective randomized trial in which biases are minimized because of the randomization. Historical controls are acceptable in some cases, but they are fraught with the usual biases, that is, comparing the outcomes of your patients with those who did not receive the therapy in non-randomized fashion. The use of response rate gets confusing and in fact, it assumes a historical control in which no therapy, frankly, equals no response. We must assume that patients who are not treated will not have their tumors regress. That's not always true, but often it is, and in general, this is a relatively fair assumption.

So then for example, if we take what I believe the best predictive factor in oncology, and that is S Receptor and Tamoxifen, it is both a prognostic and a predictive factor, which makes it confusing. In the absence of therapy, ER positive patients do slightly better than the ER negative patients, but in the presence of therapy, ER positive patients do substantially better than ER negative patients. In this case I

00176

have used Tamoxifen as the therapy, because it's the one for which we have the most data.

And in fact, one could begin to use these sorts of relative differences and develop overview analyses of the relative risk, so that in this case, one means therapy is no better than therapy in the presence of a randomized trial in which the patients are randomized, the therapy are not. And then we have a 10 percent

difference, a 20 percent benefit, a 30 percent difference, or a 40 percent reduction in the odds of the event, in this case let's say it's a recurrence. A weak predictive factor may split these patients apart statistically, but may not be important clinically, since both groups of patients benefit. A moderate predictive factor splits them further apart and a strong predictive factor splits them much further apart, so that one might not treat these patients, and one is very likely to treat these patients.

And we came up with what we called a relative predictive factor, and that is just the relative odds of response in the positive group divided by the relative odds response -- I'm sorry, I have this running on battery. Okay.

Well, I made the point.

And then one might say, all right, we will discard those that are very weak, consider those that are intermediate, and accept those that are very strong. And of course, what we do here, this is highly subjective, depending on the toxicities of therapy and the patient perspective; there are some patients who would accept therapy regardless of the toxicity in the presence of any benefit, and others who would be more thoughtful and say I'm willing to give up some benefit in order to avoid some toxicities.

Again, if we use ER in the adjuvant setting to predict relative benefits from Tamoxifen, the proportional reduction from the last Oxford overview for adjuvant Tamoxifen versus nil, and those were ER poor, was 0.06. In other words, there was a 6 percent proportional reduction in the odds of recurrence and death, whereas in those that were ER rich, it was nearly 50 percent, and the relative predictive factor for ER for Tamoxifen in the adjuvant setting in this case was well over what we would consider acceptable for routine clinical utility, and in fact S receptor is used very much in that

setting.

So the ASCO view of these proceedings are that we believe in vitro chemosensitivity assays are promising. We believe the current data are insufficient to withhold potentially effective drugs. Perhaps in the metastatic setting these might be very helpful. In the adjuvant setting, I really believe that we need prospective clinical trials to address the issue, because all adjuvant therapies are empiric by definition, since one does not have disease to measure at the time.

There is a technology assessment planned by ASCO to be convened in the winter, this winter 2000. We hope that by next summer, the results will be available and published. And we advise against noncoverage for agents found to have drug resistance for individual patients.

Thank you for your time and again, I apologize for the inconvenience of the technology.

DR. FERGUSON: Dr. Loy?

DR. LOY: I'm Dr. Bryan Loy. I'm a carrier medical director. I'm a guest of the Health Care Financing Administration. I don't

00179

need my slides to get started. In the interest of time, I will proceed onward.

I am a part B carrier medical director or CMD for the State of Kentucky. In the interest of time, I will go ahead and start here without my slides. I just wanted to get started. I am not dependent on my slides, because my presentation is of a different tact and not of a scientific nature.

I'm a part B carrier medical director for the State of Kentucky, or CMD. I appreciate being here and I am very interested in these presentations. I am going to go to the end because in my role I will be asked to implement or execute any national coverage policy decision that results from these deliberations. My concerns are the appropriate use of these

technologies for the treatment of patients, but I am also concerned about the possible vulnerabilities that a national coverage policy can create that could result in the misapplication of these technologies.

So the intent of my presentation is to describe the approach that I take at the carrier level to assess the need for policy development

00180

for new technologies and my carrier, and to discuss the impact of implementation of national policy on the carrier medical director, and to discuss some of the dilemmas that may arise as a result of the outcome of these deliberations, and then finally, to present my prospectus regarding human tumor assay systems.

And the reason I am interested in presenting it in this way is because I think it's necessary for the panel to understand that this won't be a yes no answer to a question. It won't be a yes, we will have coverage, no, we will not have coverage. Somehow, carriers and carrier medical directors will be asked to somehow implement a policy decision that is both reasonable and necessary in accordance with the law.

Let me start off by describing the environment CMDs work in as it pertains to medical policy, whether it be national policy or local medical review policy, and this has already been stated. We're currently in an environment where we're dealing with the local medical review policy in the absence of a national coverage policy when it comes to talking about resistance

00181

testing.

The practice of medicine is ever changing for many reasons, and one of the reasons is the introduction of new technologies or new applications of existing technologies. Technologies can fill voids in the practice of medicine. For example, they can provide previously unavailable information. Technologies

can also replace or supplement existing technologies. They can provide better, faster, more complete information. They can direct patient care. The role that technologies play in the treatment of patients is in part dependent on the acceptance and the implementation of the technology by the practicing providers within our jurisdictions. I am fully aware that acceptance and implementation can also be related to reimbursement and coverage decisions.

In my CMD role, routinely I will receive a call asking me to confirm my name and address from a technology company, and usually within a week or so I will receive a packet from a company representative asking me to review the material, and in follow-up I will receive a note or a telephone call asking me if we have coverage

00182

policy or local medical review policy addressing the technology in question. Less frequently, I will receive requests from treating physicians asking for coverage for a particular diagnostic or therapeutic application of the technology.

If there's a significant number of requests for coverage and we are permitted discretion as the local carrier, we will consider making a coverage decision in compliance with the Medicare carrier's manual. Policy at the local level is important in that it describes appropriate coverage within the Medicare program in accordance with the Social Security Act.

National policy, on the other hand, can have different effects on the carrier. Regardless of the number of requests for coverage, local carriers are expected to implement all national policy of the carrier, and therefore, the same national policy can have different effects on different carriers. And here's why: For carriers in states that have providers already utilizing the technology as we have heard described today, when this is a topic of a national policy, well then, the carriers and the providers will have an interest in the

00183
coverage and the utilization and the pricing language in the policy, and how the carrier implements the policy. For carriers in states whose providers do not utilize the technology, the carriers do not receive claims and therefore, there is less interest in the coverage decision.

For providers who do not utilize the technology, creating policy creates a possible future benefit. The policy becomes relevant only when the providers are utilizing the technology and billing the carrier. Over time, technologies gain acceptance by providers and hopefully gain acceptance of national organizations. Many times national organizations will publish guidelines, position papers, et cetera, describing the appropriate use of the technology for patient care. When these guidelines and positions are supported by scientific evidence, practice patterns usually become a national uniform standard of care. This is not always the case.

Sometimes technologies fall out of favor, they do not gain acceptance, or are replaced by better technologies. Unfortunately, sometimes coverage decisions are considered before practice patterns utilizing technologies

00184
are firmly established, and the policy is developed whether they be national or local carrier policies, may be premature. Subsequently, indications can be added, off label use established, as additional research results become available. Provider acceptance of the technology can also change, and some providers may utilize the technology different than providers in other states, and the standard of care can change dramatically as a result of these refinements.

These variances can quickly render policies obsolete. Creating premature policy that is silent on evolving uses of new technologies can tend to work disparate coverage between the states. If a national policy is

silent on an evolving use of the technology, then the local carrier has to make coverage decisions in response to the provider inquiries when they arise. This commonly occurs when off label use of technologies are introduced in the current medical practice. Scientifically based well written national policy can minimize many of the disparities among contractors, and an added benefit is that this process decreases the amount

00185

of resources used by contractors to create local medical review policy.

In my opinion, the Medicare carrier advisory committee is the appropriate forum for raising and addressing coverage questions regarding human tumor assay systems, as opposed to the local carrier level. I also believe that this is the forum to discuss the scientific validity of the test, and to assess the scientific validity of the clinical applications of these test results for each cancer. Implementing this national policy without operationalizing this national policy at the carrier can create some disparities in coverage if the specific clinical indications are not clearly articulated in the policy. When specific clinical indications are identified, specific codes can be employed in order to create system edits for automating appropriate claims payment. When appropriate frequencies are identified, these parameters can also be implemented for automating appropriate claims payment. Ideally, this results in efficient correct payment of claims.

It is my opinion that if the science

00186

does not support the clinical utility of the test results, then the policy should specify noncoverage at this time. If the science does support the use of specific clinical applications of the technology, then the policy should clearly specify the indications in the clinical scenario.

These indications in the case of a test, could

include the clinical setting. For example, in the case of cancer testing, the policy should specify the appropriate cancers and the appropriate times the test should be performed. For example, first line, adjuvant, and/or the metastatic settings. The frequencies of testing and what clinical intervention should have taken place should also be identified in the policy. Again, this should be supported by the science.

Implementing narrative for policies can be difficult and, therefore, clinical parameters need to be well defined. If policy language is vague, or subject to multiple interpretations, then this can lead to misapplications of the policy and can give rise to interstate coverage disparities. Moreover, if national policy is silent on an evolving use of the technology, then providers and carriers are left with a policy

00187

that won't address these evolving issues. Left unaddressed, these applications must be evaluated by the carrier, and local policy will be created to address these new issues. Again, this can lead to disparate coverage.

Polices for technologies in evolution are difficult to write, difficult to implement, and very difficult to maintain. This committee's deliberations will most likely result in one of the following outcomes: A no decision with the possibility of local carrier discretion remaining; a no coverage policy, which can be easily implemented in uniform across carriers; a limited coverage policy which may be supported by current science and lead toward a uniform national coverage, but subject to early obsolescence if applications continue to evolve; and finally, a broad coverage policy, allowing for flexibility for different and evolving practice patterns, but most likely containing vague language.

I would like to conclude by expressing my personal prospectus concerning coverage for human tumor assay systems. Human tumor assay

systems have been in existence for many decades,
but in my community these technologies have not been routinely employed for the treatment of cancer patients. I have reviewed the information supplied by the presenters. These technologies sound quite promising. In my mind, one of the indicators for assessing whether a technology that generates information is needed for treating patients is the number of requests that I receive for coverage of the technology in question from the end user of the information. In this case, it would be the practicing oncologist in my community making the request, not the producer or the laboratory producing this information that would in my mind begin to legitimize the request.

I have had no requests for coverage for this type of testing in my state.

I have asked some well respected practicing oncologists in my community and have generated little interest in these technologies. Coverage and reimbursement issues have not entered the conversation. The discussions have focused on the controversies relating the in vitro results to the long-term clinical benefit. I then questioned multiple carrier medical directors in my region and throughout the United States. Of those responses, only two states reported claims submission or requests for coverage. I do not see where the clinical use of these types of chemotherapy drug assays have been generally accepted or adopted as a national standard of care. It is not clear to me that the practicing oncologists, or those providing these methodologies are yet in agreement on the clinical applications and the clinical value of these tests.

More specifically, questions remain in my mind unanswered as to what point the testing should be used, for what cancers, what clinical scenarios, how frequently these tests should take place. I have not identified in the presenters'

packet any position statements, guidelines, et cetera, that would convince me that this technology has matured into a standard of care. Furthermore, even though some evidence supports the use of this testing for specific clinical indications, the evidence supporting a broad national coverage is insufficient, in my opinion. At this time it is clear to me that this technology is not being utilized routinely for medical decision making for most cancer

00190

patients. During your deliberations I would encourage all members of the impact panel to envision a finished document that would only incorporate scientifically based or evidence based medicine that is currently applicable for treatment of specific cancer types and their appropriate clinical scenarios. If the science only sufficiently addresses certain aspects of clinical utility, then only allow for that coverage. Allowing flexibility in policy for anticipated future trend allows for possible coverage of misapplications of this technology. In my opinion, the science supporting the clinical applications for these testing methodologies is still evolving, still coming in, and there are many unanswered questions remaining. Thank you.

DR. FERGUSON: Thank you, Dr. Loy.

It's four minutes to 12. I suppose we could have questions or comments for four minutes. Any of the panel members? Yes, Dr. Sundwall?

DR. SUNDWALL: One thing that has been going through my mind this morning as we've heard all these excellent presentations, and that is,

00191

how much patient variability is there when they have the same malignancy? Now that may have been addressed, but if it was, it's not clear to me if in fact, once you use this testing, which I think seems to have great value, potential value, but I'm not certain how much variability per tissue

type there is from patient to patient, or if in fact once you get help, that there is sensitivity to a particular drug, or combination of drugs, why isn't that applicable across the board?

DR. FERGUSON: Did you all hear the question? I guess we can have a comment from one of the presenters. Yes, Dr. Weisenthal?

DR. WEISENTHAL: The question that was asked is about disease specific activity versus patient variability. Firstly, clinical heterogeneity with a given disease is an established fact. That's shown by the fact that chemotherapy number one is not universally effective. For example, in the case of a disease like breast cancer, or ovarian cancer, first line chemotherapy will produce a response about 70 percent of the time, in the case of colon cancer, only 20 percent of the time. But more telling is the large numbers of patients that fail first

00192

line therapy that subsequently respond to second line therapy. And I'm not sure that -- but you heard of five cases today that were presented between me and Dr. Nagourney, five patients who failed high dose chemotherapy with bone marrow transplantation. $200,000 in therapy. These were patients that never responded to anything, including ultahigh dose of chemotherapy. They then got an assay and they went into complete remission.

In the case of Dr. Nalick's patient, that was someone that failed first line Taxol platinum, failed tandem stem cell transplants. The amount of money that was spent on ineffective therapy for this patient would pay to run my lab for six months. So the issue is that clinical heterogeneity is an established fact. There are many, many patients who fail first line therapy who respond to second line therapy. They should have gotten the second line therapy the first time around.

Dr. Bosanquet talked about his patients with fludaribine resistance all died, because by

the time they got the fludarabine which didn't work, they were too sick to get anything else.

Had they gotten the correct therapy the first time, they wouldn't have been dead. So the thing is, that all the laboratories that do this know that there is a tremendous heterogeneity with any given disease type. Some tumors with -- some breast cancer tumors are very resistant to chemotherapy, some are not. And the same thing holds for the clinic. That's the whole purpose for doing the testing.

DR. FERGUSON: Thank you. Dr. Brooks?

DR. BROOKS: I have a question for Dr. Kern, who is now with ImPath. Does ImPath have a different type of test, or is it going to offer a certain type of test based on a different methodology?

DR. KERN: No. The methodology is basically the test described also by Dr. Fruehauf, called the EDR at Oncotech, or DRA at ImPath. But it's based on the same technology.

May I also respond to the prior question for 15 seconds?

DR. FERGUSON: Sure.

DR. KERN: I can visually show what Dr. Weisenthal was describing. This is a data set of 40 consecutive ovarian cancer patients, all

previously untreated. You see patient number one at the top, and tested against five different drugs for ovarian cancer. Patient number one was sensitive to carboplatin, resistant to adreomyecin, so on, sensitive to Taxol. Patient number two was resistant to Taxol. You can see the patterns; it's quite different patient to patient.

DR. FERGUSON: What cancer was this?

DR. KERN: This is ovarian cancer. Yes?

DR. BROOKS: Doctor, the question is, I'm not sure what the question was he asked, but the question I would ask, is ovarian cancer

endometrioid, cirrus, mucinous, poorly differentiated, you know, et cetera, et cetera? So amongst the histologies of ovarian cancer, would you have similarity?

DR. KERN: Well, what you actually have to do is look at the clinical experience. In other words, the tests, we try to reflect what's actually going on in the clinic. For epithelial cancer, it's different from clear cell cancer, so you test different drugs and you look at, again, it has to be tumor and histology specific, and

00195

cancer specific, drug specific testing. And you do find in all different types, a great deal of heterogeneity from patient to patient.

DR. FERGUSON: Thank you. One more brief comment. Yes?

MR. KIESNER: One comment in relation to Dr. Loy's excellent presentation. I think one of the observations that Dr. Loy made was that he within his region doesn't receive a lot of requests for the coverage. I think the way the laboratory industry bills, if we get tissue from a thousand hospitals around the country, the issue of where it's billed is dependent on where the work is done. And so in our case, when all the tissue comes to Irvine, we have to bill the local carrier in southern California, so that's where all the payment questions are directed.

A second question is in relation to the technology, there are burdensome drafting requirements. I recognize that, we all recognize that. We sat through the negotiated rule making session last summer. We did have an oncology work group with about 15 people, including HCFA. We were able to deal with those issues; I'm not sure we dealt with them completely, but it is an

00196

issue that can be dealt with.

DR. FERGUSON: Okay, thank you. I guess like many other things, it starts in California and goes east; is that right? Anyway, I think on that, we will have lunch, and

reconvene at 1:00.

(The panel recessed for lunch at 12:03 p.m., November 15, 1999.)

(End of Volume I)

Transcript of November 15, 1999 Afternoon Session

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

00197

VOLUME II
(Afternoon Session - November 15, 1999)

HUMAN TUMOR ASSAY SYSTEMS

HEALTH CARE FINANCING ADMINISTRATION
Medicare Coverage Advisory Committee
Laboratory & Diagnostic Services Panel

November 15 and 16, 1999

Sheraton Inner Harbor Hotel
Baltimore, Maryland

00198

Panelists
Chairperson
John H. Ferguson, M.D.

Vice-Chairperson
Robert L. Murray, M.D.

Voting Members
David N. Sundwall, M.D.
George G. Klee, M.D., Ph.D.

Paul D. Mintz, M.D.
Richard J. Hausner, M.D.
Mary E. Kass, M.D.
Cheryl J. Kraft, M.S.
Neysa R. Simmers, M.B.A.
John J.S. Brooks, M.D.
Paul M. Fischer, M.D.

Temporary Voting Member
Kathy Helzlsouer, M.D.
Consumer Representative
Kathryn A. Snow, M.H.A.

Industry Representative
James (Rod) Barnes, M.B.A.
Carrier Medical Director
Bryan Loy, M.D., M.B.A.

Director of Coverage, HCFA
Grant Bagley, M.D.
Executive Secretary
Katherine Tillman, R.N., M.S.

00199

TABLE OF CONTENTS

00200

TABLE OF CONTENTS (Continued)

VOLUME III

00201

PANEL PROCEEDINGS

(The meeting was called to order at 1:16 p.m., Monday, November 15, 1999.

DR. FERGUSON: Dr. Sausville?

DR. SAUSVILLE: Good afternoon, all. And if I could have the first overhead, this says who I am, and the general topic that I hope you're interested in hearing about this afternoon. Anyway, my task this afternoon is to provide an overview, at least from the perspective of the preclinical therapeutics development program of NCI of antitumor drug sensitivity testing. And I will approach this, therefore, from the standpoint of one who uses tests like this, and indeed, in some cases actually tests that have been used for this purpose for the preclinical selection of drugs for more detailed evaluation, as well as from the perspective of an oncologist who has occasionally thought about using these tests in the treatment of patients. Next.

So the basis for this issue in cancer derives directly from the infectious diseases experience, wherein a number of different disease categories, such as tuberculosis, where it's well

00202
established that one has to establish that a particular patient's infected bacillus is sensitive to the agents, and a number of non-tuberculosis indications, which would include, for example, pyelonephritis or endocarditis, where it is well established from the standpoint of standard medical practice that such sensitivity tests are that valuable. Next.

The assays as applied to cancer ideally would have 95 percent sensitivity and specificity, and short of that goal, would hopefully be better in predicting outcome than the empirical choice of the physician. And the essence of the question from an oncological standpoint, therefore, is whether a particular test conveys information over and above what is implicit in the histologic diagnosis of a patient's tumor. Ideally the test would be biased in favor of detecting sensitivity rather than resisting, for this reason, and ultimately, these tests should be able to demonstrate an impact on ultimate outcome, as opposed to simply response, since in oncology, good outcome begins with the response, it does not end with a response. One ultimately has to have evidence of

00203
tangible clinical benefit that changes outcome. Next.

So among the specific assays that through the years have been utilized include the by now classical Hamburger Salmon clonogenic assay, wherein tumors that were biopsied for example, were disaggregated, plated in agarose or other solid media after relatively brief exposures to drug, and ultimately colonies counted in 14 days. There have been modifications to this, most notably the capillary tube modification used by Von Hoff and colleagues, and it seems to increase the number of patients for which valuable data are obtained. Modifications of this also include radionuclide based assays, in which radioactive

thymidine is added after three days and thus, although it is a soft agar base, one can obtain information after shorter periods of time. And there are also non-agar based assays assessing radionuclide uptake in mass culture. Next.

Technical problems with clonogenic assays include a number of artifacts intrinsic to the practice of the assay, including clumping of tumor cells, the potential of growth perturbation

00204

from manipulation of potential clonogenic cells, reduced nutrient uptake from nonclonogenic cells, with increase in the size of colonies that grow out in the treated cells. Counting evaluations with a potential large coefficient of variation, and poor cloning efficiencies. And a major limitation in the widespread use of this technique relates to the fact that in many instances, the majority of the specimens are not actually valuable, and there is the inability of this type of assay to score small numbers of resistant cells, which in a clinical scenario are thought to translate new ultimately resistance to therapy, of the sort that is manifest by the subsequent relapse of a patient with drug resistant tumor. Next.

In various reviews, actually extending from the initial use of this technique into the early '80s, the cumulative experience is that a relatively small fraction of patients actually have colony growth. And the data that is tabulated here is contained in the references that were indicated. But also, there is the finding that the tests are clearly better at predicting negative or resistant assays, than

00205

sensitive assays, such that for example, if one looks at those specimens that were sensitive in vitro as opposed to sensitive in vivo, we have a 60 percent true positive, with a range of 47 to 71 percent. In contrast to those specimens that were resistant in vitro and resistant in vivo, where there was, as you can see, a 97 percent

true negative information. Next.

This led to a so-called perspective evaluation of chemotherapy selection utilizing a clonogenic assay, as opposed to the choice of a clinician. And again, this was published by von Hoff and colleagues in 1990 in the Journal of the National Cancer Institute. And in the 133 patients randomized in a single agent therapy, of those where the therapy was assigned by a clinician, one had one partial response, and in 19 of 68 that were possible to have an assay directed assignment, there were four partial responses. Certainly there was no evidence that this was statistically different and one concluded, or this article concluded, that what one might conceive of potentially a somewhat improved response rate, did not translate into any noticeable effect on survival. And again,

00206

approximately a third of the tests could not be evaluated, and there were clearly no evidence of survival in patients either treated according to that which was recommended by the physician, or all patients that were compared versus the test population. Next.

Other specific assays which have come to the fore in an effort to meet some of the clear difficulties in the widespread use of the clonogenic assay include the so-called differential standing cytoxicity assay, or DiSC assay, pioneered by Weisenthal and colleagues. And here, one is essentially assessing the effect on whether or not cells remain alive after short periods of culture after exposure to a drug. Thus, either marrow, buffy coat or a tumor suspension after disaggregation, can be treated with drug for anywhere from one hour to four days. Interestingly, the quantification was aided by the addition of so-called duck red blood cells, which are easily distinguishable microscopically, a dye added, and then after a cytospin, one can either assess the dead cells per duck red blood cells, or live cells per duck

red blood cells, based on the differential
00207
staining of live and dead cells with either fast-green, which stains dead cells, or HD, which stains live cells.

Over variants of this approach include the so-called MTT assay, which is a dye that depends for its coloration properties as to whether or not it is reduced by living mitochondria, or a fluorescein assay, where live cells take up a dye, hydrolyze to it in a point that is detected by a change in fluorescence. But all of these techniques, again, don't then depend on the growth out of clonogenic cells, but rather allow a relatively short term exposure to the drug to define whether there is an effect on the viability of the cells. Next.

When this assay was, and again, this is in reference to the DiSC assay, was applied initially to hematologic neoplasms, there was clear evidence that there was increased cell survival, that is to say resistance in patients who ultimately were not responsive to chemotherapy that was assigned on the basis of a knowledge of the tests. So in that respect, the assay was certainly suggestive that it might eventually correlate with clinical outcome. And
00208
in addition, there was a fairly good correspondence, again, with delineation of true positives and true negatives by this assay. Next.

When this assay was applied to the somewhat more difficult clinical category of patients with lung cancer, here in an initial study with non-small cell assay, the DiSC assay was performed assessing sensitivity to ten drugs, treating with a regimen that ultimately incorporated the three most sensitive agents. In this series of 25 patients, there was a 36 percent partial response rate with a median duration of 6.5 months and with the, if you want to read, it looks to be responders and a median,

or I should say a median survival of about seven months, with an overall of about 12 months. There was clearly a threefold lower assay survival. That is to say, people with greater cell kill in responders versus non-responders. However, these authors concluded that outcome as measured by response rate and survival is within the range reported by the literature, that is to say, even though you can detect this difference, the issue of whether or not it ultimately caused

00209

a different outcome that might be afforded by treating with drugs that would be available from the literature and without knowing the patient's histologic diagnosis was not apparent. In addition, some drugs clearly had a much greater discordance in the predictive value of the test.

Thus for example, 5 fluorouracil did not seem to have any ultimate value in its performance, and on the other hand, etoposide, behavior to etoposide, was essentially predictive of the behavior of all of the of the drugs. And actually from a scientific perspective, we now recognize that since many of these agents act by inducing apoptosis, this actual result retrospectively, is not that surprising.

Interestingly, this paper also introduced the concept of so-called extreme drug resistance. That is to say, you can define patients who had greater than one or more standard deviations resistance than the median in the population, and these patients essentially had zero percent response to any of the agents. Next.

This assay was also applied in a study that was recently published from the NIH, and

00210

attempted to individualize chemotherapy for patients with non-small cell lung cancer. And from a population of 165 study patients, 21 received DiSC based regimens, and these had a 9 percent partial response rate. Whereas, 69 patients received empiric treatment with

etoposide and cisplatin; these had a 14 percent partial response rate. And ultimately, the survival of in vitro best regimen was comparable to what one would have expected from the empirically chosen chemotherapy.

Interestingly, this study also revealed an issue that also has to come up in any test in which there is a second or subsequent procedure to obtain tissue, in that the survival of patients who had any in vitro test was actually worse than those without, and this implies potentially that those people that had a sufficient volume of tumor to have the tests had an intrinsically less survival than those that did not. Next.

And the last clinical study that I'll touch on also emanated from the NCI and was published in 1997. This attempted to use the DiSC assay in limited stage small cell, and here

00211

we turn the somewhat, and consider the use of the test in what may be considered in its most favorable scenario, because this disease which is traditionally, and now actually standardly treated with the combination of radiation therapy and chemotherapy, would potentially treat empirically with a regimen known to produce a high level of response, and then come back after finishing consolidation with radiation sensitivity with either a chosen regimen based on the in vitro sensitivity or a standard approach using an additional three drugs that the patient had not seen previously that would be regarded as standard or part of the standard care of patients with small cell lung cancer.

And in this study, there was actually a trend towards somewhat improved survival in patients who could actually receive the in vitro best regimen, but it certainly was just a trend. And most interestingly, of the 54 patients that were entered, the minority of the patients could actually be successfully biopsied in this very, shall we say well coordinated, well resourced

clinical trials scenario. Next.

So in terms of summarizing what I list here as my own disinterested perspective on whether or not chemosensitivity testing is what one would might consider to be ready to prime time in widespread use, I would offer that from my perspective, no method has emerged as a quote-unquote gold standard, owing to methodologic variation and the definition of what constitutes resistance or sensitive tests. The unfortunate fact that one cannot get reliable data from most if not many patients. And in the few completed prospective or randomized trials, there is little assurance that ultimately there is a difference effected by the test.

What we ultimately need if tests of this nature are to be potentially useful, is probably better drugs, because in point of fact, since most of the drugs are unfortunately inactive in many of the diseases in which these tests would be used, knowing that they won't work is not actually terribly valuable.

We need a method that is applicable to all specimens obtained in real time with the diagnostic specimen; that is to say, to require a second test, or second procedure, in order to obtain the specimen, inevitably indicating or introduces potential biases in studies related to those patients who could withstand or undergo these procedures, as well as of course, making the test, the performance of the test more costly than one might potentially desire.

But on the other hand, I think the future holds potentially with newer approaches, including gene expression arrays, serial analysis of gene expression, there may be better, and hopefully more useful techniques to assess this in the future. But whatever the test, be it some permutation of a currently available test, or one of the newer methodologies here, its ultimate value should be established in prospective

randomized trials where one uses the diagnostically guided as opposed to the empirical treatment before assessing whether or not it is openly valuable.

And I thank you for your attention.

DR. FERGUSON: Thank you, Dr. Sausville. I think you've gone in shorter time than even I asked for, and so I'll open for a question or comment. Yes, Dr. Hoffman?

DR. HOFFMAN: Yes. I would like to ask Dr. Sausville his opinion about the assays that

00214

were discussed this morning, based on three dimensional culture and other new third generation techniques that address these problems and have shown to be able to assess greater than 95 percent of the patients' specimens, have shown survival benefit, have shown very high correlation to response. I would like Dr. Sausville's comments on this morning's talks.

DR. SAUSVILLE: Again, I wasn't here this morning, and indeed, my brief was not to comment on specific assays from this morning's activities, but to offer an overview of problems in the field in general. And I would certainly say that if the tests that were proposed this morning seem of interest, the real question is have they been evaluated in prospective randomized studies. Because unless they have not, or I should say until they have, one, and since as far as I'm aware, they have not, it would be, I think premature to conclude that they are, therefore, of widespread general use.

DR. FERGUSON: Dr. Weisenthal?

DR. WEISENTHAL: Now would be as good a time as any to address the issue of the requirement of prospective randomized trials for

00215

acceptance of this technology. I think it's a very important issue, several speakers have raised it, and the issue is this: Should these tests be used in clinical medicine until it has been established in prospective randomized trials

that patients treated on the basis of assay result have a superior therapeutic outcome to patients treated without the assay result? The cop-out way to answer this, which I'm not, this is not my answer to it, but what I could say if I wanted to cop out, and it's perfectly valid, is that never has the bar been raised so high for any diagnostic test in history.

Dr. Sausville began his talk by pointing out bacterial cultures done in sensitivity testing, including one of his examples was serum bactericidal testing. Serum bactericidal testing, for those of you who may know it, is something that Medicare does reimburse for. It's very controversial, it's much more controversial actually than cell culture drug resistance testing. The performance characteristics are certainly inferior based on sensitivity and specificity. And furthermore, there has certainly never been a prospective

00216

randomized trial showing survival advantage or therapeutic outcome, you know, higher cure rate or anything, whether you use serum bactericidal testing or not, or any other form of antibiotic sensitivity test.

We're talking about laboratory tests, not a therapeutic agent, and I think that one would be advised, at least first of all, to judge them on the basis of the way that other laboratory tests have been judged, and that is, do they have acceptable accuracy, sensitivity and specificity?

However, moving on to the question of the prospective randomized trial, all of us, no one more than those of us who have been working in this field for 20 years, would love to see prospective randomized trials, physician's choice therapy versus assay directed therapy. This has been the Holy Grail. I hope before I die, I will be able to participate in such a trial. I mentioned earlier, the fact is that there have been a lot of energetic, very talented people,

that have devoted their careers to this, and the best example is Dr. Dan von Hoff, who is the most energetic. He and I were clinical oncology

00217

trainees together at the National Cancer Institute. We both started working in this field in the same lab at the same time. And Dan had -- you know, my CV lists about 50 publications; Dan's CV is probably closing in on 2,000. And he's organized more prospective randomized trials and things like that. He was unable to successfully get a study initiated and patients accrued, and completed. I have devoted enormous amounts of effort to getting those trial done, and for one reason or another, they didn't accrue patients, and things like this.

I want to point out that in Medicare, Medicare has a problem, and the problem is not in the year 2000, the years between 2010 and 2015. The budgetary crunch in Medicare is going to come in 2010 and 2015 when those of use who are now 52 years old are going to be 60, 65, 75 years old, and we're going to be getting cancer. What's going to happen over the next ten years is that there's going to be an ever increasing array of partially effective and very expensive cancer treatments. We're seeing that now. Drugs are being approved at a very rapid pace. We don't have a clue how to use them.

00218

We brought up the idea about using the test as a litmus test, like should you pay for the therapy. Well, the only way that you're going to be able to ever use the test as a litmus test is if you do the prospective randomized. And I would submit to you that the way to get the prospective randomized trials done is as follows: Look at the data that you heard about this morning. Surely, you must be convinced that there is a germ of truth in this. You know, there is a consistent, overwhelming, and I think that study after study is showing that these tests do predict, they can identify the

difference between good treatments and bad treatments. So it is not much of a leap of faith to say that if only someone could do the trials, then there's a good chance that they would turn out to be positive, and if they do turn out to be positive, by the year 2010, we will have a wonderful tool to triage therapy, to triage patients, right at the time when Medicare most needs it, when the budgetary crunch comes, when we've got all these expensive cancer therapies. You know, I gave you the example of the five patients treated with the bone marrow transplants

00219

at $200,000 a patient, who did not benefit from that, who then got an assay and had a great result. What if they had gotten the assay? It has the potential to be enormously cost effective. But the only way that it will be used in that way is if you do the trials, but the only way -- it's a catch 22 -- the only trials will ever get done -- I personally believe that if Medicare approves this, it will be the shot heard round the world, Swann, ECOG, CLGB, they will be lining up to do trials. You guys, you know, come back and maybe approve it conditionally, come back in five years and see what's happening.

DR. BAGLEY: Well, you know, today -- it brings up an interesting point, and I think you bring up the comment that, you know, never has the bar been this high. Well, I would take exception to that. I think the bar is not any higher for this than for anything else that we are currently looking at. And it is not something that we are not used to hearing for other things too, and that is, gee, we're paying for things that were never subjected to any scrutiny, so why should we subject it to scrutiny. I mean, that's -- we hear it all the

00220

time, and that's just not going to work in this day of evidence based medicine. And I'll tell you because, you know, how much it costs isn't an issue that we're really here to talk about today.

Because, you know, two years HCFA reorganized and changed the whole focus of coverage, moved the coverage office away from that part of HCFA that pays for things and looks at program integrity, and moved it into the place at HCFA that looks at quality and clinical standards.

And that's exactly the focus we ought to be doing because, you know, what it boils down to is, it's not just why not pay for it, it doesn't cost that much, or it might save a little money. But it's let's pay for it because it's the right thing to do, and represents quality medicine. And when that happens, you know, we shouldn't just pay for it, we should pay for it, we should promote it, and perhaps, if the evidence is there, we ought to insist on it. I don't think the clinical community or the beneficiary community would tolerate us insisting upon a pattern of behavior, or even promoting a pattern of behavior, without evidence, and so why should we pay for it without evidence? And

00221

that's the change we're trying to make, that's been the whole point of changing the coverage process, putting together advisory committees like this, is to say, let's look at evidence and let's make decisions about what we pay for based on quality, and once we know what quality is, let's not just pay for it, let's not stop there, but let's pay for things that we are willing to promote and perhaps even insist upon. And so, that is the reason for the focus on evidence, and it's going to be there. And the fact that we may not have subjected past technologies to the same evidence, doesn't mean we can't go back and look at them, time willing, but it doesn't mean we should lower the bar for new technologies.

DR. FERGUSON: Do you want to respond a minute, Dr. Sausville?

DR. SAUSVILLE: Yes, I do wish to respond to that. And I want to thank you for that perspective, because clearly, there is nothing that ever, that doesn't lack for good

intentions. Clearly, the desire to convey useful patient benefit goes without question. And the efforts that were cited over the past two decades have really been enormous efforts in that regard.

00222

But one distinction that I must point out is that when one considers the bacteriologic analogy, the diagnostic specimen, that is to say the bacteria growing in a bottle, equals the test specimen. So that is one intrinsic difference.

In many cases, cancer related sensitivity testing requires additional efforts to get and process tissue different than the routine. So it is a point where the analogy is not exactly apt, I think. And you quoted the endocarditis issue, and you're right. It is controversial as to whether or not ultimately sensitivity testing is beneficial, because among the lethal consequences of endocarditis are a series of almost anatomical problems, valve problems, thrombi, et cetera, that are not in any way predicted or dealt with by the sensitivity testing. So again, it's -- I think that the two are, recall each other, but have important differences in thinking about the ultimate value of the tests.

DR. FERGUSON: Dr. Sundwall?

DR. SUNDWALL: Just a quick question. Dr. Sausville, I am a family physician, not an oncologist, but I was very perplexed by your

00223

statement. If I heard it correctly, you said, knowing what drugs won't work is not all that helpful. I don't understand that, given the morbidity and the difficulties with chemotherapeutic agents. I've had many patients suffer terribly from chemotherapy, and how can you say not knowing what won't work isn't that helpful?

DR. SAUSVILLE: Because the context in which -- and I respect your point, and I don't certainly mean to in any way imply a lot of sole searching on both the part of physicians and

patients that goes into the decision to entertain therapy. But in oncology, frequently the treatment is driven by the histologic diagnosis, so if for example the initial diagnosis of small cell lung cancer, if one could have a pattern of drugs that have more or less a susceptibility, I am not aware that such tests would be considered definitive in saying, well, because you happen to have a resistant small cell lung cancer, you should not receive any therapy. So in that case the therapy, or choice of therapy, is ultimately driven by the histologic diagnosis that's apparent. Consider the opposite point. Somebody

00224

with a chemotherapy refractory neoplasm, manifested, such as pancreatic or renal, which are problems which as far as I'm aware, are not considered responsive to any set of agents routinely.

Again, the information of whether the patient has that dire situation is implicit in the histology. It's not clear that any tests that can be done ultimately defines a drug that can change the outcome that is at the present time ordained by the histology. So I take your point, that being able to reliably choose drugs that convey a useful clinical benefit is very worthwhile and a goal that should be pursued. I am not sure that the current tests actually allow the clear delineation of such agents.

And in that regard, you can tell a patient who has the unfortunate diagnosis of pancreatic cancer, that they are likely not going to respond to a medicine chosen on that basis, or chosen after having gone through an additional test to obtain tissue and then tested for assay resistance.

MR. KIESNER: I think Dr. Sundwall asked a very interesting question, and I think

00225

there are at least two clinical strategies for using this type of information. I think on one hand you can say, we're going to select a drug,

and another, there may be a different clinical setting, and I will give you two examples. I think this is very important.

Dr. Alberts spoke this morning about a clinical situation where he would be referred a patient from another hospital, and that patient may not, may be unaffected by the primary care, he has relapsed, the tumor is growing, and they send him, they send the patient to him. Doctor, what can you do to help me? In that situation, there may be three or four or five different drugs, single agents, none of which have been determined to have a significant clinical benefit over the other drug in that situation. If I am a patient and if any physician can tell me of the five drugs, Frank, two of those drugs you're resistant to, what has he told me? He's said, I'm not going to use those two drugs, I've saved you from the possibility that you're going to get those two drugs and not benefit from it. It's very, very well documented that these tests are able to identify resistance, and if I'm a patient

00226

and if my physician in that setting can identify the resistance, I believe he has done me a real service.

The second situation is one which I experienced personally. And I'm not mentioning it because it's personal, I'm mentioning it because it's exemplary of the position that a lot of families can be in relation to elderly Medicare patients. In -- I'm from Minneapolis. My father was in St. Mary's Hospital. He was ten years past Medicare age and was being treated for cancer. We saw what the drug was doing to him. If his physician could have come to me and said, Frank, I have two or three other drugs, two or three other choices I could try, and I have done a test and I could see that they are all resistant, I don't think we should go any further. From the family situation, it was a very difficult situation to make, do you go further. We made the situation not to. But to

this date, if I would have had an assay that would have told me the drugs that the physician was considering will not work, I would feel I would have been served, our family would have been served, and my father would have been

00227

served. Elimination of drugs, identification of drugs in those types of clinical settings that don't help, or help you stop therapy, I think is something worth considering.

DR. FERGUSON: Thank you. Very briefly.

DR. SAUSVILLE: So my response to that is the essence of the issue, and it also pertains to the question before, is whether or not one could have reached the conclusion that drugs would not have benefitted your relative by the diagnosis itself, and not have ultimately had to rely on a test. And here the performance properties of the -- the unfortunate performance properties that when a drug is predicted to be sensitive by these tests, the outcome is unfortunately not any different in many cases, than when things -- and in fact, in all cases that I'm aware -- of when drugs are seen as resistant, is the essence of why we are in a quandary about how to appropriately use this.

DR. FERGUSON: Thank you. Harry. Dr. Handelsman?

DR. HANDELSMAN: I'm Harry Handelsman. I'm at the Center for Practice and Technology

00228

Assessment, Agency for Health Care Policy and Research, and our office was asked by HCFA to review the 1990 article by Kern and Weisenthal on the use of suprapharmacologic drug exposures. And I'm going to briefly synthesize what I think was the essence of that article and then give my personal critique. Unfortunately, some of this is going to be repeating some of the data that you heard earlier today, and that's unavoidable.

Bayes' theorem suggests that drug sensitivity testing in vitro will be accurate in

predicting clinical drug resistance in tumors with high overall response rates only if the assays have a specificity of greater than 98 percent for drug resistance. A 1989 review of the literature by the authors indicated that a 30 to 50 percent false positive rate, and a false negative rate as high as 15 percent.

This reported assay, which was developed by Kern, uses a soft agar culture with products of concentration times time higher than those which can be achieved clinically and used drug exposures 100-fold higher than other contemporaneous studies. Response assessments were made by retrospective and blinded chart

00229

reviews. The authors reviewed 450 correlations between assay results and clinical response over an eight-year period. The assay was calibrated to produce extremely high specificity for drug resistance. Two assay end points were used, colony formation and thymidine incorporation.

Overall response rates were 28 percent using the colony formation end point, and 34 percent using the thymidine incorporation end point. At the assay lower cutoff value, the assay was 99 percent specific in identifying non-responders, fulfilling the Bayes prediction. Patients with drug resistant tumors could be accurately identified in otherwise highly responsive patient cohorts. The demonstration that the post-test response probabilities of patients varied according to assay results in pretest response probabilities allowed the construction of a nomogram for predicting probability of response.

In 1976, it appeared that no method of predictive testing had gained general acceptance, and during the subsequent decade, high false positive and false negative rates continued to plague the field of in vitro testing.

00230

The clinical advantages of developing a highly specific drug resistant assay include:

The avoidance of the use of inactive agents in treating responsive tumors; the avoidance of drug related morbidity of inactive agents; the identification of drug resistant tumors for timely consideration of alternative therapies; and obviously, the cost savings of avoiding the use of ineffective agents. Alternative assay methods are available. However, the use of cell culture assay had the advantage of measuring the net effect of both known and unknown mechanisms involved in drug resistance.

It is indeed possible to estimate the post-test response probability for specific drugs in specific tumors and patients. This can be achieved through the determination of assay results and the application of a constructed nomogram for assay predicted probability of response.

In general, efficacy studies in both in vitro and in vivo tumor models provide an opportunity to obtain data on both efficacy and toxicity, and to refine dose and schedule information for clinical trials. In vitro

00231

testing has been extensively applied to determine the potential efficacy of individual drugs and remains an attractive alternative to testing empiric regimens in phase I and phase II clinical trials. In vitro testing can differentiate active and inactive agents, but cannot serve as a substitute for in vivo studies, despite providing elements of both positive and negative predictive reliability.

Combinations of agents, which are the most widely applied treatment strategy, are best evaluated using in vivo models, where both toxicity and pharmacokinetics can be adequately studied. Although in vitro assays can provide primary drug resistance date, the most relevant outcome from such assays is improved patient survival, and there have been no clinical trials demonstrating such a result. In addition, it remains to be determined if in vitro testing will

be found to have direct clinical applications for disease or patient specific therapies. There have been encouraging reports of survival advantage of patients treated with in vitro directed therapies, but these require confirmation from larger numbers of patients and

00232

variety of tumors.

Both randomized and non-randomized studies comparing tumor responses to chemotherapy selected by in vitro testing with empirical chemotherapy have produced conflicting results. Response rates appear to be better with in vivo selective agents. However, the impact on survival has not been adequately addressed. Ideally, in vitro assays should be correlated with both response and survival data. The most significant issue in the realm of cancer chemotherapy is that of the resistant mechanisms. The ability to identify ineffective agents in these assays, albeit potentially important, does little to elucidate the mechanisms problem. The assay described in this article can perform its intended task of identifying resistant tumors, and determining a probability of response, but its clinical utility has not been established.

DR. FERGUSON: Thank you. I think we have time for a few questions. Panel, or others? Comments? Yes?

DR. FRUEHAUF: I think that was a nice summary of the paper and I think the issue of

00233

survival was addressed this morning.

DR. HANDELSMAN: Excuse me, if I can interrupt. The issue of survival was predicted, but it wasn't on a comparison with alternative therapies.

DR. FRUEHAUF: That's true. It was survival in a blinded prospective way, looking at people just getting empirical therapy and asking the question, if you're not going to respond, will you have inferior survival? And we've

addressed the issue of, do you want to get a drug, as a person who has cancer, that won't work and won't benefit your survival? And I think that's the important point that this paper is establishing, the utility of knowing that a drug will not be of benefit to a cancer patient.

And I have dealt with neuropathies that are incapacitating to professional tennis players. I have had to deal with all sorts of toxicities, and many of these people progress through therapy and die of their disease, and the quality of their life during that progression was significantly adversely affected by getting ineffective therapy. So the clinical utility question to me as a practicing physician, is to

00234

not harm people with ineffective therapy that will not, which has been demonstrated not to benefit their survival. And I think most people understand, if you don't respond to the therapy, you're not going to live longer. And we're not trying to say that the test will predict a drug that will help people live longer, for drug resistance assays. We are trying to say, and you stated, and Dr. Sausville stated, that these assays accurately predict drugs that will not be effective. And then the question is, so what? I think the answer to the question so what is, so I don't want to give those drugs to my patient.

DR. FERGUSON: Thank you. Okay. I guess we can go on. Harry, thank you very much. Dr. Burke?

DR. BURKE: I think we're going to have a lot of fun this afternoon. My name is Harry Burke. I'm a consultant to HCFA. I'm an internist. I'm a methodologist. I'm only here for today.

The first couple slides that I am going to present are not HCFA's position, they're my personal review on the subject, and shouldn't be considered HCFA's policy. I am going to address

00235

three issues today. First, the levels of

evidence, which has been raised several times by various speakers. I'm going to talk about test accuracy. And then I am going to talk about the Kern and Weisenthal article that Dr. Handelsman just gave us an introduction to.

But before that, I would like to make a couple comments. First, the extreme drug resistance is really a therapy specific prognostic factor. It really has to be looked at in the context of other therapy specific prognostic factors. Dan Hayes was right. It's like ER and PR, and these other factors. And there's a scientific rationale underlying therapy specific prognostic factors that must be dealt with. The utility of the test depends on the characteristics of the test; that's clear. But it also depends on the efficacy of the treatments if it's a therapy specific prognostic factor, because they're inextricably linked together, and you can't separate the two. And it depends on the prevalence of the disease or the resistance in the population under study. So it's really those three factors together that must be taken into consideration when looking at something like

00236

this.

Let me make another point, and that is, when we talk about the utility of this test, we can't be talking about the utility for individual patients. We have to be talking about the utility of the test for a population of patients. So we can't switch back and forth between the two, because we're really mixing apples and oranges when we do that.

I'd like to make a couple comments about what has been said earlier. Fruehauf, Weisenthal and others have suggested consistent findings across studies, and they've made a claim that that proves something. And I would like to suggest that, yes, consistent findings across studies can be due to robustness of the underlying phenomenon. But it can also be due to consistent biases across the studies. And so if

you are going to make a claim for consistency of 35 studies, that all suggest the same thing, and that's a robustness claim, you'd better be prepared to tell me why it isn't due to biases in the 35 studies themselves. So you really have to look at each of the 35 studies and you have to ask the question, are these really consistent?

00237

You can't just wave a hand and point to 35 studies. You have to rule out the alternative hypothesis.

Secondly, I'm a little confused. Kiesner pointed out that we could use this task at the bedside at the discretion of the patient, or I mean of the physician, while Kern suggested that it could be used to deny a particular drug. And I need to know, which is it? Is it that the evidence is so convincing that it can be used to deny a particular therapy, or that it's not that convincing, and it's just one of an armamentarium of tests that are available to the physician.

First, let me just do a little background. Comparative clinical benefit is what I'm looking at. This is my gloss on reasonable and necessary. It could be defined as the test or treatment providing a measurable improvement over all the current relevant tests and treatments at a cost commensurate with the measured improvement. I also suggest that FDA approval is prima facie evidence of safety and efficacy, but if that isn't there, I think safety and efficacy must also be demonstrated. And a comparative clinical benefit study of a

00238

prospective, or of the test or treatment, must compare itself to the other tests or treatments. It doesn't stand alone. And so when you say well, this test or treatment is really good, you have to say what the other tests and treatments are, what you're comparing it to.

Now I would -- I am not totally a believer in randomized clinical trials for everything. My suggestion is that there may be

three levels of evidence that can be adduced: A strong evidence, which is either a large prospective randomized clinical trial, or two large retrospective studies where one study independently replicates the other study. I think that's good science as well. Or two medium size randomized prospective trials. I think all of those would be strong evidence. If I saw a really large retrospective study that was independently replicated by independent investigators, independent institutions, I would take that as fairly strong evidence.

Moderate strength are medium sized prospective trials, a large well designed retrospective study that hasn't been replicated, or two medium randomized prospective clinical

00239

trials, medium size.

Weak evidence. Small properly designed and implemented prospective randomized trials, I think are weak evidence, and I think are well recognized as that, and I think Pito and others have suggested meta-analysis to overcome the weaknesses of small randomized clinical trials. Alternatively, two medium sized retrospective studies that were done by independent investigators might be good evidence.

But insufficient evidence, small systematic studies, I consider them really exploratory rather than evidence. Case series, I think are well considered as anecdotal. And any study that's not properly designed, implemented or analyzed must be considered fatally flawed.

Large is 500 patients, medium sized, 250, small is less than 250. You know.

Okay. Test accuracy. What is a properly designed, implemented and analyzed study? Well, test accuracy of course is an association between each patient's predictions based on the test, and each patient's true outcome. That's test accuracy. The factors that affect test accuracy include the study

00240

population, were the patients who were selected easy to predict. Because you can select patient populations, and we'll get into that later, that are very easy to predict by just about any test. The test characteristics: Was the test assessed in the clinical setting which it's intended to be used for? The reproducibility: Does the prediction variability increase across laboratories and reagents? And finally, the method of measuring the accuracy: Was the correct method used?

And I'm going to focus on two of these, the first and fourth, which are the most problematic.

Okay. Sample size, or study sample characteristics. The composition of the study, the study population, makes a difference in the observed accuracy. A sample with only extreme cases, i.e., the predictors are extreme values of their range, will be easier to predict than a sample with many intermediate cases, the predictions are mostly in the middle of their range. For example, for women with breast cancer who have many positive lymph nodes, their outcomes are fairly easy to predict. Women with

00241

metastatic disease, their outcomes are pretty easy to predict. It isn't a hard task to do. What is hard to do is to predict the women with small tumors and with no lymph node involvement or metastatic involvement, that's really tough to do. So, if you just pick an extreme population, it turns out those are pretty easy predictions to make, but it turns out that most patients aren't in the extreme, so it's relatively unuseful. Okay? Thus, the sample must be representative of the real world in which the test is to be used.

Measurement of test accuracy. There are several ways to assess test accuracy. The correctness of the accuracy assessment method depends on, so when you select a method of test accuracy, whether there is a preexisting threshold, in other words, is there something out

there that says everybody above this should be positive, everybody below should be negative, does that already exist, or do you have to construct it? The number of tests to be assessed. And whether the assessment is performed on one population or more than one population.

And just very briefly, this is really

00242

hard to read. I can't get the lines on tables to work out for me, so this is lineless. But it turns out that the sensitivity and specificity pairs are really, have one threshold, they do one test, and its one population. Okay? Positive and negative predicted value, there is one threshold, one test, and two or more populations, because really, the positive and negative predictive value we're talking about are different prevalences, therefore, different populations. And the area under the receiver operating characteristic includes all thresholds, two or more tests are assessed, and one population. So in other words, we use the ROC as a best unbiased measure of test accuracy.

In terms of the measures of accuracy discussed above, without changing the test itself, there are only two ways to change the accuracy of the test. One way, of course, is to change the threshold of the predictions, and then your sensitivity and specificity would change. And the other way is to change the prevalence of the disease in the population, because then your negative and predicted -- positive and negative predicted values will change. Okay?

00243

Prevalence's effect on accuracy. The optimal prevalence for assessing the accuracy of the test is to use a population composed of 50 percent disease, 50 percent unaffected. In this situation, the prevalence itself provides no advantage to the test. As the prevalence departs from 50-50, the impact of predicting the prevalence becomes more prominent. In other

words, if the test acted as a naive Bayesian classifier, then for each patient it would always predict the most frequent outcome, in other words, it would predict the prevalence. So for example, if there was a 90 percent prevalence in a diseased population, then the naive Bayesian classifier would say disease every single time for every single patient, and you would be right 90 percent of the time. That's pretty good. Okay? That's a pretty accurate approach. As the proportion of patients with or without the event moves, either toward a hundred percent or zero, the naive Bayesian approach becomes more effective, more efficient in its predictions. So it's only at 50-50 for binary outcome, that you neutralize the naive Bayesian classifier approach. In other words, if the true prevalence

00244

of the disease in a population is close to a hundred percent, it's almost possible for a test to add predictive information. Okay? That's really an important idea. So, as you get towards high prevalences, almost no test will be helpful anymore. Okay?

Changes in the prevalence of the disease in a population, as reflected by corresponding changes in the test's positive and negative predictive values. If one were allowed to report the positive predictive value, or the negative predictive value of tests just by itself, then one might be tempted to create or select a high prevalence population for assessment of the test, because the test would appear to possess a high predictive accuracy, okay? Until of course, it was compared to the naive Bayesian classifier, at which point it would cease. Thus, both the positive and negative predictive values of the test must be assessed. Then, if the prevalence is not 50 percent, the test must be compared to the naive Bayesian classifier. Further, both the sensitivity and specificity of the test must be assessed in terms of the cutoff that was

00245
selected, and the prevalence, because it turns out that although it's commonly thought that prevalence doesn't affect sensitivity and specificity, it certainly does, and there are a number of papers that demonstrate that.

So, a better way to assess the accuracy of the test is to use the ROC. This measure of accuracy is impervious to changes in prevalence and reflects the characteristics of the test across all sensitivity and specificity pairs.

Well, okay. So now, I was asked to take a peek at Kern and Weisenthal's paper as well, and it turns out that they really are very sophisticated in their use of data and results. It's probably one of the most sophisticated papers I've ever read, and I have read quite a few. I'm going to talk about those areas of the paper.

Overview of the study. Kern and Weisenthal used two in vitro tests, which have been mentioned, as surrogate outcomes for response to chemotherapy in patients with different types of cancer. If a patient's tumor demonstrated drug resistance in a test, i.e., after the patient's tumor cells were exposed to
00246
the drugs for a certain period of time, and the cells did not achieve a threshold inhibition, the test was interpreted as predicting that the patient would not clinically benefit from receiving the drug.

So, we go back to our levels of evidence, and we can ask overall about this study, where it would lie in our levels of evidence? Well, the colony formation test is really Level III, it's weak evidence. The thymidine incorporation is really Level IV, insufficient evidence.

But, not letting that bother us too much, let's talk about the study itself. It's a retrospective chart review, subject to several biases, including therapy selection bias, who

received the therapy, and study selection bias, which patients were included in the study. And the study was not validated on an independent population, but it was done on the same population. It was done from 1980 to 1987 in the United States.

The study characteristics. Initial population was 5,059 patients. From that, they winnowed it down to 450 patients that they

00247

actually studied, about 9 percent of the initial population. They looked at eight different types of cancer. They had 332 colony formation patients, 116 thymidine incorporation patients. And the non-respondent prevalence was 71 percent of the population.

One thing that the study wasn't very clear about, it said that virtually all patients were treated with standard chemotherapy, but then later on it said, most of the patients whose specimens were analyzed did not receive chemotherapy because they underwent curative surgical procedures. And I didn't understand that distinction.

We'll assume that all 450 patients in the study received chemotherapy. The percentage of patients who receive chemotherapy today may actually be much higher. The criteria to decide which patients received chemotherapy is not reliable. This is really not a very acceptable approach to a study. If in fact you're going to predict who's going to respond to chemotherapy, I think you really have to say how chemotherapy was selected, what the selection criteria were.

Also not provided were the patient

00248

characteristics of the study population, and this is really critical information. For example, if the population was composed of patients who had already received primary chemotherapy, had incurable disease, and were undergoing salvage treatment, then this study would not be applicable today, and in addition, the results

would be biased. So we really need to know what the chemotherapy selection criteria were, and what the patient population characteristics were, neither of which are provided to us. There is no basis from which to understand the results that we are seeing.

Now, the function of the test is to predict clinical non-response to chemotherapy using suprapharmacologic drug doses. Now, we're interested in the non-response rate per drug per cancer type per test type. That's what we're interested in. So there were eight drugs. Now I'm not going to talk about combination therapy because that's a whole other subject. There are eight drugs, eight cancer types, that means there were 64 bins, okay? So that means per cancer, per treatment, so there were 64 of those combinations for each of the two tests, for a

00249

total of 128 accuracy assessments. And excuse, the lines aren't there, but you see 64 bins. And for each bin, you would want to know prospectively, hopefully randomized, you would want to know, for disease one, treatment one, what does the test say, okay, about this population? In that one cell. And then you would want to follow that population over time and see what actually happened to those people.

So for breast cancer and a particular chemotherapeutic agent, you would like to see, did the test predict for that chemotherapeutic agent for breast cancer, successfully. And you'd want to do that for each of the 64 cells. And in fact, you must do it for each of the 64 cells.

If there were the same number of patients per cancer type, then the 118 patients tested for thymidine incorporation would be at 1.8 patients per bin, for this study. And for the 332 patients tested for colony formation, there would be 5.2 patients per bin. These frequencies would be too low to be meaningful.

Now, out of the eight drugs tested and reported, the only drugs to use today, and are

they not used in combination, the efficacy, the
00250
efficiency of these tests must be demonstrated with each chemotherapeutic agent in use today, and for each combination of agents, each type of cancer.

Now, just a couple final points. It's unclear why this study provided two sets of thresholds instead of one. Further, although two thresholds were tested for significance, three were presented in the text, shown in the tables and figures. The first threshold is 45 to 75, and the second one was 15 to 40. In this study, the thresholds that were selected to assess on the same population that were used to determine the optimal thresholds. This elementary mistake, reporting the results from the population used to create the threshold, rather than the results of an independent population, always results in the overestimation of test accuracy.

The outcome was standard response criteria. We are never given a definition of what standard response criteria are. We don't know who got the chemo and why. We don't know the study population. We don't even know the outcome. We are never given definitions of any of those three. It's absolutely critical that
00251
the specific response criteria employed by the investigators be revealed if that is their outcome.

Now of course, Rich Simon, who many of you know at the NCI, and others, have pointed out that response is an unreliable outcome and should be avoided if at all possible. So, okay. So rather than the 64 sets of results that we were looking for, two sets of results were presented, one for each of the two thresholds. Each of the results is across all eight cancers, all eight therapies, and the tests, and the results are there for the first threshold, 60 something percent sensitivity, 87 specificity, 43 and 99. Clearly, the sensitivity goes down as the

specificity goes up. Neither sensitivity or specificity pair is very high. Combining all results into one conglomeration provides no information regarding the utility of the test for each drug in terms of each cancer type. The study should have reported the area of the ROC curve, both tests, for the 64 sets of results. Thank you.

DR. FERGUSON: Thank you very much. We're actually at our time for a break.

Perhaps -- it is almost 2:30. If we take a 15-minute break, I think, yes, would you please come back up, because there may be a couple of questions for you.

(Recess from 2:25 p.m. to 2:45 p.m.)

DR. FERGUSON: I wonder if there are any in the audience, or panel for that matter, who would like to ask Dr. Burke some questions related to his presentation? And also, Dr. Burke has promised to give us the last few slides. Are there questions for Dr. Burke from members of the audience or from the panel?

Dr. Weisenthal, did you have a question that you wanted to ask, or a comment?

DR. WEISENTHAL: I want to thank Dr. Burke. He started off by paying me compliment, and he said of Dr. Kern and I's paper, that this is one of the most sophisticated papers that he's ever read. I've also been talking to critics of these technologies for 20 years, and that's the most sophisticated criticism that I've ever had, so I want to congratulate you on that.

There are several points that were raised in your talk which should be addressed.

Just to begin with, the study by Kern and Weisenthal that you spent the bulk of your time reviewing, just to begin with that, you brought up several methodologic criticisms and raised questions about patient selection and so forth. I want to remind everyone here that that was

published in the Journal of the National Cancer Institute. I assure you it underwent rigorous peer review. When we submitted our first draft of the manuscript, the reviewers there had certain problems with it and they had certain things they wanted clarification of.

DR. BURKE: But that's an appeal to authority.

DR. WEISENTHAL: No, no, no. Dr. Burke, had you been one of the reviewers, no doubt you would have raised those issues at the time and we would have responded to those. And I'd like to ask Dr. Kern now if he can respond, so we're in consideration of that you were one of the -- you know, you can't blame us because you were not the reviewer of our paper. Had you been there and helping us to get the essential information out there, I'm sure it would have been a better paper. But we'd like to address

00254

those issues that you raised at this time, if that's okay.

DR. BURKE: Absolutely.

DR. KERN: One of the points was the selection bias. How could you end up with 450 correlations out of 5,000 patients in the study? Well, the 5,000 patients was an overview of all the tests that we had done in the laboratory. It wasn't meant to imply that the clinical study was based on 5,000 patients. And in fact, at the Department of Surgery, UCLA, where I was, most of the patients were treated with surgery or radiation, not with chemotherapy.

Secondly, many of the patients that received chemotherapy received adjuvant chemotherapy. So the inclusion criteria of the study to get to 400 patients included, first, patients had to have advanced disease; second, they all had to have objectively measurable disease, either by CT scan, x-rays or so on. Okay?

Now, as far as another comment that you made about one study, but it's not been

independently validated, I think I may ask Dr. Bosanquet to address that issue, because he published an article in Lancet a couple of years after our paper.

DR. BURKE: Did you want to address any of the other issues that I brought up?

(Inaudible response from audience.)

DR. BURKE: I mean, this is not an opportunity for us to get into whether the study has been validated or not at this time. That was just an issue that I raised, and perhaps at another forum that can be addressed further. I think we have time limitations.

DR. KERN: Well, I will try to answer.

DR. BURKE: So keep going. There were a lot of issues.

DR. KERN: Bring up a couple of the issues, remind me of them. Let me see what you consider a serious objection.

DR. BURKE: Well, the selection, the patient characteristics, the criteria for who got what treatment.

DR. KERN: Let's go one at a time. Who got what treatment was determined independently, not by the assay, but by the disease type. The patients went on standard protocols. Most of the patients who ended up at being UCLA, an academic center, were all on some sort of clinical trial, randomized trial protocol.

DR. BURKE: What were the standard protocols? What was the response criteria that you used?

DR. KERN: The response criteria were the ECROG criteria of partial response and complete response.

DR. BURKE: And what percentage was each in terms of your study?

DR. KERN: I'm sorry, I don't understand.

DR. BURKE: In other words, in terms of response, global response measured, and what

percentage of these patients were partial responses, what were complete, and then at that time, how were those defined in your study population?

DR. KERN: Okay. The responses were, again, just by objective measurements. It was retrospective, but scans, x-rays. And the complete response, obviously, complete disappearance of the disease. Partial response was by the criteria of two dimensions and the shrinkage of at least half in two dimensions.

00257

Standard criteria.

DR. BURKE: But this was a retrospective study where you went back to the charts. We all know about the paucity of information and the error of information, and in follow-up information not being in the charts. How did you manage those issues in your retrospective study?

DR. KERN: Well, the follow-up -- obviously, there are problems, and I'm not trying to say there's not biases in it. We all know the disadvantages of retrospective chart reviews. The only thing I can tell you is what actually was done, two oncologists reviewed the charts and made their best decisions of what the responses were, based on measurable criteria.

DR. BURKE: What did they do when they disagreed? What did they do when information wasn't there? What did they do make sure it was accurate information? Do you want to continue with this?

DR. KERN: No, I cannot say that I can answer every question. I mean, I'm not an expert in your field.

DR. FERGUSON: One brief, and then

00258

we'll let Dr. Bosanquet speak.

DR. WEISENTHAL: This is really important, okay? You know, you talked about an eight-by-eight table, and we only have one point. You know, a study like this is never

going to be done again in the history of the world. Never again are you going to have 330 patients treated with single agents. The important thing about it was that this was an honest blinded study in the following fashion, and that is that the clinical results were determined independent of knowledge of assay results. The clinical results were reported to the Department of Biomathematics at UCLA; they were like the stakeholder in this, they had the clinical assessments. Likewise, they received independently from the laboratory the laboratory assessments, and then the correlations were made as stated.

DR. BURKE: Let's just deal with that issue for a moment, because Dr. Kern sat down and you stood up. So we've got the 64 bin table, right.

DR. WEISENTHAL: Yes.

DR. BURKE: And the issue is, how do we

00259

know the utility of this test for a chemotherapy in a disease?

DR. WEISENTHAL: Okay. You're making the same criticism as Maury Markman has made. What Maury Markman says is as follows, and he says that he notes that there have been no prospective randomized trials.

DR. BURKE: That's not my question.

DR. WEISENTHAL: Wait a second. But it's the same thing. He says that even if some day there were to be a prospective randomized study, that that would only apply to that one particular situation, and it would not tell you anything about all the other situations.

You know, the sort of information that you're asking for in the real world will not be available for 20 to 50 years, if ever.

DR. BURKE: No, no. I understand the mitigating circumstances. But the question is, if you want this test to predict a particular chemotherapeutic regimen in a particular disease, then I want that information, and I don't have it

in your study.
DR. FERGUSON: Okay. I am going to ask for Dr. Bosanquet to give his response, and then

00260

we are going to go ahead. We will try to have a little more time at the end.
DR. WEISENTHAL: There's an extremely important point. Basically he started out his talk denigrating -- in other words, I made the point that we have 35 studies consistently showing the same thing, and he denigrated that, and he said, oh, that's just due to consistent bias. And I would like to prove to you that that is not true.
DR. BURKE: Excuse me. I didn't. I posed an alternative hypothesis. I said there are two hypotheses for the 35 consistent studies. Assuming that they are consistent, which we have no evidence of, but assuming that they are consistent, it could be due to two things. It could be due to the fact that there is a phenomenon there, or it could be due to consistent study bias. And until you eliminate the alternative hypothesis, you haven't done science.
DR. WEISENTHAL: I would like to then eliminate the alternative hypothesis and prove to Dr. Burke that we have indeed done science in this setting.

00261

DR. FERGUSON: Can you do that in the final hour?
DR. WEISENTHAL: Okay.
DR. BOSANQUET: It was stated that there was no independent validation of this. We actually took the data that we published in the Lancet the following year. This paper that we're discussing is 1990; we published this work in 1991 in the Lancet, using CLL patients. And we also looked at extreme drug resistance in these patients.
We got this. We found 22 of 119 patients had extreme drug resistance in vitro,

and none of these patients responded. So here is one of the things that we would speak to, which was independent validation in a completely different set of circumstances in a different laboratory, and we find exactly the same thing, extreme drug resistance, no response.

DR. FERGUSON: Using the same cutoff points that were determined by Kern Weisenthal, I guess.

DR. BURKE: Just to respond briefly to that, two points. One, that is not a replication of the 1990 paper.

00262

And number two, I do suspect that that's exactly correct, that it is disease and treatment specific. And that's exactly my point. That's exactly my point. You have to talk a specific disease, a specific treatment, how does the test do? Not a conglomeration of diseases and treatments together. That's exactly my point. Thank you.

DR. FERGUSON: Thank you. Dr. Burken.

DR. BURKEN: Hi, everybody. Can everybody hear me okay? I am Dr. Mitch Burken, a medical officer with the coverage and analysis group at HCFA. What I'd like to do is try to tie together some of the presentations from earlier in the day. There will be a lot of material in here that you've seen before, but what I want to do is try to wrap it up, and wrap it up in a way that's consistent with Dr. Bagley's opening remarks around 8:00 this morning, looking at the broad sweep of the evidence, not spending as much time on specific papers as much as trying to see the bigger picture, cutting across many assay formats.

Well, as I said, for the first several minutes I want to be as conceptual as possible,

00263

and then we'll get more into the bulk of the evidence itself.

But let's think about why we would order any type of lab test, okay? A lab test has

its maximum clinical utility when the disease probability is most uncertain. In other words, we heard a little bit about the 50-50 point, and the naive Bayes condition, and so forth, but let me just try to restate that in a slightly different way. If we have any type of lab test we're looking at, and exploring questions of clinical utility, okay? What's the probability that the patient has a disease? If a patient is very unlikely to have the disease, then what kind of information do you have when you get a lab test result? It's certainly not very very high.

And the reciprocal situation, where we have a very very high probability or prevalence of disease, and then the lab test doesn't really add a whole lot, because we're almost positive the patient has the disease. It's when we're unsure of ourselves, and when we are at that 50-50 point, that's when a lab result can really begin to add value.

Well, where are we in the Medicare

00264

program? The panel is charged with trying to demarcate what's reasonable and necessary with respect to human tumor assay systems, okay? And we need to find a spot in this, or we need to kind of bracket an area of this graph where lab testing -- and again, we'll talk about the HTAS in a second. But where is lab testing most reasonable and necessary? Where does it add information?

Let's talk about now about applying this more generic situation to human tumor assay systems. Well, let's talk about the chemosensitivity scenario. We talked all day about how this testing can assist clinicians in selecting effective single agents. Okay? Conversely, the chemoresistance scenario is where this assay, or where these assay systems can avoid ineffective agents. And what's our reference here? The reference is data from published clinical trials; maybe they're in peer reviewed journals, maybe unfortunately they're

just in abstracts that are available at ASCO meetings. But again, there is information from clinical trials that does provide a backdrop against which one can look at this lab testing

00265

and the added value thereof.

So let's go back to our graph again. In vitro testing has the greatest clinical utility when the presumed sensitivity or resistance, because remember, they're really reciprocal functions of each other, is most uncertain. And going to our X axis here, what's the real question? The question is, is tumor X sensitive or resistant to drug Y in patient Z? Therefore, we need to be specific as to what questions we're posing.

Well, let's talk some more about clinical utility, because as I said, what we want to do is look at the broad sweep of the evidence. We've talked about different outcome measures today; we've talked about clinical response; survival; we've talked even a little bit about quality of life, although in the packet of materials there is really not a lot of quality of life literature to discuss, so we won't really get into that.

And in looking at clinical responses and outcome, we need to identify robust two-by-two data using a valid gold standard, and from there we can look at different performance

00266

measures. In this case, I think it's valuable to look at positive predictive value as a marker for chemosensitivity, and negative predictive value as a marker for chemoresistance. One could also talk about the sensitivity or specificity, but let's try to keep it just a little bit simpler, let's focus in on some concepts, and not worry so much about the math. There are others in the room who may be more expert, but let's try to keep it simple, and not get too wrapped up in the numbers, but let's try to get wrapped up until the themes and the concepts.

As we discussed earlier this morning, as Dr. Bagley emphasized, we have to insure that the biases, such as insufficient sample sizes, don't substantially influence our results.

Going back to our graph now, in 2-D rather than 3-D, let me emphasize a point that I've said already, but let me reemphasize it again. That if you are at the extremes of this utility function, okay, where the lab test, whether it's human tumor assay system, or a serum sodium, or whatever it is, or a chest x-ray, any type of diagnostic test, if you're at the extreme regions of this utility function, it doesn't

00267

really matter what your predictive values are. If the predictive value is high, it can be offset by the fact that you are in an extreme region of the utility function where those numbers don't really mean as much. Okay? And we'll talk more about that. Okay?

Well, what kinds of measures do we need to evaluate test accuracy? The ones that I talked about below, predictive values, but there are also sensitivity, specificity, area under the ROC curve, but let's talk about something else. What about some of the physician concerns. In the lab, what kinds of things can a physician tell his or her patient when a particular tumor can or cannot be assayed by the lab?

On the right side of the slide are what I would call the quality control measures, and we're not really going to spend time on those in this particular, at least my particular presentation, but you've heard from FDA earlier. So let's just kind of stay on the left-hand side of the slide for now.

Well, just to get back to a couple of those issues that really cut to the heart what physician concerns might be, you know, is there

00268

sufficient assessability or evaluability of the tumor cells from the submitted specimens? And we found out that some of the earlier clonogenic

assays had very very -- had relatively low assessability or evaluability rates. But let me pose another question.

Even if a particular assay format is evaluable 90 percent of the time, it still might mean that 10 percent of the time the physician speaks with his or her patient, and they really just can't get an adequate result. So I think that's an issue. You know, even if it's 90 or 95 percent, there is still some percentage of the time when you don't have a result and you come back. Okay?

There are other issues that come into play. What's the effect of tumor heterogeneity? We talked a little bit this morning about tumor heterogeneity, but there's another type of tumor heterogeneity as well, and that's the type that can occur within the same patient, so that a primary tumor and its metastatic lesions have different in vitro patterns. And again, that is a consideration to keep in mind when we're thinking about this type of testing.

00269

So as I mentioned, in vitro results for solid tumors from one site may not always provide the same result as other sites. However, there is a paper back in 1986, we'll get to it a little bit later, I'll touch on it again, but it shows that this problem may not be quite as pronounced in clonic lesions such as leukemias.

Well, there is a whole host of in vitro assay formats and we can, as I say, just kind of go through those. But I think it's important to mention at the end here that we will not in this presentation be going through the clonogenic literature. When we reviewed this material at HCFA, we didn't feel there would be a lot of value in discussing the older technologies that had the lower evaluability rates, and just felt it would be better to present it to the panel this way.

Well, what kinds of criteria do we use to evaluate the literature? Again, our goal here

is to be broad based. We looked at peer review journals in English. There were some manuscripts pending publication that were necessary for panel discussion. There were a couple of the assay formats that were relatively recently developed

00270

that we felt we would not be fair to the requesters if we excluded some of the manuscripts. There was, for example, Bartels chemoresponse assay was FDA approved back in 1996 and package inserted in the summary of the safety evaluation data was included as a way of evaluating that. And we did not look at abstract data.

Based on that, what types of additional search methods? Well, we -- again, we looked at articles submitted to HCFA prior to November 1st, 1999.

DR. FERGUSON: Mitch, can you speak into the microphone?

DR. BURKEN: Right. When we started reviewing this sometime, sometime before the panel itself, we found that the Fruehauf and Bosanquet review article from the 1993 PPO updates, crystallized many of the issues. And what we did, based on it, there were some summary tables that were actually presented this morning, where they looked at groups of studies. And I again refer you to summary table seven and eight from the 1993 PPO. And as a result, we really focused our efforts, our literature efforts on

00271

the EDR, as well as also some of the other thymidine assays, because there were some other thymidine uridine incorporation assays pertinent to bring to the panel. And then we also did a lot of sampling of DiSC and MTT, using a MEDLINE search, and we did not have any time limit on our studies.

And when we went through and did our literature search, then we had to figure out what we would want to present to the panel. And since clinical response was one of the outcomes we

looked at, as well as survival, we needed to have confidence in the viability of our two-by-two tables. So as a result, any study that lacked the clinical criteria -- either the -- the clinical criteria either had to be documented or referenced, you know, for clinical response. We only looked at adult patients.

And just for the record here, in the rather extensive handout which has been provided for this session, we do list the pediatric studies that have not been summarized in this presentation, but there is a notebook of all the studies that are being presented in this presentation, are available. I know it's kind of

00272

hard to read the whole notebook tonight, but as a supplement to the materials you already received, there are papers in here such that anybody that has any questions about any of the bullet items from this afternoon's presentation can go back to this, as well as your other materials.

We included both prospective and retrospective two-by-two data designs. The only thing we did exclude for this, again, for this panel presentation, were descriptive type studies that didn't use any quantitative summaries. There were some studies that went beyond two-by-two tables, used regression analysis and some other techniques, and those were included.

Now talking about all these studies, you know, how can we present these studies to the impact panel? Can we group them or pool them, or do we need to go through them individually? It was something that we really had to spend some time thinking about. And we came to the very, very strong conclusion that data from the individual studies should not be pooled. The reason being is that they're, the studies are so heterogeneous, they use different cutoff points, different tumor drug combinations, different

00273

clinical response criteria, that we just felt very uncomfortable about doing a meta-analysis

for the purposes of presenting data to the panel, okay? So therefore, each study must be presented on its own merit, and I think that's a fundamental approach to presenting data this afternoon, and probably lengthens the presentation a little bit, but we feel it's important.

So now, let's just go through the evidence. Let me just walk through the handout with you. I don't -- there is a lot of bulk on my slides, but again, it's in the handout and it is really set up to be a reference guide to trying to put it all together.

The assay formats I start with are not based on cell death versus cell proliferation. It's not done that way. I went from the assay formats that we concentrated on, as in the EDR, the DiSC and the MTT, where we really had the most literature, and then towards the end I have some of the other formats where there was a little less literature that came up, based on the criteria that were described in previous slides.

The Kern and Weisenthal article from

00274

1990 is a complex article that was referred over to Dr. Burke and Dr. Handelsman for separate review. Again, a central piece of evidence, but highly complex.

But let me go through some of the other articles that pertain to EDR as well as some of the other thymidine uridine incorporation formats.

Eltabakkh, '98, shows, you know, some PPVs and NPVs. There were some confidence intervals that are reported. As you can see, the NPVs in this study is actually fairly low. Let me start, as I said, rather than to go through all the bullets, let me try to highlight what are some of the themes. As I said, there were a hundred new patients with ovarian cancers. We find out in this study, all the patients were recruited prior to chemotherapy, which is important when we think about selection bias.

And we found 75 evaluable patients, so we went from about a hundred down to 75, which is really pretty good.

Fernandez-Trigo is a study, again, that also has some case loss of about roughly 25 percent. But in this case, there was a very rare

00275

site cancer that was selected, so I would just keep that in mind.

Moving on to some of the other thymidine uridine formats, you know, I talked about the CRA, the Bartels CRA, and it turns out that there was a study by Elledge back in '95 that enumerates the findings of clinical trial. And one little, kind of I suppose warning to the panel when evaluating this paper, I mentioned that NPV is a marker for chemoresistance and PPV being a marker for chemosensitivity. Well, in this particular article, not the article but the package insert, and the summary safety and evaluation data, it's flip-flopped, so you have to be a little bit careful there. It turns out you have to reverse that, so you have to really be wide awake when you read these two-by-two tables. And I mention that down here, that the two-by-two table design differs from the other studies presented, even differs from the Elledge paper, which is -- the Elledge paper comes out before the FDA submission data. And this was a prospective blind enrollment of 60 relapsed breast cancer patients. The interesting thing about this particular assay format is that it was

00276

very specific for breast cancer and 5-FU.

Well, what about some of these earlier five-day thymidine uptake assays? Sondak in '84 had a series of 142 patients with successful assays out of a pool of 219, with 33 clinical correlations. Quite a bit of case loss here, even though, again, you know -- well, the numbers are small, but the NPV, again, is high. You know. But one, again, has to be concerned about possible selection bias.

Sondak in '85, 819 mixed solid tumors, again, if you use different cutoffs, you're going to have different PPVs and NPVs.

Sanfilippo, in '81, there were several studies on three-hour incubation, rather than the five days, you know, and there were -- as I said, you can see the numbers here. I think the interesting thing about this particular study in '81 was the use of subsets for high proliferative and low proliferative non-Hodgkins lymphoma cases.

And in '86, Sanfilippo, the same group, went ahead and studied 169 patients with various types of germ cell testicular tumors, but only 29 cases were available for clinical correlation.

00277

Again, we didn't really know how many people were previously treated or untreated, and that could inject some bias.

More three-hour uptake assays, two studies by Silvestrini in 1985 and Daidone in 1985 from the same institution as Silvestrini. Different tumor types.

Well, let's kind of move along, and we finished up with the thymidine/uridine incorporation assays, and let's move on to the DiSC assay. And as I say, there were several papers that were reviewed in the Cortazar and Johnson article, which is the review article in 1999, which did a MEDLINE search, and targeted 12 studies, and four of those 12 studies were DiSC assay approaches in solid tumors, three of the four studies being small cell lung carcinoma, the other study being a non-small cell lung carcinoma. And I think, you know, in each of the studies, the test groups did at least as well as the control groups. The survival data was not particularly convincing. In the Gazdar study, the survival rates were similar; in the Wilbur study, the survival rate comparisons were not am. Again, that is survival rate of assay

00278

directed versus, you know, empiric therapy

groups. In both the Shaw and the Cortazar articles, their survival rates were really not -- there was really not enough of a difference to really hold much discussion.

But I think where we find a lot more evidence, again, based on our structured review, is looking at this, and hematologic tumors. And we start with Dr. Weisenthal's study in '86 where there is 70 cases. What we did was we subtracted out the 29 cases of ALL. Again, it's just a judgment case one makes as to how you want to treat the pediatric tumors. I can tell you that the pediatric performance table was just about the same as for adults, and there weren't significant differences, so I think what one could --

DR. FERGUSON: Mitch, closer to the microphone please, or maybe you should hold it.

DR. BURKEN: Yeah. As I said, one can scan through some of the pediatric studies quickly, and I think get a flavor for that. But moving on, looking at the adult data, we had PPVs and NPVs that were over 80 percent.

Dr. Bosanquet in 1999 had a fairly

00279

elegant study, and it was reviewed earlier today. I will leave the details to the group.

Another study that was also discussed earlier today was a study by Mason that did some modeling. In this particular case, not only was there some clinical response data and survival data that was looked at, but there was some modeling done using regressions, where if you -- I know that the print is a little small at the bottom, but I just wanted to mention that if you look at the life years gained per assay, the modeling here said that if you had a simulated 50-year old with stage C chronic lymphocytic leukemia, there would be a life years gain would be about six months, and about three weeks if it was a simulated age 70 stage AB female. Again, these are all simulations that are based on the assumptions in the regression modeling. But it's

a little bit more of a sophisticated approach, as I said.

And continuing on, there's been, as I said, a fair amount of work in hematology. Tidefelt in 1989, with more than 90 percent of the patients not being previously treated. There is a complex predictive value studies in this

00280

paper with varying anthracycline concentrations and different treatment regimens. But if you flush out all 40, actually 40 out of 53 patients were available for clinical correlations, and you can see the PPVs and the NPVs.

To go on now, to continue more on the hematologic DiSC studies, Bird in '86 was a small study. That's the one I quoted earlier because there was peripheral blood and bone marrow that were used as two separate sites. And there seemed to be reasonably good concordance between in vitro testing, between peripheral blood and bone marrow. But again, I caution, the sample sizes are fairly small here. Bird in '88, again, another small sample study.

I'm going back just some more. More small sample studies. Dr. Bosanquet in 1983. Dr. Beksac in '88. I think the interesting thing about this study was it's kind of a mixed retrospective prospective approach, so it was a somewhat complicated study even though there were only 16 patients.

Dr. Bosanquet's 1991 study that was actually just up on the screen a few minutes ago, showed 67 patients with CLL where there was a

00281

survival benefit. But just before we leave DiSC and the hematologic applications, you know, there were some articles that didn't have documented clinical criteria, and you see down, Dr. Bosanquet's article. Again, we certainly looked at the survival data, but we didn't feel that the clinical criteria were adequately specified in this particular paper even though, as I said, it showed some survival data.

And Kirkpatrick from 1990 was a paper, again, all in the backup book here that has some pediatric data.

Moving on to MTT, several articles did not have documented clinical criteria so that for the purposes of this panel presentation, we didn't feel it would be useful to present clinical response data. And it's interesting that three -- we talked about the 12 studies from Cortazar and Johnson. Three of them were Yamaue, 1991, 1992 and 1996, but unfortunately, none of those articles had, you know, adequately described criteria, and we just didn't feel we could construct good enough two-by-two tables. And the survival rates in these studies do not compare test versus control groups. And so, the

00282

pediatric neoplasm studies are listed there. Veerman, I think was also mentioned this morning. But again, you know, it's our choice.

Well, let's move along then to some of the solid tumor studies for MTT, and this goes more or less in chronological order. We had Suto in 1989, with GI solid tumors. Again, a very small number of clinical correlations are available.

We had Tsai in 1990. This was from cell lines from 25 patients with small cell lung cancer. In this case we have regression modeling as opposed to two-by-two data.

Furukawa has a larger sample, but again, only 22 patients available for clinical correlation. So the numbers may be high, you know, an NPV of 100 percent and a PPV of 75 percent but again, with small sample sizes and this degree of case loss, one really has to wonder about the possible selection bias. And then we see some survival benefits in this study.

Saikawa in 1994, 50 patients, 40 of whom received post-surgical chemotherapy. This was basically just divided up into two groups, an

00283

adaptive group versus a non-adapted group. Again, we have some survival data here as well.

Sargent, '94, 206 confirmed or suspected epithelial ovarian adenocarcinoma patients. 37 were previously untreated. And again, we have a -- we were able to have survival data on 37 of those 206.

We have a more recent study by Taylor, again, stage three, four, previously untreated adenocarcinoma. 43 available for clinical correlation, or roughly 50 percent out of the starting 90 were finally available after you consider tumor evaluability and clinical correlation. And we have a couple of subgroups here for all treatments and platinum only.

Xu, 1999, it's in your packet, your original green book. 156 advanced breast cancer patients. And they actually noted in the study itself that the source of selection bias -- well, they didn't say they had selection bias, but they did say that they preferentially recruited worse prognosis patients in the MTT directed versus the control group, which was certainly a source of concern for those reviewing it.

And just a -- hematologic MTT tumors,

00284

there's just a lot of those. If you go several slides back, a lot of those studies like Veerman and Hongo, and Hongo were excluded because they were pediatric studies, but if you look at 23 patients with de novo AML and five in CML blast crisis, 21 were available for clinical correlations, with again, good looking predicted values, but I think people should evaluate the robustness of the numbers.

Then, just to kind of close out, again, we wanted to be fair and not just presenting the thymidine incorporation assays as well as DiSC and MTT, so we did look at some studies from FCA and some of the other assay formats. Leone had 78 cases in 1991. This is again, for those of us that are trying to keep up with the different abbreviations, this is the fluorescent cytoprin

assay, and see, the Leone study.

And then Meitner in '91 actually extended the Leone data set and worked it up to a total of 101 cases with similar NPVs and PPVs.

FMCA is a similar fluorescent method, a little bit more recent. Some of the literature, a Larsson study had 43 samples with 27 clinical correlations. Again, the numbers look pretty

00285

good. In this circumstance we did find blinding. I'll tell you a little bit about blinding towards the end, but not terribly often did we see evidence of blinding in the studies.

Csoka is a more recent article, as I mentioned. 125 patients with newly diagnosed or relapsed ovarian cancers. 45 available for clinical correlation. He did have a breakdown of previously treated versus drug naive patients. Blinding again was reported. And there was, in a small group again, an NPV of 100 percent.

Again, moving along, Dr. Nagourney was kind enough to submit a manuscript to HCFA on his apoptotic assay. One thing I would mention is, or question I would raise is, the manuscript was just a summary manuscript, and from it we were not able to determine how the EVA assay improved treatment management beyond empiric treatment regimens for the refractory patients. So that is a question.

Then we also looked at several of the HDRA papers that were submitted to us by Dr. Hoffman and his company. Many of the articles that were submitted to us are in a -- again, all of it is available to the panel in a notebook

00286

form, but what we did is we pulled out, and they're also in here, the four, three articles in a manuscript that are clinical correlations. Many of the other patients were experimental and pharmacologic studies.

I might add that there was a lot more material that was submitted to HCFA than my presentation would suggest, many many more papers

that we looked at. But many of them were experimental pharmacology studies, and we didn't feel that this particular venue looking at medical necessity would be quite the right place to get into a lot of extensive experimental pharmacology.

So looking at these four papers from HDRA, again, small sample sizes, but again, you can see the NPVs. Again, very few of the studies had confidence intervals calculated, as you can see throughout my presentation.

Furukawa in '95, this was presented earlier in the day. Post-surgical stage three to four patients. Mixture of gastric and colorectal tumors, and similar findings that we've seen.

Kubota, 1995, stage three four gastric cancer with somewhat -- when I bulleted these for

00287

you, what I have done is I've not gone into all the different subgroupings, so therefore, some of the sensitive groups range from a sample of 20 to 38, and the resistant groups from 89 to 99, but I have not gone into great detail to specify all those subgroupings. I am trying to give the main message here.

And then Dr. Hoffman's article that's a manuscript in press. Again, more gastric and colorectal tumors.

Well, as we try to summarize all the literature and take this kind of view from the hillside here, the Cortazar and Johnson article helps do that a little bit, in the sense that they've selected out 12 prospective trials. I mentioned that four of them were the DiSC studies that I outlined, the three MTT studies that I didn't present to the panel because of the lack of documented clinical criteria, and five of those 12 studies were the earlier clonogenic assays.

Overall findings from the 12 study review, showing that only a small percentage of patients have actually been treated with an in vitro selected regimen, and that's certainly

00288

consistent with many of the studies that I have presented along the way, demonstrating case loss. And most of the patients have had advanced stage solid tumors. The overall assessability rate only being 72 percent, but I think it's really quite fair to mention that five of those 12 studies were from the earlier clonogenic methods, so that would have a negative impact on the overall evaluability rate since again, we're talking about five assay formats that were not as technically advanced as DiSC and MTT.

And these trials, the response rates among the directed therapy patients were at least as good as those achieved with empiric therapy, and five of the 12 trials illustrated survival data for the directed versus empiric therapy, but it was difficult to determine overall trends in these five studies, including three DiSC trials. In only one of those trials was there randomization, and in that particular trial all the experimental arms consisted of small sample sizes.

So where are we, at nearly the end of the day here? I think we found in going through this systematic review that there is not strong

00289

convincing medical evidence to support the overall clinical utility of human tumor assay systems. The comprehensive literature review demonstrates that there were many different tumor drug combinations among different studies and this made it difficult to really make conclusions about particular tumor drug combinations because of this variability. And that's really kind of what I would call a structural feature of reviewing so many articles. Many of them had small sample sizes. We had frequent selection bias, recruiting documented or possible refractory patients.

Remember, let's go back to our utility function where we are thinking about being in the center of that function or at the extremes, and

if you are recruiting patients into the study who are at the extremes of that utility function, then there is a concern that regardless of whether the negative predicted values are high or not, you're not getting a lot of clinical utility. And in the same vein, by recruiting advance stage patients, you may be getting yourself into a situation where without lab testing, you pretty much know that a patient

00290

isn't going to respond anyway, therefore your negative predictive value or your positive predictive values are going to be adversely affected by such selection bias.

And I've noted before that there was only rare or occasional documented use of blinding.

Well, that's the broad sweep, but when we go down and we look at it in a little more detail, we really should note that there were a relatively higher number of clinical correlations available for DiSC and MTT assay formats. As a result, the human tumor assay systems may have a greater potential clinical utility for hematologic neoplasms such as CLL, where there has been really a fair amount of work, then solid tumor. And when considering this literature, let's never forget, you know, the importance of evaluability and heterogeneity in making determinations.

And again, we are still in the tough spot of trying to apply single agent drug tumor interactions to multiple agent regimens, and I think a certain amount of inferences have to be made from these, you know, in vitro studies.

00291

Thank you.

DR. FERGUSON: Thank you. Dr. Bosanquet?

DR. BOSANQUET: Thank you, Dr. Burken, for summarizing that. I'm glad to see the up to date work has been included in the production.

You spent some time on your clinical

utility curve at the beginning. I wonder if you could explain to us all how that was mathematically derived, because I would have drawn a different curve.

DR. BURKEN: Well, one could, I suppose one could argue that rather than being a triangular distribution, it could be a normal distribution and have a slightly different look to it. But I think what we ought to do is agree on the fact that a lab test is most valuable when you're most unsure of whether the patient has a disease or not.

DR. BOSANQUET: I quite agree with that.

DR. BURKEN: I think that's a critical point, and I think that ought to be established, and that the value of any lab test is going to drop off considerably if you're at the extremes

00292

of prevalence.

DR. BOSANQUET: Well, that's the bit that I would necessarily disagree with. You have drawn a triangular curve here, if I can call it a curve. We all agree, I think, with the 0 percent on the left and the 0 percent on the right, or the low added information at both left and right, and the very high added information in the middle. But just the shape of that curve, and you have spent some time on it, and I just would ask you again, how did you mathematically define that? Because if you look at the Bayesian curves, I think you would find a mathematical, if you define that mathematically from the Bayesian curves, I think you'd get rather a different curve, and your conclusions form this bit of the talk would then be different.

DR. BURKEN: Yeah. Let me just say that I, you know, I'll admit up front that this could have been a normal curve rather than a triangular function. But the most -- rather than getting bogged down in the mathematics, I think it's important for the panelists to consider the question of mapping out -- let me go to this next

graph. What we need to do is we need to kind of
00293
map out an area where we feel laboratory testing is reasonable and necessary. Now we're not -- I'm not standing up here and telling you that that cutoff point -- it happens that I have the yellow box here at maybe 20 percent or 80 percent. I'm not coming out and telling you that there's any mathematical validity to making the box 20 percent and 80 percent. What I'm trying to do is illustrate a concept of how a lab test becomes less useful as it drops off away from the 50-50 point.

DR. BAGLEY: Would it be fair to say that although the Bayesian curve which you showed, which is mathematically derived, deals with the probability of a correct diagnosis, whereas what we're dealing with here is not the probability but the clinical utility of increasing that probability? I mean, as the certainty of the disease goes higher, the probability, you know, based on a combination of tests, is also going to go up. But as that probability becomes more certain, the clinical utility or the incremental value of that additional information becomes less. And I think this is an expression of the value of the
00294
information.

DR. BOSANQUET: I quite agree with you, but you see, many of these tests are used on resistant patients, where the pretest probability of response is very low. And what Dr. Burken is implying by this curve is that if you have less than 20 percent pretest probability of response, then these tests aren't very useful. And I would challenge him, and I think he admitted that this is not mathematically defined.

If we could just have a look at the one slide that I've got? We've seen this slide before, and the important thing is, if you take a pretest probability of response of, say 5 percent, Dr. Burken was suggesting that anything

below 20, the test was not going to add any information. But if you look at this, the information is added at very low levels, because if you take a patient who has a pretest probability of response of 5 percent, you can split that into test sensitive patients who have a probability of response of 20 percent, and those who have a probability of response of 1 percent. And I think that's, I would disagree, and I think that would be a useful addition of

00295

information to those patients with very low pretest probability of response, so anything from 5 percent on. And therefore, I would suggest that the curves you were showing were somewhat misleading. That's all I'd like to say.

DR. BURKEN: Yeah. What I'm going to do is I'm going to let Dr. Burke pitch in a little bit with some of the mathematics. But again, I want to emphasize that the schematic that was put up did not, you know, was not -- that box did not mean to imply that 20 percent or 80 percent or 30 percent would be some type of cutoff that this panel would be expected to respond to. The diagram is simply there to conceptually show in actually probably more even a qualitative way than a quantitative way, that there is simply less value from a lab test among, in a situation where you are very sure of a disease, or you think the probability of disease is so low, that's also another scenario where the test wouldn't be terribly useful, and that the inference from that diagram -- so let me flip it later on to this one.

And again, please don't read any cutoffs in here that, where the red triangles

00296

begin at 20 percent or 80 percent, please don't read it that way. But the purpose of that diagram is simply to show that at the extreme regions of low probability or high probability, if you have studies with selection bias, where patients who were recruited into a study are in

the extreme regions of that utility function, what it can do is detract from the power, or I don't want to use that word power because that's a statistical term and I'll get myself into trouble with some of the statisticians. It can detract from the ability to use positive and predictive negative values as a marker of clinical utility. You just have to be aware of what kinds of patients you have in your study before you can go to bat with high NPVs or PPVs to make a case for a lab test, any lab test.

DR. FERGUSON: Dr. Burke.

DR. BURKE: Thank you. There's a couple points that have to be made. One is, when you -- I mentioned briefly the 50-50 situation, which is the fair test for a test. But what happens is if you look at the accuracy of a test, it's very difficult to find therapy dependent prognostic factors that are really accurate.

00297

It's really hard to do. If you're looking at estrogen receptor status, the area under the ROC is about .62. Okay? So it turns out that these factors are fairly weak. This test is a therapy dependent prognostic factor. And the issue becomes, if your test has an accuracy of .62, but being a naive Bayesian gives you an accuracy of 90 percent correct, okay, then the issue is, what are you going to be? So the bottom line is that you have to look at the marginal utility of your test in relation to the population you're using it in.

And I think what Mitch's slide is pointing out is not a particular shape, but the fact that it becomes harder and harder to exhibit any marginal utility as the prevalence of the disease goes up. Because eventually, you're going to become a naive Bayesian, because that is the correct approach when the prevalence becomes very high. And in fact, it's true, that is the correct thing you should become, because your test is not as good as predicting the prevalence.

Now, the other thing is, you could say well, maybe my test can help a little bit at the limit. But the problem is, your test has a variance associated with it. And then the question becomes, as the prevalence goes up, it becomes almost acentotic, and so you have very very little room to move, and your test variance can take up that room, so you'll never know whether you're doing anything good or not.

DR. FERGUSON: Dr. Fruehauf?

DR. FRUEHAUF: I would like to thank Dr. Burken and Dr. Burke for telling me that I should be a naive Bayesian, although I don't know what that is yet. I think statistically, I'd like to use this curve, because I agree with this. I think this is true. I think your points are valid and I'd like to use this as an example of how there is a relationship between what we're talking about and what you're talking about, because I don't think they are separated.

Now, let's use this curve. Here we have the probability of response in the middle of 50 percent, because we're using these tests to predict response, not whether disease is there, so we're making an assumption that disease is equivalent to resistance or sensitivity; true?

DR. BURKEN: Well, I'm a little concerned. You know, we may not want to overplay this one issue.

DR. FRUEHAUF: This is your model, and you're relating it to in vitro drug response, so please tell me, how does this curve relate to in vitro drug response? What is the relationship between the X and Y axis, and treating patients with chemotherapy?

DR. BURKE: Well, let me tell you the way I would choose to use it, okay? And just to make sure we're on the same wavelength here. The way this graph is designed is that that block, the granite block in the middle that's gray, wherever we should have those cutoffs, and we

won't argue about that, demonstrates that there is a lot of information added by lab testing in that region, because there is enough uncertainty about whether a patient is either resistant or sensitive to that particular drug, and again, they're reciprocals of each other, so I can use them interchangeably.

DR. FRUEHAUF: Okay. Can I go from there? I understand that.

DR. BURKE: But let me say, let's not talk about that it measures response. It's the

00300

height of the graph, the Y axis, is how much value there is from the lab test result.

(Inaudible question from audience.)

DR. BURKE: Basically it is simply saying that the greatest uncertainty is at 50 percent prevalence, right, of response, non-response, whatever the case might be that is your gold standard.

DR. FRUEHAUF: Sure.

DR. BURKE: And the issue simply is that if you set your study for a 50 percent response or non-response, and then you test your drug or whatever in that population, then the prevalence isn't going to help you or hinder you in your predictions.

DR. FRUEHAUF: Yes, I appreciate that. So let's take Tamoxifen and breast cancer, previously untreated breast cancer. If we give Tamoxifen to women without knowing their receptor status, there's a 30 percent response rate across the board. Okay? No test. Everybody gets Tamoxifen, 30 percent response rate. Well, that's okay, but can we do better? Let's get receptors, and now treat according to the results, and see if we change and enrich for

00301

response, and eliminate people from what can be toxic therapy because of side effects. And what we find is that if you have estrogen receptor in the tumor, there is a 75 percent response rate. And if you don't have those receptors, there is a

10 percent response rate. So my question to you is, can you relate that knowledge and that test to this curve for me?

DR. FERGUSON: I am going to take the prerogative of cutting this off right now so the panel can have a discussion. I'm getting hints.

DR. WEISENTHAL: Before this, you said that I would have the chance to respond.

DR. FERGUSON: I did actually. It's going to have to be very brief.

DR. WEISENTHAL: The point I was trying to make, Dr. Burke was implying bad science or sloppy science, or whatever. And also, HCFA in their review specifically excluded the pediatric ALL patients. I just want to make the following very brief points.

Firstly, medicine is imperfect. Secondly, medical oncology is inadequate. 70 percent of all the treatments that we give don't work. More than half of the chemotherapies for

00302

non-FDA approved indications, none of which would stand up to the level of rigor that Dr. Burke is asking here.

So I think it's very important that you have to look at the information as a whole. Earlier in my presentation I presented what I call the central hypothesis, and the central hypothesis simply stated is this: You test tumors in vitro, you get a spectrum of responses in vitro, and that the responses in vitro are related in some way to the responses in vivo.

I showed you 35 studies which were admittedly, and you know, you got some detail there, of variable quality, and some were very marginal quality, but some, particularly the ones that were excluded, and I don't think there is a meaningful biological reason for excluding pediatric patients, other than they don't get Medicare, but the disease is enough similar.

But if you look at the work that was done at the free University of Amsterdam, which was excluded, and I want to tell you about that,

that would stand up, I believe, to Dr. Burke's level of rigor. What they did there was they first of all did their training set studies where

they did their retrospective analysis. They got their criteria for sensitivity resistance and so forth. And then in a prospective blinded fashion, using those criteria which had been established from the retrospective study, they prospectively tested it in the cooperative group study in a double blinded fashion, published peer reviewed in the journal Blood, which is one of the most rigorously peer reviewed journals around, and it showed absolutely astonishing great results. And you have to include that paper in the context of everything that you've heard.

Now, the point that I wanted to conclude with is that if all we're talking about is validating an estrogen receptor, it would be very simple. We're talking about one test, a very common disease, breast cancer. But what we tried to do, beginning 20 years ago, is we said we have this morass, we've got hundreds of diseases, hundreds of potential therapies, which are increasing every year dramatically, and there just has to be some way of matching patient to treatment.

DR. FERGUSON: Okay. Thank you.

OPEN COMMITTEE DISCUSSION

DR. FERGUSON: Are there some questions?

DR. HELZLSOUER: I would just like to throw out one comment, with the estrogen receptor analogy is that the reason we know all this is because they were done in clinical trials. They were evaluated in clinical trials, and I didn't want to lose that momentum for tomorrow's discussion.

DR. FERGUSON: Other questions from the panel? Let me just ask a point of order here.

(Discussion off the record.)

DR. FERGUSON: Mr. Barnes?
Mr. BARNES: Actually, I have a question for Dr. Burke. Would it make any sense to go back over data, or would it in fact be too hard or impossible, to do a disease specific analysis based on test by test? I mean, it seems to me that we're all bumping up against the fact that there is a bunch of different tests, and about 30 different types of cancers.
DR. BURKE: No, I'm -- I, for all my strong comments, I'm agnostic as to the test itself. I have no opinion on it one way or

00305

another. But that is the only way to evaluate the claim that they really want to make.
MR. BARNES: Right. But what I mean is, using the data that are either in the articles, or could the data be generated some other way, to reevaluate.
DR. BURKE: Well, let me make two brief comments. One is, it's striking that there isn't a large cohort study for a particular disease, which is what one would expect, given the frequency with which this test seems to be done. But number two, yes. For some diseases, for example CLL, it may in fact be the case that there is sufficient evidence, okay, to evaluate a particular test for that particular disease. And the reason why I say particular disease is because diseases are kind of strange, as you well know and I well know, and CLL is a very strange disease, but it has its own characteristics associated with it. And so yes, you'd want to look at CLL in terms of CLL. Why CLL? You're saying well, for this test, given the characteristics of CLL as a disease, does this test help us? In early stage disease, in late stage disease, for particular treatments, if

00306

there are effective treatments, because remember, for therapy significant prognostic factors, if there is no effective treatment, then there is no need for therapy specific prognostic factors.

MR. BARNES: Right. Well, let me ask my question a different way. Of the 12 studies, or 13 or 14 or whatever they are, is there a way to go back to them and dissect out the histologies, CLL or whatever, according to test result, specific test by test, and get data? So in other words, do you think that anyone, not necessarily you, could go back to the actual publications and dissect that out?

DR. BURKE: It depends on the publication, it depends on the study. Some yes, some no. It depends on the adequacy of the study. Some studies you, like retrospective studies, it would be very difficult to do that. Prospective studies that were done properly would be much easier to do, because you would have complete information, which you most of the time don't have in retrospective. But yes, it could be done, if the data were there.

MR. BARNES: I'd just like to add a couple of comments as well. Many of the studies

00307

on solid tumors and even hematologic tumors or mixture of tumors, and the studies do not specify histologic subtypes, and it becomes very difficult to create a laundry list of studies by disease type. I think if you go through the handout from this presentation, you will see that unfold, because I do specify the, you know, whether it's mixed or what tumor types it is, you know, and sometimes it's just very difficult.

DR. FERGUSON: Other questions from the panel? I have a question I'll throw out. It seems that in a number of the papers that we've seen, when there were comparison groups that they were, if there were patients whose cancer was sensitive to a drug, they were given that drug, whereas the other quote, control group, was one who showed resistance, and those were allowed to have physician's choice in the chemotherapy. Now that seems to me to bias the two groups if the test has any validity at all, so that they really aren't comparable. That is, the ones with the

drug, the cancer showed sensitivity, were treated by guidance from that test, whereas the ones that didn't or were resistant, were treated by physician's choice. And then one -- that group

00308

comes out worse, and why wouldn't we expect that? And I guess I'm asking for a comment or an explanation for why that's the best thing to do, because it does not seem to me to be the best thing to do. Yes?

DR. HOFFMAN: My name's Robert Hoffman. We performed such a prospective study. I think in the previous retrospective studies we showed very extensive correlation between survival and response in the drug response assay, in our case the histoculture drug response assay. So we then designed a trial as you mentioned, comparing outcome of patients who were treated by assay guided therapy if their tumors were responsive in the assay, to clinician's choice in the resistant patients.

I think it would have been unethical to treat the resistant patients with the resistant drugs as a matter of course. So that was, I think, the criteria in our study. Of course the next step, I think, would be an absolute randomized trial where you separate the patients beforehand, but I think if knowing someone is resistant, given not only the data from our studies, but the very very extensive data

00309

presented by the other groups here, I think it's not, and respectfully in my opinion, it's not ethical to treat with a resistant drug.

DR. FERGUSON: Are there any other questions or comments? Go ahead, Dr. Klee.

DR. KLEE: The study that Dr. Bosanquet alluded to, at one point in the presentation they were talking about this MRC study, I guess is ongoing, the randomization for, which really is randomizing against use of the drug testing versus not using the drug testing. Does that -- that seems like a rather fundamental type study,

and I was surprised that hadn't been done earlier, it's ongoing now, but it would be sort of the basis of much of the clinical trial work that's been done on a lot of the therapeutic side of things, so it just surprises me that there was no published study along that line. And apparently there are numerous difficulties in trying to carry that out, but I don't know why that hasn't been done or what has precluded doing that.

DR. FERGUSON: Yes, Dr. Weisenthal?

DR. WEISENTHAL: As one who participated in the design and funding of such

00310

studies, what I have to tell you, it's one of these things that's easier said than done. In 1985 I had a large grant from the VA, had 31 VA hospitals, it was a cooperative VA study in multiple myeloma, standard therapy versus assay directed therapy. It was several years in planning, we had two national investigators meetings, one was held here in Baltimore. A tremendous amount of work and everything went into that. What happened was that eight months into the study, accrual was running only about one-fourth of what had been projected, they decided that the study just would not be ever completed and so it was cancelled.

Subsequently, we got a study going in the Eastern Cooperative Oncology Group, which was to lead to a randomized trial in non-small cell lung cancer. Again, in the first six months the study accrued six patients, although we had 51 hospitals eligible to contribute patients, and that was closed.

And I keep mentioning Dan von Hoff, who's the most energetic effective clinical trials organizer I've ever seen, tried several times, and never completed a single prospective

00311

randomized trial. It's just much easier than said, for all sorts of reasons, that we could discuss over a margarita.

DR. FERGUSON: Thank you. Yes, Dr. Burke?

DR. BURKE: The cooperative groups and other randomized trialists are collecting frozen tissue, and an issue that may be available to you. I mean, they know the different treatments, they know the outcomes, they have snap frozen tissue. The question is, can these assays be done on snap frozen tissue, because if they could, the outcome is already known.

DR. FRUEHAUF: It would be really wonderful if we could use frozen tissue for the assays, and this is really one of the technical issues of doing a prospective randomized study, and we did this with the GOG. GOG-118 was a prospective study, wasn't randomized, but to be obtain fresh tissue at surgery, send it to the laboratory, and I am a member of SWOG, and I attend GOG meetings, I'm a member of ASCO, and I can tell you that tissue banks are a great idea, but they haven't really reached fruition because of the logistical problems of moving tissue from

00312

one place to another is very difficult. And so we have not -- you can't use snap frozen tissue; it has to preserved in a live state in media, and transported so it gets there within 24 hours.

DR. FERGUSON: Thank you. Other questions from the panel? Yes, Dr. Helzlsouer?

DR. HELZLSOUER: I have a question in terms of these assays, and I'm having a little trouble lumping them all together. But is it my understanding that there is only maybe one that tests combinations routinely, all the rest are single chemotherapy assays?

DR. FRUEHAUF: The question of single agents and combinations is kind of a tempest in a teapot in a way. Every lab that I know of tests drug combinations. We test drug combinations at Oncotech, AntiCancer tests drug combinations, Dr. Nagourney tests drug combinations, Dr. Weisenthal tests drug combinations. I think one of the issues that is fundamental is, is there drug

synergy, and if you don't test two drugs two together where there could be synergy, what are you going to miss in the information? And this goes to the issue of why we use multi-agent therapy. You are an oncologist and an

00313

epidemiologist, and I'm sure your thinking like many oncologists is, we use multi-agent chemotherapy because of the gold ecomin hypothesis, that there are multiple subsets within each tumor that are differentially sensitive to different agents in the combination. So when platinum was added to testicular chemotherapy regimens, that additional activity killed a subset that was there microscopically. Even though people had CRs, they weren't surviving.

So single agents should be active in combinations. So our view is, if you test a single agent and it can't reach its drug target, there's extreme resistance to that single agent, and it can't reach its target because of protein or rapid adduct repair or what have you, that single agent isn't going to add a synergistic effect in the absence of its own effect. So, Dr. DeVita, in the third edition of Principles and Practices, made a statement in his chapter on chemotherapy that combinations should always be made up of active single agents. And so we look for single agent activity against the cancer in the salvage setting, and then we'll move this

00314

agent up into the adjuvant setting, when it's been proven to have activity.

So single agent testing is predicated on finding out if the single agent would have no benefit as a single agent, it's unlikely then that it would have benefit in a combination, but we all test combinations.

DR. FERGUSON: But let me -- as I recall reading these papers, that none of them actually routinely were testing two agents simultaneously on one. I mean, the majority of

the papers that we read and that Dr. Burken presented single agents. Maybe serially they would test several agents, but not together in one Petri dish routinely.

DR. FRUEHAUF: Yes. I think that Dr. Nagourney presented evidence, and I will let them speak, but just for our role, we tested the concept of whether single agent testing was predictive in combination therapy in breast cancer. So we took the single agents and looked at their activity as single agents, and added up their scores as I presented this morning, and that was predictive of how the person did in response to the combination. Now other people

00315

have tested combinations as well, and I'm sure they will comment on that.

DR. FERGUSON: Okay. Was I misreading these papers?

DR. HELZLSOUER: That's the same way, I interpreted them the same way, they were all single agent tests, and not combinations.

DR. FERGUSON: Yeah. I mean, all the published stuff we saw was single agent.

DR. HANDELSMAN: The bulk of it was, but not all of it.

DR. FERGUSON: Okay. Yes?

DR. HOFFMAN: Technically, to test combinations is entirely feasible. We're dealing with most of the tests with culture dishes, culture wells, with medium. You can add one drug, two drugs --

DR. FERGUSON: I don't disagree with that.

DR. HOFFMAN: You can add ten drugs. Most of the studies have, as has been mentioned, have focused on single drugs to understand their individual activity. We've done a study as yet unpublished that shows predictivity to the combination treatment for ovarian cancer as

00316

predicted by Cisplatin alone, but to mix drugs in the cultures is technically trivial.

DR. NAGOURNEY: Yeah, if I might just address that. Actually we specifically do focus on drug combinations, and as Dr. Fruehauf alluded to, most drug combinations are basically additive, and in some cases subadditive or antagonistic. There are a small number of combinations that are genuinely truly synergistic, and which are extremely attractive and interesting as therapists. One of the most attractive are the interactions between alkylating agents or platinum, and antimetabolites, a couple of examples of which were cited in some things we referenced, one paper in the British Journal of Cancer, indicating true synergy between alkylating agents and CDA, and that observation has now resulted in a 100 percent response rate in an ECOG trial.

Similarly, Cisplatin and Gemcitabine as a related combination in solid tumors is presenting us with really one of the most active combinations we've ever seen in medical oncology, but those are actually pretty rare. So I think for the most part, most drugs are intelligently

00317

given as single agents, but there are a few very beautiful examples of synergy, and they can be test.

DR. FERGUSON: Is this in response to that?

DR. KERN: Yes. Just briefly. In the now famous, or perhaps infamous Kern and Weisenthal paper of 1990, we had a cohort of 105 patients that were treated with combinations, and we showed that --

DR. FERGUSON: I don't doubt that the patients are treated with combinations. The issue was, was the test done with two drugs?

DR. KERN: That's correct. All the drugs were tested singly and in combination in the laboratory, and correlated with the clinical with the clinical response.

DR. FERGUSON: That wasn't clear, at least to me.

DR. KERN: I understand. It's in Table 5 of that paper. Thank you.

DR. HELZLSOUER: Another concern I have which hasn't been addressed, and we didn't really have it in our packet, were the reproducibility issues of these tests. Then I just heard that

00318

you have to have the fresh tissue within the lab within 24 hours, and this may need some clarification. Also, we're dealing with home brews that are being done in certain labs, so the tissue has to go to that lab, there won't be kits. So what will be the accessibility of this? Not just -- so we have the reproducibility issue in doing that, but then in general, how would these be able to be done if there is only a few labs doing these?

DR. FRUEHAUF: Well, I think that if there's a favorable decision today, there will be many more labs doing this.

DR. HELZLSOUER: Well then, I would like to have more information even yet on reproducibility.

DR. FRUEHAUF: Yeah. The reproducibility thing is very important. And we're inspected by the College of American Pathologists to fulfill CLIA regulations, and we have to show precision, we have to show sensitivity and specificity, and we do that by looking at thousands of cases in our database to show that the population patterns remain constant.

00319

Most of the laboratories, and we work by getting specimens from all over the country, and we set up a system where Federal Express takes the specimen immediately after surgery and brings it to our laboratory. The other labs use similar courier processes. So it's not that it's hard to get a motivated person in the pathology department to send the specimen.

DR. HELZLSOUER: Let me ask you this, about your reproducibility studies. So they're

done using your known samples, so you have your known controls; is that what you're saying within your lab?

DR. FRUEHAUF: That's correct.

DR. HELZLSOUER: Is that what the regulations are? My experience with that, with dealing with laboratories, is that's usually not very reproducible when you're dealing with sent specimens that you do not know. So I wonder if those studies have been done in these assays to determine for samples unknown to you --

DR. FRUEHAUF: Yes.

DR. HELZLSOUER: Sent specimens.

DR. FRUEHAUF: That was done. SWOG did a study in the '80s where they looked at

00320

concordance between laboratories, and they sent the same specimens to different laboratories. And they found a concordance level of about 80 percent between laboratories for the same result, and I think that is really significant considering the variability of biological specimens.

DR. BURKE: (Inaudible).

DR. FRUEHAUF: It's from a book, Tumor Cloning Assays, that was published. Somebody might know that better that I do.

DR. BURKE: My question was, did you have a citation on that.

DR. FRUEHAUF: I can provide that to you after the meeting.

DR. BURKE: Because I share your concern about -- I mean even the most common tests have difficulty, most prognostic factor tests have a great deal of problems with reproducibility. CAP has been trying to do standardization for years in this area, on even automated type tests, and it's very, very difficult to do.

DR. FRUEHAUF: I can tell you what we've done. We have cell lines that we study.

00321

We have 25 different cell lines with

characterized drug response patterns. And we send these as unknowns into the laboratory on a periodic basis, to make sure that every day, every week when we're running the assays, we are getting the appropriate result for these cell lines, which are unknown to the people in the lab who are doing the assay. So we have an internal validation process with 15 to 20 cell lines that we run routinely to validate the Cisplatin result is appropriate for the ovarian cell line, for instance; that the adreomyecin result is appropriate for the breast cancer cell line, and this is an internal validation process which, we use the same one for doing markers, for doing HRCC New, and P-53 as phase for actions, where you have to have internal validations you run in your laboratory to confirm that every time you're running the test, you're getting the same expected results.

DR. KASS: Could I ask a follow-up question on that? In that particular study that you referenced, one thing that was of interest to me, we have seen lots of different types of laboratory tests, and I was wondering if in that

00322

study they addressed the results comparing the different types of assays that we have heard referred to today, have any studies been done to look at the comparability of the DiSC versus the MMT, versus whatever?

DR. FRUEHAUF: Yes, and I think that other people do this all the time. Dr. Weisenthal does three separate assays on each specimen that comes into his laboratory. What we did for GOG 118 internally, we ran a DiSC assay, and we ran an EDR assay on the same specimen, and we looked at the cut points of low, intermediate and extreme resistance, and we found that they were exactly concordant with a very small, one to two percent difference. So, the cut points are very important, reproducibility is very important, but all the people who are doing this have been doing this for 15 or 20 years and

have -- there was an NCI consensus conference in the '80s that addressed these specific issues of quality control, because this all stems from the NCI funding these laboratories originally to develop this technology. And it was partly done for drug discovery and it was partly done for helping patients get the right therapy. So the

00323

consensus conference looked at the issues of coefficient of variation, out wires, how many standards you needed to run with each assay, et cetera. And they set up a profile of quality control requirements internally in the laboratory that would be necessary, and they compared the different laboratories that were doing the testing, so that there would be a uniformity of process. And so we incorporated into our laboratory procedures those quality assurance and quality control measures, along with the internal standards being run all the time. And what we are doing now is using these cell lines to send to the other labs as proficiency tests, because we have to have proficiency tests to maintain the quality assurance.

DR. FERGUSON: Thank you. Very briefly, Dr. Kern.

DR. KERN: Yeah, very briefly. There was a study published by NCI a few years ago where we compared four laboratories, UCLA lab, Sid Salmon's lab in Arizona, Dan von Hoff in Texas, and Mayo Clinic, Dr. Liebe's lab, they were all sent -- all labs were sent 20 compounds, blinded, coded. Most of them were anticancer

00324

drugs; some of them included sugar and salt. And we published on the very close reproducibility of all four laboratories. I can provide you that reference.

DR. FERGUSON: Thank you. Dr. Loy?

DR. LOY: I just wanted to ask a question that remains in my mind, and that is, when is the optimal time to biopsy? Certainly you would expect tumor biology to change after,

or posttreatment, whether it be chemotherapy or radiation, and I'm just wondering if there's any studies to talk about or clarify when the most appropriate time to biopsy is, and if there's any predictive value in testing those tumors that have not previously been treated.

DR. ROBINSON: My name is William Robinson. I'm with the U.S. Harvest Medical Technologies Corporation. We didn't send literature to the panel, but we did get in on this at the end, thank the Lord. One thing we wanted to draw reference to was that question about timing, because according to a research paper that came out of NIH in 1981, they felt the most appropriate time was within the first four hours of biopsy, because I think according to

00325

Dr. Wing, that the gethaco protein does get inducted very early on, so therefore, you don't get a real response, a clear response to what the tumor looks like in vivo as opposed to what you actually see in the Petri dish. Some of the literature we sent actually does show you, for those who can actually see this, that we were able to pick up a metabolism very early on, within about minutes. So if it's a case where you're going to compare MTT tests and the DiSC tests, I think the idea is you want to get the tumor in the closest condition that it appears naturally, so as far as automation is concerned, and that's where we come in, we think that this is the kind of tool and the kind of forum for discussion as to how you combine therapies, that this makes this a very good and useful meeting. Thank you.

DR. LOY: Thank you for that, but my question was more directed towards when in the course of the history of the disease, is it pretreatment or posttreatment?

DR. FRUEHAUF: Acquired resistance is an important question, so that if somebody has a biopsy and you get a result and you treat the

00326

patient, and then the patient's failed primary therapy and you want to go back to your result. The question is, is that result still valid to treat the patient now, who's had intervening therapy? Is that part of your question?

DR. LOY: That is part of my question, but please address that issue.

DR. FRUEHAUF: So first, of course, you have two kinds of variability up front in newly presenting patients; you've got sit, inter-site, and for synchronous lesions, and so we studied in paired cases synchronous lesions and metachronous lesions. And we looked at extreme resistance frequencies for the various drugs, between sites and over time, for ovarian cancer. We presented this at AACR. We found that there is a very low frequency of a two-drug category shift, of about 5 percent, in terms of synchronous lesions. So if you looked at platinum resistance in an ovarian cancer patient and you compared the primary ovary with the peritoneal metastases, only 5 percent of the time was there a significant difference in the result. It went up to about 8 percent when it was over time, so the difference over time -- now, I think the key is,

00327

there's not a lot of heterogeneity and change in resistance patterns, but there can be a decrease in sensitivity, so that if you're using the assay to identify ineffective agents, an agent that's ineffective initially, after intervening therapy, was still inactive later. It was a loss of sensitivity that was occurring. So there is a robust ability to say if the drug wasn't going to work up front, it's unlikely after failure or progression that that drug is now going to work in the relapse setting.

DR. LOY: Have you found the same thing or have studies been shown to show the same thing to be characteristic of hematologic malignancies, which are known to transform after chemotherapy?

DR. FRUEHAUF: I would leave that to one of my friends who does this research on that.

DR. FERGUSON: Mitch, do you have something?

DR. BURKEN: Just a quick comment on a study. Just -- I didn't get it in before. The issue came of up of concordance or discordance between different assay formats. And you know, there have been several studies; as a matter of fact, some of them were listed this morning. One

00328

of the studies, I'm not sure whether it was listed or not, was by Tavassol in Oncology in 1995, where there were 17 patients that had head to head FCA and EDR, and there was some discordance. At least 12 of the 17 patients had, or 12 of the 17 patients had at least two drugs that had different patterns. The problem with those kinds of studies, as I said, you run up against complicating factors like the tumor heterogeneity that we talked about earlier, where the differences may be due to the fact that there's just intrinsic tumor heterogeneity. And so, it does open up I think another vista of ways of looking at test accuracy.

DR. FERGUSON: Dr. Bosanquet, did you have some response?

DR. BOSANQUET: Can I address a couple of these points? We very early on looked at different biopsy sites for the hematologics, and compared drug sensitivity. So we looked at blood, bone marrow, lymph node, and found almost identical drug sensitivity from those three sites, in CLL and non-Hodgkins lymphoma, similar diseases.

We have also -- I also concur in the

00329

ovarian data that John Fruehauf has just mentioned. We also in our laboratory find almost identical results between a situs and a primary tumor in the ovarian setting.

The point was raised about the timing of the biopsies. In 1988 we published a paper in Cancer, which hasn't been mentioned, in which we looked at drug sensitivity before and after an

intervening period of time. If there was no intervening chemotherapy, there was no difference in drug sensitivity from one to the subsequent test. If there was intervening chemotherapy that was not the drug that you were testing -- I'm sorry -- if there was intervening chemotherapy, for instance, with Doxorubicin, and you looked at the difference in chlorambucil sensitivity before and after the Doxorubicin, there was usually a slight increase in resistance, and chlorambucil resistance.

If you looked at the drug that had been given in between, so you tested Doxorubicin, then you gave Doxorubicin, then you tested Doxorubicin again, you saw a greater increase in resistance between the two tests. There was one anomaly to this finding, this universal finding, which is

00330

becoming a standard chemotherapy in CLL in Britain. And that is that we found that if patients were treated with chlorambucil, and this is just in CLL, if patients were treated with chlorambucil, they became 10-fold more sensitive, or there or cells became 10-fold more sensitive to the steroids. And this is an anomalous finding, which is really quite exciting. And if you look in the original literature on steroids, not much use in untreated CLL. But we found them, high does methylprednisolone for instance, to be very effective in previously treated CLL, supporting this finding from the laboratory. So that's, as far as I'm aware, the only time that increased sensitivity is induced by treatment.

DR. FERGUSON: Other questions from the panel members or comments? If not, we will reconvene tomorrow morning at 8:00.

(The panel adjourned at 4:33 p.m., November 15, 1999.)

Transcript of November 16, 1999 Morning Session

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

00331

VOLUME III
(Morning Session - November 16, 1999)

HUMAN TUMOR ASSAY SYSTEMS

HEALTH CARE FINANCING ADMINISTRATION
Medicare Coverage Advisory Committee
Laboratory & Diagnostic Services Panel

November 15 and 16, 1999

Sheraton Inner Harbor Hotel
Baltimore, Maryland

00332

Panelists
Chairperson
John H. Ferguson, M.D.

Vice-Chairperson
Robert L. Murray, M.D.

Voting Members
David N. Sundwall, M.D.
George G. Klee, M.D., Ph.D.
Paul D. Mintz, M.D.

Richard J. Hausner, M.D.
Mary E. Kass, M.D.
Cheryl J. Kraft, M.S.
Neysa R. Simmers, M.B.A.
John J.S. Brooks, M.D.
Paul M. Fischer, M.D.

Temporary Voting Member
Kathy Helzlsouer, M.D.

Consumer Representative
Kathryn A. Snow, M.H.A.

Industry Representative
James (Rod) Barnes, M.B.A.

Carrier Medical Director
Bryan Loy, M.D., M.B.A.

Director of Coverage, HCFA
Grant Bagley, M.D.

Executive Secretary
Katherine Tillman, R.N., M.S.

00333

TABLE OF CONTENTS

00334

TABLE OF CONTENTS (Continued)

VOLUME III

00335

PANEL PROCEEDINGS

(The meeting was called to order at 8:05 a.m., Monday, November 15, 1999.)

DR. FERGUSON: Miss Tillman, our right hand, is here, and has some announcements and pronouncements.

MS. TILLMAN: Good morning, and welcome again. First of all, I am going to read the conflict of interest statement again. Conflict of interest for the Laboratory and Diagnostic Services Panel meeting November 15th and 16th, 1999. The following announcement addresses conflict of interest issues associated with this meeting, and is made part of the record to preclude even the appearance of an impropriety. To determine if any conflict existed, the Agency reviewed the submitted agenda and all financial interests reported by the committee participants. The conflict of interest statute prohibits special government employees from participating in matters that could affect their or their employer's financial interests. The Agency has determined that all members and consultants may participate in the matters before the committee today.

00336

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

In addition to that, we request that anyone with a cell phone please turn it off, so it doesn't disrupt the discussion this morning.

Also, as the speakers either come to the microphone or the panel members begin to speak, if you could identify yourself for the record, since we have a court reporter here, and it would make it easier for him to identify who's speaking.

And also, anyone who would like a transcript of the meeting can contact Mr. Paul Gasparotti, with Salomon Reporting Services, and he can make a transcript available for you.

Dr. Ferguson?

Opening Remarks - Introduction

DR. FERGUSON: Thank you. This morning, we'll be primarily discussing among the panel members, and then voting on the questions, or the questions proposed as points to vote on.

00337

I would like to remind those in the audience that we would like to keep this restricted to the panel's discussion, except on points of reference where we may need some points clarified from members who presented yesterday, until the 11 to 11:30 session, which we can open up to four or five minute remarks by some who presented their work yesterday.

Open Committee Discussion

DR. FERGUSON: Now I would like to start this morning and kind of go around among the panel members to get their ideas and their comments and critiques and concerns, and questions on what we've heard, and anything they think that is important that we should know about. Maybe I could start over there on the far right.

DR. MINTZ: My concern here is trying to hone in on -- these tests are reasonable, and it's reasonable in a setting of a malignancy to do this. The question I think with which I'm wrestling is when are they necessary. And it's hard to hone in on the data that we have seen on specific situations where we find them necessary. I look forward to further comments

00338

this morning on can we identify situations and disease states where we feel collectively that this test is necessary.

I am just beginning to read the articles by Dr. Bosanquet here in CLL, but I do find them interesting, and I'm ready to be persuaded, but I would like a little time to look through them. But as we address these questions this morning, I am interested in hearing further from the participants yesterday as to where we can find specific situations where we might deem this a necessary test. And I look forward to my colleagues trying to identify that situation. I at present have not identified such a situation, but I am open to being persuaded.

DR. FERGUSON: Very good. Kathy?

DR. HELZLSOUER: Yeah. I agree that intuitively these make sense to be used, and I am wrestling with the same things. I think cancer is too large of a disease entity, and it seems that there probably are settings, and maybe CLL is one of them, where these tests are appropriately used. There is the issue of metastatic versus primary settings, or adjuvant setting, or trying to sort out once they've been

00339

previously treated.

And I'm not convinced, although we heard a lot about quality of life, and I believe that's a very good clinical indication for this, that if you can avoid unnecessary chemotherapy, that's extremely relevant and important, but I'm not convinced yet, given the specificity of the overall test results, that we have 80 percent,

plus the 10 to 20 percent problem with acquisition of tissues appropriate processing, how many will be spared. And my readings of some of the graphs yesterday and some of the articles here, that if you still have 20 percent that, that's the specificity in the combined group, would you feel comfortable eliminating that for an individual, because there is still a chance 20 percent of the time they would still be sensitive in vivo, they will still respond. And when you get down to a metastatic setting when people will choose something for even a 1 percent benefit, it's hard for me to see that you will be eliminating a lot of chemotherapy. So that's one thing that I would like more clarification on.

I think that the problem is that there isn't much information on the clinical outcome,

00340

although there's a correlation with survivors, it's the problem we always have with reviewing issues, that responders always do better than non-responders, and it's probably a good marker for responders. But I think we have to see how we can clinically use that in either choosing chemotherapy, and I think the compelling argument is the issue of avoiding unnecessary chemotherapy, but I'm not sure I have the evidence to say that would actually be done in practice.

DR. FERGUSON: Thank you. Miss Kraft, do you have some?

MS. KRAFT: Cheryl Kraft responding. First of all, what was pointed out yesterday was two different percentages of how many people don't respond to any type of chemotherapy or cancer treatment, one being 70 percent and the other being 76.3 percent. So it's clear that we don't know how to manage cancer patients so that they can survive. So the question to ask is, will the tests that are available to us, these human tissue assay systems, help us in prolonging or help the physician in treating the patient?

From what I can tell in the studies

00341

that I've read, that the use of these tests and the way the doctors use these tests in treating patients, that there were no negative effects to the patient or consequence to the patients, with the exception of one trial that was outlined in one of the studies. So that being, do these tests then have, test for drug resistance at a sensitivity that is great enough so that the physicians can interpret the benefit to the patient? Well due to the fact that drug resistance is growing and is definitely multifactorial, as one of the articles said, and the heterogeneity in cancer tumors is great, then are we not, and myself, trying to make sense out of all the articles I have read, and in which cancers and which drugs should be treated for which specific cancers, are we not maybe trying to fit a heterogenetic tumor into a box? I think analytical people try to fit everything into a box.

And so what I would like to put forth to the panel is that maybe we should step out of the box and we need to look at, since again, all these tumors are heterogenetic, should we look at just continuing to do what has been done, that

00342

being continuing to test all types of these cancer tumors against all the drugs available to us and see, and continue to treat patients accordingly? Now none of the patients in none of the articles were denied treatment of drugs that they were considered to be resistant to, so taking that into consideration, maybe the studies should continue to be done.

However, during that time, this panel needs to think of should the patients, even though this may not be definitively designed for a specific tumor and a specific drug, should the patients really be denied a test? And this is a laboratory test we are talking about. Should they be denied a laboratory test that could possibly benefit them?

I think, again, this laboratory test is a tool for a physician. The physician should take advantage of all the tools available to him to treat a patient. And since studies show that only 25 to 30 percent, again, of patients do respond to the test and/or the drugs and/or the correlation of the drugs and the chemotherapy that we have available to them, should we not consider, due consideration to looking at the

00343

advantage of these human tissue assay tests and the resistance that has been found to chemotherapy drugs?

DR. FERGUSON: Okay. Thank you. Dr. Hausner?

DR. HAUSNER: Dr. Richard Hausner. For me, I would like to take the approach to try and put my comments in the context of my own clinical experience, my own day-to-day, I'm a working pathologist, although I am on the active clinical faculty of Baylor College of Medicine and the University of Texas Health Science Center in Houston. I practice in a community hospital, but I have very long reach in terms of my clinical experience. I have a big practice. And I can tell you that in Houston, Texas, where there is quite a bit of health care going on on a daily basis, not once ever in my life, with all of the cancer patients that I've seen, have I once been asked to harvest tissue for this procedure. Not ever. And I can tell you that if any of the patients in my practice had had this testing, that we would have been involved in the harvesting by definition, because the surgeons would have surely asked. So I know that it

00344

hasn't happened.

Nevertheless -- and I came in here reading the source material with that bias, because I had that bias from the very beginning. But nevertheless, somewhere around the middle of yesterday afternoon, my thoughts began to crystallize, and they crystallized during the

time that, in the afternoon session when the data was put up to a tremendous amount of scrutiny and a very sophisticated critique, and I thought that it held up pretty darned well. And I have come to the conclusion that while over the past 20 years of the research that has developed for this technique, it clearly was a research tool and not ready for prime time, that the decision was correct not to allow this into Medicare's realm and therefore, give it the validity to go forward.

Because what is someone's exciting front line technique comes very close to someone else's quackery, and at some point it would have been premature to allow this. But I believe now that the third generation technologies clearly take this beyond a research tool and that from this point forward, I would hope that the

00345

clinical studies will be conducted to refine where this could be best used.

Another analogy would be that a, that if this technique is not permitted in its current state, then the panel ought to reconvene and consider removing microbiologic sensitivity testing from the armamentarium of physicians, if this is not approved. The truth, I believe, lies somewhere in the middle, therefore, and just like so many other things we do in medicine, that this is a useful tool, imperfect as it is, and the ground rules may have to be carefully defined, but to turn the test away in its entirety, I believe would be inappropriate.

And in closing, I would point to the final paragraph of Dr. Weisenthal's paper in which he talked about whether we use the civil or criminal criteria of preponderance of evidence versus beyond a reasonable doubt. Beyond a reasonable doubt, we don't have. Preponderance of evidence, I believe we do. And therefore, my conclusion is, as a rough sketch, is that something ought to be done towards bringing this test into, as another tool for physicians to

use.

DR. FERGUSON: Thank you. Dr. Kass.

DR. KASS: Thank you. Mary Kass.

First of all, I think my first question was about the testing methodology, but I think that there is overwhelming evidence to show that these tests meet all the normal QC, all of the normal standards that all other laboratory tests have to meet. I think that they're valid, I think that they are reproducible, so the third generation of tests for me is no longer a concern in that respect.

The question has been raised about necessary versus clinical utility. I don't know how to define a necessary laboratory test; I think that's really in the mind of the user. When I was in training, which wasn't all that long ago, the emergency room of a downtown urban hospital in Washington, D.C. didn't even have a laboratory open from midnight until eight a.m. because there were no laboratory tests that were necessary to make clinical diagnoses. But we've come a long way since then, and I think medicine has grown and realized that there are many things that can help physicians do a better job in taking care of their patients. So the clinical

utility of this test, I think has been demonstrated, to certainly my satisfaction.

The fact that the test is difficult to do because you have to acquire fresh tissue, it has to be shipped in a certain way quickly to a laboratory, that doesn't bother me either. That doesn't change its utility. I remember when we first started doing flow cytometry, the transport of specimens to do flow cytometry on was a big challenge to us. Now we do it routinely and we don't lose specimens in the transport process.

It is very intriguing to me that this particular methodology may be very helpful in evaluating new drugs, the number of new chemotherapeutic agents that are rapidly being

introduced to try to help us have a greater impact to the treatment of cancer. I think that anything that we could use to help define which modalities have a greater possibility of working and which don't, would be very helpful. I think it also allows the earlier consideration of other treatment modalities for patients, rather than going through a whole course of chemotherapy and waiting for the end point of no response. Earlier in the course of that, a clinician may

00348

have an opportunity to switch a chemotherapeutic drug, or remove one which has a very toxic side effect from the treatment regimen.

I guess in summation, I think that we haven't done a terrific job in treating most of the solid tumors. I think everyone is very disappointed in the fact that we haven't been able to have greater success than we have. I think that this is another tool, one of many, that could be available to clinicians that might help, certainly in terms of the quality of life, if we could remove drugs from the treatment regimen that were not effective, and perhaps in a better outcome.

I think the patient that testified yesterday, that's one case, it's anecdotal. However, I've practiced pathology for 32 years; I have never seen a patient with widely disseminated pancreatic carcinoma that survived. You have to take notice of that. I think that's worth listening to.

So, I think that's the summation of my comments.

DR. FERGUSON: Thank you. Miss Snow?

MS. SNOW: I'm Kate Snow. I'm the

00349

consumer voice for this panel. I listened very intently to all of yesterday's testimony and I agree that Mr. Stein was very compelling, and I too have never seen a pancreatic cancer survivor. However, I did not know how old this gentleman was, or if he had other comorbidities.

I believe that if I were a cancer victim, I would want this study available for my use. I would feel it was reasonable and I would also very much feel it was necessary.

Listening to the quality of life and the cost of life that could be gained, and to decrease the burdens for individuals was also very compelling. If it takes the guess out of the therapy that's used, I think it's a very good tool to have available to us.

I struggle with whether or not this test will be available in a way where those of us in northern rural Michigan will have access to this kind of tool or not, and what that might look like in the future.

I do feel there is a possibility for a cost effectiveness. It may need some more research and looking into exactly how cost effective this could be, both for the medical

00350

community as well as the beneficiary.

And I think that's all I have to say for now.

DR. FERGUSON: Thank you very much. Dr. Loy?

DR. LOY: I'm Dr. Bryan Loy, and I listened also very intently yesterday to the presentations being made. I have a couple of comments, first of all regarding the presentations. I noticed a number of cancers were being elaborated on. I am still not clear at what point in the clinical progression of the disease, or how often the testing should take place.

However, having said that, this does sound like this is a tool that be could be very useful. But having listened to the presentations yesterday, again, we were focusing on specific cancers, and to try to take that tool and apply it to all cancers at this point in all clinical scenarios, doesn't seem to be quite reasonable at this point. We really didn't talk a lot about the sarcomas, or trying to talk about such broad

fields as hematopoietic neoplasms. I think at least in my mind, I would need some more

convincing evidence to try to apply this technology wide spread, and I think that this is certainly germane to a policy type discussion.

The other piece that's still lacking in my mind is where this really fits clinically. Because some cancers are clearly curable with chemotherapy, or they're curable with radiation therapy in combination with chemotherapy, or they're curable with surgical resection, or any of those combinations. And trying to really fit this into that niche is going to be quite difficult to do from a policy perspective.

Having said that, I think that there certainly is some promise. I think there is some utility that has been potentially demonstrated here, but I am not clear on where this fits yet.

DR. FERGUSON: Thank you. Dr. Murray?

DR. MURRAY: Thank you. I am Robert Murray, and I've kind of grouped my comments into four areas. The first point is that I believe what we are supposed to be doing is looking at the questions that were presented, the six specific questions that we would like to come to grips with and arrive at answers to. I sense that we have been taking the view from 35,000

feet and not from the detail level that we need to or that we were asked to. I am concerned about that, and I really think that we have looked mostly at number 5, which asks, is there evidence to support clinical utility? I sense that from the speakers who have voiced their opinions and also reflecting my own that the answer is yes, there is evidence for utility in certain cases.

The second point is, we're stumbling over reasonable and necessary versus clinical utility. Reasonable and necessary is in the statute, and Dr. Bagley gave us a couple examples of how you can assess reasonable and necessary.

I view it as a term of art. I don't think we look first at reasonable in isolation, and then we look at necessary in isolation. We'll certainly get thoroughly enmeshed in what kind of necessity. As was already mentioned, no laboratory test may be necessary, mathematically necessary. You can certainly find alternatives. But nonetheless, utility is perhaps an equivalent term that we have in fact focused on. But again, we are looking at a very high level.

And Dr. Loy's comments yesterday, and

00353

his comments just a moment ago, I think remind us that we need to come to conclusions that are going to allow a very high legal of specificity. We can't just say, I don't think we should say, at the end of this session, yes, there are some situations when some testing might be appropriate. That is simply not the level of guidance that we need. Some of us went through negotiated rule making over the past year and we realize how difficult it is to draft a national coverage decision with the uniformity and specificity. So I am concerned about the fact that we are, we seem to agree that there are some situations in which there is utility, but we're far from reaching the level of specificity that we ultimately will need.

The third point is just my own reaction. Spending my life generally in the laboratory, I tend to analogize all of the situations, the questions, to existing laboratory tests. There is no question that many laboratory tests which are routinely approved currently have nowhere near the evidence, nowhere near the accuracy and predictive value that the tests that we're considering today, that we heard about

00354

yesterday, have already demonstrated. Yes, we do have to look at outcomes. We have to look at outcomes measured in different ways. We have to look at evidence. But the evidence, even if the bar is raised higher, the evidence that we have

heard certainly exceeds the evidence that we have for many, many tests currently in use.

My fourth and last point is actually two very minor specific objective questions, and perhaps Dr. Bagley can respond to one or both of them. In Dr. Weisenthal's paper that he included in the packet, there is a reference to a Medicare hearing in April of 1998 which seemed to indicate that it was a decision of what I would assume was an administrative law judge, ruled that these tests would be covered. And my question is, which I'm not asking for an answer now, but sometime before noon, does that decision affect our decision here? If a judge has already ruled that they are coverable, then what are we debating?

And the last and very minor point, a question that perhaps one or perhaps several of yesterday's speakers could answer, are any of the tests that have been suggested, the tests

00355

performed that are being currently offered on the market, are any of them covered by patent protection? Are we doing anything, are we making a decision on issues that would force or encourage or would support limitation in the availability of the test? And again, I am not asking for an answer now, but if sometimes perhaps during the open discussion, I am curious what level of patent protection there is currently, could these tests be offered by any laboratory if they were approved? That's all of my comments.

DR. FERGUSON: Thank you.

I had a number of things, mostly in the form of questions myself. But I guess some can be considered comments. First, there are several different tests done by several different groups that we were exposed to. Not all seemed to be equal or equivalent to each other, they were used in many different kinds of cancers. This leaves a large number of combinations and permutations for us to grapple with. And it's hard to put

them, as a matter of fact, I would say it's impossible to put them all in one basket and say, you know, treat them all together. At least I would find it difficult, given the amounts of data and studies that we saw, all for different tests and so on, so this in my view makes it a difficult job.

A second, that many of the studies that we saw were on the small side, small numbers of patients.

Number three, it wasn't always clear to me how the patients were chosen for these studies or from what populations they were chosen. In other words, what the denominator was, how did these patients get into the study. Sometimes it was. I'm giving sort of an overall, at least what my concerns were. It's clear that these patients had to be self selected in a way that there was an accessible tumor to be biopsied or surgically removed, that some patients who perhaps had recurrences weren't available, once they had tumors that recurred, because they were deep or in bone, or inaccessible in some other way, or weren't willing to put up with biopsies and so on. So that there were patients that might possibly benefit but couldn't because they didn't have tumor available. Whereas the tumors, the easily accessible, perhaps is in leukemia patients and lymphomas, where tissue is reasonably easily available, and maybe that's a different group. I mean obviously, maybe they're self selected in that way to be better and more responsive. But any way, it is a bit of an issue, I think. It's hard to treat them all equal when you need tissue in order to do this test.

The fifth point, it seemed like to me on most of the studies we were dealing with advanced tumors, mostly recurrent after stage 2. I wondered how many actually stage one and stage two type patients had been studied.

Number six, that in -- it seems to be a number of papers alluded to the fact, or studied the fact that even cancer cells from the same patient were different, in other words, that the primary site tested different than the metastatic site. Which brought up the notion, and this was again stated and makes it somewhat difficult, that patients had been treated, their cancers now become more resistant and test differently with these tests. So this is just another factor which makes the testing, you know, when you test, after treatment, before treatment, and whether

00358

you test a metastatic site, the primary site, do you still have that and so on. This all adds other things and as Dr. Loy mentioned, when is this test most useful? And so this just raises to me another set of questions.

Then I think what was mentioned by Dr. Barnes yesterday, in a number of areas there are several histologic types of cancers, so we weren't always given that kind of information, and whether they all test the same or might test the same. Ovarian cancer is a multidimensional animal, as I understand it.

So, those were my concerns. Having said that, I also felt that in some of the studies that were presented, I was impressed with some of the leukemic studies and some others that there is some usefulness and that it needs to be mined, but mined carefully and under the right conditions.

Just another comment about randomized trials. Where I sat at the NIH as chair of the technology assessment committee for the American Academy of Neurology for a number of years, the number of randomized trials with outcome measurements for diagnostic tests, I don't

00359

believe I could have counted on one hand. I would have to look very hard to find those tests. I remember seeing reference to one or two, but -- and there may be more, but I think

there is no question that for diagnostic tests, randomized trials with good clinical outcomes are extremely rare and I believe that, however, they should be done. We need better standards.

Dr. Bagley?

DR. BAGLEY: I would like to bring up a couple of other, or reiterate a couple of other notions which I just want to sort of bring to the forefront for us to keep in mind as we consider these.

Dr. Loy brought up a very important point, I think, yesterday. And it's one that's easy to lose sight of when, as Dr. Murray said, looking from 35,000 feet. And that is that any recommendations that you make, that are then placed or implemented in the policy, need to be done with some specificity. We normally don't write policies that simply say, pay for test whenever a patient's physician thinks it's necessary. Now that might be a reasonable policy, but Medicare isn't designed to work that

00360

way. And in fact, we've learned from long experience that if we do things that way, that while it works 99 percent of the time, the 1 percent of the time that it doesn't work, it is a disaster, because there are, there is fraud and abuse in Medicare. It is a very very small proportion of what goes on, but it accounts for a large portion of the dollars, and they're the dollars that belong to the beneficiaries of the program, and they need to be protected.

And perhaps that's the reason that Congress gave us the admonition that we shouldn't just pay for medical service, we shouldn't pay for medical service that a patient or a physician thought was reasonable and necessary, but actually the prescription that's written in the law is written in the negative. It says, Health Care Financing Administration will make no payment for a service unless it is reasonable and necessary. That means there has to be some policy determination and there has to be some

review, some process by which we determine the things that are reasonable and necessary.

Now with diagnostic tests, it's perhaps a little more difficult than it is for therapies.

When we're talking about treatments, we're talking about options that a patient can take, and in fact, they can select from one of the options, and as long as there is evidence that they are reasonable choices, it then becomes a little easier to come to the notion that it's reasonable and necessary. But diagnostic tests become a little bit more difficult, because diagnostic tests, after all, give us information.

Patients want information, physicians want information, and we're all taught that the more information we have, the better off we are, more information gives us better results. But that's not always the case.

First of all, that, when Medicare views a service or a test or a drug, or anything else, as a covered service, therefore, it will be paid for by Medicare, it's easy to lose sight of the fact that that doesn't -- it's paid for by Medicare. It means it's paid for by the beneficiaries in the Medicare program. Medicare is after all a program which is funded by the beneficiaries, and the future beneficiaries, which is all of us. And in fact, the payment comes from that source, and in fact it doesn't

come entirely from that source. Some of it comes directly from the pockets of the patients who are receiving the service.

And what we're talking about here is a combination service. Some of what we're talking about may ultimately come under the heading of laboratory testing, but a great deal of what we're talking about is not laboratory testing, but it's physician service. It's interpretation, physician interpretation, and it really comes under the heading of consultation, it's a physician consultation. And when it is paid for

by Medicare, so called, it means that 80 percent of it comes from the premiums which are paid by the Medicare beneficiaries, premiums that are paid for the part B Medicare service, which is optional, although most beneficiaries do opt for that. But they pay a premium every month and that premium pays the service. That premium is determined in some part by the amount of payment in the program. It's a health insurance premium. And the remaining 20 percent comes from the beneficiary. So these tests are not tree.

Now, it doesn't matter if they're free or not. If someone has a fatal disease and they

00363

offer hope and they offer an improved way to treat it, then we don't really put a price on that, nor do the patients. But we need to remember that there was that last notion that I put up yesterday when we talked about what was reasonable and necessary. And that was, once something's safe and effective and once something has demonstrated utility, and the risks outweigh the benefits, then perhaps we should also look at the issue of whether or not it adds value.

Now value is not a new concept, it's not a new concept in considering medical treatment. We've always done that. We've always done that with diagnostic tests. As long ago as when I went to medical school, the notion was given to us very early that tests are not something to be used indiscriminately. When you order a test you should consider, is the information needed, is it going to make a difference and therefore, am I properly using this resource? So the notion of does it add value to the patient's treatment is very important, and I think that needs to be kept in mind.

And then the next point, which follows

00364

on what Dr. Mintz said, is it reasonable, is it necessary? I think it is an important concept, because for it to be reasonable and necessary

both, it needs to offer not only information, not only information which may be correct, but information which is likely to influence the course of treatment. Is it information the patient and the physician need? If it's going to guide therapy, then it should give us a decision in which we should do this. Now there is an interesting, there's an interesting interplay between what's reasonable and necessary.

We have recently been looking at the same thing in terms of what's reasonable and necessary in terms of a test with regard to using PET scans for many diagnostic uses. And in many ways it's the same kind of a process we're going through here; we're saying when is it reasonable and necessary, and for what conditions, because it's used for many many things, and the evidence is stronger for some uses than others. And as we approach that we say, well, if it's reasonable and necessary, then it's diagnostic information which is useful. Now we had just such a situation when we considered the use of a PET

00365

scan for evaluation of a single pulmonary nodule. The argument was that a single pulmonary nodule evaluated with a PET scan in which the nodule turns out to be not metabolically active or occult would eliminate the need for a biopsy and in fact, we can do a PET scan, if we get a negative result, we don't needed to a biopsy. But that's certainly a powerful argument, and for a patient making a decision about whether or not to have an invasive biopsy, it certainly is a reasonable option.

But we then looked at it and said, well then, if we use PET scans to eliminate the need for an open biopsy, how would we view an open biopsy that was performed after a negative PET scan for a single pulmonary nodule? Then we would be left with the dilemma of saying if the PET scan was reasonable and necessary because it could prevent an open biopsy, then was the biopsy after the negative PET scan reasonable and

necessary? It's hard to say they were both. And we are faced with the same dilemma here. We have a test, which we are told is useful to patients, because it will allow us to more accurately select their chemotherapeutic agents. We can

00366

avoid toxic agents which won't be effective, or we can select new expensive agents which are only used when they are effective. Now that's a persuasive argument.

But then we had the organization which represents most of the oncologists in the country stand up and say they are neutral on this procedure, but the one thing they're sure of is if we allow this procedure, we shouldn't pay attention to the results. That's what they said. It should not be used to withhold therapy. Which means, if a drug is shown to be resistant we shouldn't withhold the drug, based on the test, or if a drug is shown to be not sensitive and it's an expensive drug, we should use it anyway. That seems to me to be hard to understand. That you can take a neutral position about a test and say it looks okay, we think it's reasonable to do it, as long as we aren't asked to pay attention to the results.

And that gets back to the final point I made yesterday, is that in terms of looking at the evidence and one of the things is to look at the evidence and say, where does it take us clinically? And not just say, it's good enough

00367

to pay for but not good enough to pay attention to. We should say, it's not only good enough to pay for, but the evidence is so strong in a given area and perhaps it's a given tumor, perhaps it's a given kind of patient, perhaps it's for given drugs, but if the evidence is strong enough that we should pay, we should not only pay, we should promote and at some point we should insist, because after all, if it was reasonable and necessary to do the test, if we then ignored the test in future therapy, would in fact that

therapy be reasonable and necessary?

So, I am simply putting those problems that we deal with in writing policy into context for you, because I think you need to keep those in mind as we answer these questions. Because as Dr. Loy said yesterday, we need to have specificity because reasonable and necessary as a test means it might be reasonable and necessary to pay attention to the results when that happens.

And then just finally Dr. Murray's question about what happened in terms of the fact that an administrative law judge overturned a claim denial. We are talking about policy here,

00368

policy which says, this is how it's going to apply to the entire Medicare population. And when we write policy, it applies to everyone. When we write a policy that says we will pay for PET scans for single pulmonary nodules, it means we pay for single pulmonary nodules for everyone. And we don't pay for another use that we haven't dealt with. Now that means that we've written a national policy and it applies to everyone, every carrier, every beneficiary, and every administrative law judge in the appeals process. It is binding on everyone. That's why we make national policy on bright line issues, when we know which side of the line the coverage policy ought to be.

The opposite is true when we don't have a bright line. We leave it to the carriers to make policy based on input from the carrier advisory committee and also to review claims on a claim by claim basis if necessary. And when carriers review claims on a claim by claim basis and make a denial of a specific claim for a specific individual, that individual by right can appeal that claim, and that appeal process if carried to its conclusion has a hearing before an

00369

administrative law judge. That administrative law judge hears the facts and can overturn the

carrier's decision to deny that claim, but can only overturn that decision if there is no national coverage decision in place which can influence that.

So what we're talking about here is a binding national process, not an administrative law judge. When the administrative law judge, and administrative law judges are with the Social Security Administration, they are not medically trained, and an administrative law judge hears an appeal, overturns it, it applies to that beneficiary and that claim only. It is not precedent, it does not apply to other beneficiaries, other claims, other carriers, or Medicare as a whole.

DR. FERGUSON: So, you remind me, your slide yesterday said that these procedures are not, there is a national coverage policy; isn't that correct? So the administrative law judge couldn't have overturned the noncoverage.

DR. BAGLEY: The administrative law judges are bound by national coverage decisions. In areas where we have a national noncoverage

00370

decision and an appeal for an individual claim goes to administrative law judge, the administrative law judge is bound by the national noncoverage decision. There have been cases where administrative law judges have overturned claims denials which were ultimately in conflict with national coverage decisions, and your question is, how can that happen? Well, it happens, and the solution to that is that there needs to be an overturning of the administrative law judge's position, the denial of the claim. That at times doesn't happen and the claim is paid, and if it's not appealed by the government, the claim is paid, so that process can lead to claims payment. But in general, the policy we're writing, the fact remains that the policy that we're writing is binding, it is national, and if all of the appeals don't go both ways, that can happen.

DR. FERGUSON: Thank you. Mr. Barnes?
MR. BARNES: Well, as you know, I'm the industry rep on the panel and as such, I have tried to be a liaison with the proponents of reimbursement, which I understood was my job. And to some extent I sort of feel like I'm

00371

sitting at the wrong table at this particular moment, and you'll understand. But let me offer just a couple very quick thoughts.

The process, a couple comments on process, I guess. One is that the industry representatives here yesterday heard some criticisms of studies, and in fact a lot of attention yesterday afternoon was paid to quality of studies and science, and it, while I'm not an advocate for the industry, it does make sense to see if they might have any particular thoughts to share with this panel today, and I would encourage the chair to allow that opportunity.

There was a lot of interaction with regard to Dr. Burke's presentation, but I think Dr. Burken's review of a great number of studies, and in particular the statement that a couple of things were reversed on some of his slides which panel members had been looking at overnight might prompt the industry to want to clarify a couple of points. So I guess I would encourage that to be allowed.

A couple of panelists so far have encouraged the panel to pay attention to much greater detail, to really define what cancers,

00372

what drugs, what tests, and I think that might have been a general sense of the panel yesterday. And so, my second comment on process is that the way that we have conducted this session over the last one plus days really isn't very conducive to all of the details that you three gentlemen have said you'd like to hear about. And I don't have a solution for that, but I think basically, I just don't see how that could have worked very well, given all the

different variables that you'd like to hear more details on, so it makes it kind of difficult.

The comments yesterday on the quality of the research did not, I agree with several panelists, did not seem to outweigh the general notion that there clearly is benefit to patients from this test, and that from time to time I guess kind of made me a little bit frustrated. I was sitting here thinking that there are a number of small companies or even very small labs who can't quite present the randomized clinical trial that many of us would like to see, the conclusive once and all for cause and effect, RCT. But they do seem to present a good mass of evidence suggesting patient benefit, as has been

00373

mentioned, and I certainly, if I personally or someone in my family was involved, I would like them to have access to this test.

Speaking about access to the test, I think it was Mr. Kiesner in his early comments yesterday who pointed to the numbers of hospitals that send specimens in the volume of testing that they do. I have spoken a little bit to these folks, and it would appear that a number of people in U.S. managed care organizations have gone through an assessment of this testing technology and have indeed decided that it is something that is of value and something that they are willing to submit information, rather tissue for. So I guess I would be interested in hearing more about who some of those plans are and what kinds of evaluations have happened previously.

The negotiated rule making came up, and I understand there in fact was a consensus document that was put together, and I'm sorry, I am not familiar with what happened during negotiated rule making, but my understanding is that there was a movement in the direction of recommending reimbursement for this testing, and

00374

also that the ASCO signed on to, I'm not sure

what the term would be, but did indeed agree, unlike yesterday's discussion where we heard they were neutral, and I think I would like to hear some more about that. I think there were a couple panelists involved, and it would be interesting to know some more about that. Evidently, the gentleman from ASCO was not aware of that.

In my real life, a health economist, I appreciate Dr. Bagley's information and perspective on valued added. That was the last step in the stair case of defining reasonable and necessary. I would be interested in hearing some more from Dr. Bosanquet, if that's possible. I understood he had quality adjusted life year information suggesting utility, and I believe a study that has been published that hasn't really been made available to the panel, it's in the Technology Assessment Journal, which most of the panelists probably don't see. But I think that kind of information has in fact been accumulated. So perhaps he will have a chance to share that with us later. Thank you.

DR. BROOKS: My comments are a little

00375

bit along the lines of a number of the others, and I just want to start out by saying --

DR. FERGUSON: Do you want to say your name for the record?

DR. BROOKS: Yeah. John Brooks. As a pathologist, I certainly came to this without much knowledge or interest, in a sense of one who would give chemotherapeutic drugs, and kind of evaluated it in the same way as I would evaluate any new upcoming test that we have actually to evaluate, almost every week I would say, in the clinical laboratories. As a pathologist, generally, my thought would be that we like to see more tests done and certainly, useful tests are very helpful to people. So you know, in evaluating the information that I got beforehand and that we heard here, I was certainly impressed with how much had been done. I was actually a

little bit surprised at how much had not been done, however. I mean, in some settings, it certainly seems to me like the data is there for utility. I am not doubting that the test, I mean it's been mentioned before by a number of people that, you know, I actually believe that the test does test resistance and so forth, and that we

00376

may have to decide which of the tests might be recommended, or maybe two tests, A and B, whatever, but histology specific type of data wasn't necessarily there, except in certain situations.

You know, the hematopoietic malignancies certainly seemed to have really pretty good hard data. For example, one question that kept occurring to me is how often in an individual tumor type, not a site, not just ovary, but a specific type in that ovary, because that's what the clinician diagnosis is. And that kind of information, for example, is available in CLL. And the articles provided to us this morning tell me that you know, for example, 12 percent of patients with CLL, which I view as a pretty high number for a very uniform type of disease, you know, where the markers, et cetera, are quite similar case to case, showed a difference. Whereas, if I had had such data in small cell, et cetera, I certainly would be much more persuaded that the test is useful to people.

In other words, if we're trying to define a policy, and suppose we had a tumor whose

00377

data showed it never, it always had the same resistance pattern, then you wouldn't want to necessarily, although I may be foolish, to order over and over again to get the same result.

All that said though, you know, I do view that there is data there that shows some clinical utility. The question in my mind then becomes, how would you write a policy, et cetera. What I would like to see, just

apropos, forget chemotherapy is something that says, you know, people who have diabetes ought to get a glucose test, and we'll pay for that every time. Okay.

So then I come to the issue of denial, because as it was brought up before, we had ASCO there saying that, you know, sure, go ahead and do the test, be we might not pay any attention to it, and we sure want to give the drug that's resistant. I mean, that's kind of what I heard. I was thinking about that overnight and I was wondering, well, okay, can I think of something to explain that position? And even though I'm a pathologist, I actually talk to people a lot. We're actually not in the closet, and we're out in the public, and I love patients, and in fact,

00378

patients call me all the time. Okay.

So let's take diabetes. So suppose we have a diabetic test, suppose we don't have a glucose test, and the new proposal is, I've got a better glucose test to tell if somebody has diabetes. And if they have a sensitivity specificity of 80ish percent, or maybe even 90 percent. So okay, now I have a patient and I am a clinician treating this patient. The patient has a negative new glucose test, and it's negative. Should I be denied giving insulin to this patient? It sounds pretty reasonable. But maybe, you know, the test is imperfect, and maybe by looking in the eyes and by looking at the weight and by looking at this and that, I could figure out a way as a clinician in an imperfect world, and certainly not with our glucose tests that we have now, that you could in fact see people should have insulin because of a, not gut feeling, but as a group of symptoms and signs that I see in the patient, that I should be able to give insulin to that patient.

So with that said, I'm a little bit torn in fact, as to the issue of if a test is really good and it shows a negative result,

00379

should you deny the drug being given. And I guess that's kind of where I am. And in a sense, you could look at things that happened in other arenas. I'm aware that when they were doing the surgical treatment, you know, surgery versus medical treatment for coronary artery disease, and the surgeons said that, you know, the cabbage, or bypass was better, it took us ten years to figure out what that was all about, mainly because there was no clinical trial.

So another question that occurs to me, and I don't have the answer, and I want to hear what happens the rest of this morning, is if this is approved, do we want just America to use it willy-nilly? That is, this person uses it over here to treat this person, et cetera, and nobody gathers any data. And I don't mean gathering data by the companies, I mean publicly gathered data. So another question therefore that occurs to me is that if we approve it for specific histologies or, you know, certain restricted diseases or otherwise, should the -- should this be for a period of time and should the data be gathered by an independent source? Now that's not to say a clinical trial. A clinical trial,

00380

you know, you have mentioned is very difficult to start, especially if it is only directed at the DiSC assay or whatever. But for example, you could just have the results go into a central repository, and the physician who used the test would be required to say exactly how he treated that patient, for how long, and was it on protocol or off protocol.

My final question is, what does this do to clinical trials? And we have large public groups, cancer, breast cancer groups, prostate cancer groups, et cetera, working very hard to define what are the best therapies for each of these cancers. Now those cancers at least are more defined amongst themselves. If we bring in such a test as this, do we undercut what's being done in those trials, because, in other words, I

want to dovetail them, but I need to know how. I don't want to see suddenly everything being done willy-nilly based on a test that after all, only seems to have an 80 percent or so sensitivity. In other words, there are patients who are resistant who may respond. I think that's the question.

So with, those are the questions I have in my mind, and I would like to listen to more.

DR. FERGUSON: Paul Fischer.

DR. FISCHER: In preparation for this meeting I called a couple of the thoughtful oncologists in the Augusta area and asked them about their knowledge and experience with this test. They were uniform that in their opinion that it was not something that was useful because people did not behave the way the test would predict when they were given chemotherapy. And what I realized after listening to the folks yesterday was that there are really two cultural views here. The one cultural view is histology driven. We look under the microscope, we see a particular histology, and we therefore know what drugs to give. The other world view is the test tube driven world view. And when they were showing slides yesterday, one of the speakers said, well, I know what to give because I see all these cells got killed. So it's a really different way of believing what is going on in the world.

And then the question for me becomes, do the champions of the technology who spoke yesterday, do they represent pioneers or nuts? And I think that's what HCFA has to decide, because does HCFA advance the standard of care where the average oncologist says we don't believe in this, but HCFA is going to pay for this so let's try it, or do they support the current standard of care, which is to let the histology world decide what to do.

The problem I have as a family doctor

is that in my total practice experience I believe that probably my patients have been hurt as much by chemotherapy as they've been helped. And that's even given some of the wonderful drugs that do respond beautifully to chemotherapy. But I regularly protect my patients from oncologists who have this histology world view. If there is a tumor, it needs chemotherapy. And if you didn't respond to the first one, you get the next one. So eventually you're given drugs that are more and more toxic, less likely to benefit the patient, but because you've still got tumor, you need another course of chemotherapy.

Now my way of dealing with this in practice is to be very selective of who I send my patients to, and some oncologists are conservative and some are not, and I avoid the

00383

latter. But I, you know, on a weekly basis sit down and talk with a husband and wife and talk to them because they had some advice from an oncologist that they should go through this chemotherapy and to be quite frank about it, it's not always a very balanced view of whether the patient is going to benefit from it. So I really believe we need to move beyond the current situation and put some brakes on this histology driven world view which encourages more and more chemotherapy given with less and less benefits.

And I'm not sure we need to say that we have proven that this technology is useful for every tumor and every drug. You know, obviously they haven't. But clearly they have in CLL. You know, I'd like the people who give chemotherapy to stop for a second and say well, gee, maybe there are some other ways to think about who needs what, and this seems as good as any approach currently available, and I would therefore, I will vote to support some sort of funding for this when we get to the end of the morning.

DR. FERGUSON: Dr. Klee.

DR. KLEE: Hello. My name is George

00384
Klee, and I guess as I looked at the data that we had sent to us early on, it looks like you have a valid laboratory test. But where the issue sort of comes is where does this fit into the practice of medicine and how does it improve patient care in the longer run? You know, if we were to look it in sort of a protocol design, a selection of which drug should be used in which diseases, that seems to be a legitimate application that could go forward. Looking at it on individual patients is where the issue seems to come into play, and if so, which patients and for what decision purpose are we looking at it?

The data that was sent with the packet seemed to indicate that the best utility for the test was in the negative predictive value. That is, in response to determine which patients are not likely to respond to chemotherapy. And I was pretty well convinced of that until yesterday's presentation by Dr. Burken. And going through the numbers that were in those slides, and I tried to tally them up last night, there are very few that are up in that 99 percent negative predictive value. You know, there was a few of them that had numbers less than 20 that had 100

00385
percent negative predicted value, but there was a lot of cases that were, a lot of publications that were referenced there that had negative predictive values in the 20 to 60 percent range, which doesn't really make too much sense if you have an a priori odds of 70 percent, that apparently these are not prevalence adjusted.

And although the presentations went through a lot of explanation of the Bayesian theory, it would be nice to have these numbers prevalence adjusted. But even so, if you start with this at a 50-50 with those lower odds, it's getting up to a point where it's not looking like this test would really be that useful. You know, if you had something that's, you know, you have a priori odds of 70 percent, and you can only get

it up to 80 percent chance that this drug is not going to respond, that's not very much of an improvement over the prior. You know, if you're taking it up so you've got a chance of one in a hundred that this drug is not going to respond, then I think we've got something we can use in a clinical decision. So I guess I would like to see some of that information further clarified and presented in the form that, you know, several

00386

presenters have indicated that the Bayesian theory is needed, and it's the post odds that we're looking at in terms of the negative predictive value on these tests.

I guess I'd like to expand on several of the issues that were raised by other members of the panel here in terms of, we've got a multitude of diseases, and where does this fit in? You know, we've got a multitude of drugs, we've got a multitude of subclasses of these diseases in terms of whether it be histology or whether it be in terms of just other clinical classifications as to the stage of the disease and things like that. It looked like it's too big of a matrix that we're trying to deal in here, and it didn't look like one size fits all for the answer to that.

And I agree with the general assessment that the leukemias, CLL in particular, seemed to be one where the focus looked a little more channeled in an area where we can say that there is a definite improvement based on at least a few trials that have been out there. But at the same time, when we look to see, there was a question I was asking there yesterday, has it really been

00387

compared to using this test versus not using the test. That seems like where you would find out whether or not there is a benefit for the test. And those studies have not been done.

And I would like to maybe inquire a little bit more of Dr. Bagley, that at our orientation session, you had mentioned that there

is some new activities going on with HCFA where tests can be prospectively monitored in terms of utility, and perhaps introducing this in a disease management strategy with controlled output, you know, similar to what Dr. Brooks was alluding to, but is that something that through the funding mechanisms of HCFA, that this could be carried out, since NCI doesn't look like they have carried this out recently. They looked at it many years ago. And if this is not part of the clinical trial approach, we need an alternate way to do this. It doesn't seem like it's something that's a yes no answer, perhaps a controlled introduction in a very focused disease with the output monitoring required as to what happened to these patients, was there benefit in terms of quality of life years or in terms of survival, or in terms of any other parameters you

00388

might put up. But it looks like if we're going to take the practice of medicine to this higher plane that has been alluded to several times, we are going to need to do it in a controlled manner. Somebody's going to have to pay for it. It's not reasonable, I think, to have the burden of this put back into the people that may be making the test, but it needs to be looked at in terms of those that would be benefitting from it and as it is part of health care policy, then it looks like it should be, you know, integrated through, whether it may be a joint venture between NCI and HCFA, to say how do we carry this out in a manner that we can say, give this a controlled trial over X number of years, see what's going to happen, only apply it in a very limited focused area, for example CLL, and then try to see, did it make an impact, did it change the way that we're caring for these patients? Did it change the benefit in terms of life expectancy or quality of life years for the patients.

DR. BAGLEY: Let me answer that real quickly, and I'll try to make it a quick answer,

because that's a long complicated issue. And it
00389
has to do with when should Medicare become involved in issues that are not quite settled, how should we continue to look at things, and what can do? I'll give you some examples and tell you where we do do it then, and then tell you how limited that option is.

In two areas of diagnostic testing, we have recently written cautious coverage policies, not that different from what we're dealing with here today, new technologies have a lot of promise, clinical community doesn't know quite how to use it, they are not so sure they're going to use it to replace something else, we're not so sure it's going to improve care, but we cautiously advance coverage, and say we'll not only pay for it, but in paying for it we will collect some information. Now we can't collect very much information. By law, we can only collect the information necessary to process the claims. But we can interpret that to the point of saying we can get a certain amount of clinical information because that's what we process claims in.

Those two example list are magnetic resonance angiography of the head and neck, which
00390
is new and controversial. We added some coverage for MRA for head and neck vessels, and at the same time we cover it, we gather some information on what the indications were and how it's used, so that we can continue to evaluate it and say is it making a difference.

We're doing the same thing with the other example I mentioned, PET scans. PET scans for single pulmonary nodules which, you know, the promise was, it's a better less invasive way to monitor these patients. We're collecting the information from the claims in such a way that we will have an idea of what the experience was, and are these patients avoiding biopsy, or are they getting some other ones. That's one place to do

it.

But remember that in a very limited way we have gone forward and said that we are very cautious. And we have to be very clear that we're not doing research, because research is not only not Medicare's job, it's prohibited. So we are not doing research, but we are at least monitoring early diffusion of technology.

The third place we have done that is in another area called lung volume reduction

00391

surgery, which is a new surgical procedure. It has the promise of helping patients with emphysema. It was early touted as a miracle cure, some people still believe it is, and we are promoting it very heavily. And we looked at it early on, and realized that this would be another issue such as is turning out to be, bone marrow transplant for breast cancer, something that after ten years not only was never proven to be beneficial, in fact it turns out it may be harmful. And so, lung volume reduction surgery, we're gathering information at the same time we're paying for it in limited clinical trial.

But in general, it is very difficult and may not even be possible, although it's tempting at times for an issue like this, to say gee, it's early, let's pay for it and keep an eye on it, and then we will sort of stimulate the research.

In the previous panel we had a multiple myeloma, which is at the current time investigational, and the question was, is it ready to move to prime time? We made it clear to that panel and I think it ought to be clear to this panel and future panels, that when Medicare

00392

covers in an area that -- the day of the clinical trial is gone. We can accumulate information from case series, we can watch patients, but the day of the clinical trial is over when Medicare starts to cover it. There will be no randomization for that procedure ever again,

because we have made the decision by coverage of saying randomization is no longer necessary, we have the answer.

DR. FERGUSON: Dr. Sundwall?

DR. SUNDWALL: Thank you. I feel fortunate going at the end of this discussion, so I get the benefit of all these experts, and I don't mean to be flip about that. I really appreciate the expertise of the panel, more analytical than those of you who carefully reviewed the studies. My perspective is that of a family physician like Paul, who is taking care of patients, diagnosed cancer, followed them through chemotherapy, and seen sometimes the benefits, more often the heartache and morbidity that's associated with that. That doesn't mean I'm not a believer, it's just that I think we need to have a very healthy skepticism about current cancer treatments.

00393

I am also coming at it from the perspective of someone who's spent most of my adult life in health policy here in Washington in different capacities, and I'm very familiar with the tension between those advocating something new and wonderful, and the payers who in their responsibility to monitor how we spend our funds, make sure they are done appropriately.

However, I must admit that through the course of the discussion yesterday, I was reminded of a quip I once heard about economists, which was intended to be funny, but the economists were described as people who, when something is proven to work in practice they want to find out if it works in theory. And it seemed to me the preponderance of evidence was that this is in fact a useful tool, it's information, I think in some respects it has been oversold in what it promises, but I look at it more simply as Grant described, it's information. And it clearly would be useful to me and my patients in making decisions about chemotherapy.

I think to put it in context of

reasonable and necessary is wise. It is reasonable. Whether it is necessary, I think must be left to the discretion of those experts who are going to be using it.

And Grant, I just have a quick question for you, because I'm not clear in my mind, maybe I should be at this stage of our panel's life, but I'm not. Let's assume for the sake of discussion that this panel has consensus that it is reasonable to pay for. However, to determine what's necessary is going to, because I know about this having participated in the negotiated rule making, that won't be done willy-nilly as you were talking to Dr. Brooks, that won't happen, because Medicare doesn't willy-nilly pay for anything. It will either be then subject to development of a national coverage policy, or left for medical review. And how do you envision the next step? Will it be an LLMP or will it be, will you ask us, not us, but will you convene another panel or will you ask us to reconvene and develop what we in negotiated rule making did? For those of you who don't know what that was, we did wrestle over several months and developed coverage policies, was it 23 tests or 24?

DR. BAGLEY: It seemed like 124, David.

DR. SUNDWALL: For those of us who participated, it was a 13 month process in total, and it felt like going to the dentist for three days at a time, but we did it, and it was a useful thing. So do you think this kind of testing would be subject to a national medical review policy or would you leave it at the discretion of the carriers?

DR. BAGLEY: Well, given the fact that the whole negotiated rule making process was driven by mandate from Congress, and the goal was that there be national policies for laboratory tests, and even when they weren't national policies they should be uniform across all the regions, I think it would be our goal to in one

fashion or another to try to make the policies as nationally uniform as possible. That means making them as specific and data driven as possible, and that means being as specific as possible in what we're dealing with. And most likely, it's going to require something beyond what this panel will do in terms of mining through the literature. So should this panel make a recommendation that we be, that we go down a path of looking for specific kinds of indications, of drug tumor combinations in which

00396

the evidence is somewhat better, then I think we would undertake the task of trying to mine the literature for that purpose. But I don't see reconvening this panel. But I also think our job would be to try to find a consistent policy which could be applied nationally, whether it be regionally uniform or national, I think we would try to have some uniformity.

DR. SUNDWALL: Because I think that would calm a lot of the concerns that I have heard about today, the idea that if some of us say that it's okay, then it's just universally available. I think there needs to be guidance and expertise on how it's appropriately applied.

The last comment I would just make is, I think from a health policy standpoint, we keep talking about quality access and costs, and access I think is important, and this is a useful technology and ought to be accessible to Medicare beneficiaries appropriately applied. I do think the cost issue is promising. I was disturbed by the testimony of the American Cancer Society clinical oncologist yesterday. It seemed very self serving. We're neutral, but by the way, cover chemotherapy, and don't you dare tell us

00397

what we can't use. And I really thought that was not constructive, and I think it's troublesome.

And I mean, you must hear this all the time. A typical provider tells HCFA, I don't want much, I just want more. But I thought the

oncologists were not constructive, and I'm perplexed why they were not more supportive of this when in fact they were during negotiated rule making, and I assume this young man was speaking on behalf of the society. But I do think in the event this is approved, I think one of the questions you posed before us merits discussion and hopefully we'll have it through all of the tests we talked about that, and that is, no, I don't think HCFA should pay for things that don't work.

DR. FERGUSON: Okay. Miss Simmers?

MS. SIMMERS: I think this fits into the category of the tools that a physician should use in order to dose, or what course of treatment to endeavor and to go in that dialogue between the patient and the physician and the family.

I think the one presenter and one panel member pointed out something that I think about all the time, at least in one region, and I think

00398

it unfair to have Medicare beneficiaries pay their insurance premium and not get access to that test. So I'm very much the 35,000 foot view of the Medicare beneficiaries' needs and how well what we decide to recommend would serve those beneficiaries. And I think Dr. Bagley's point about they help pay for it is very valid. They do help pay for it, and as insurance carriers are paid, I think the same level of access should be available to Medicare beneficiaries that there are provided for those who have private insurance.

I have some concerns, however. Certainly the accessibility of the test, although I am convinced there are ways to address that, I think there would be a duty to promote this as, promote its accessibility not only in very small regional areas, but nationwide, should we decide to cover it. I think there are some valid research and scientific questions to be answered. I do believe that this is a diagnostic test and should be held to that criteria, and not

to that of therapy. I think that was clear, and although I think the science could be better, it's not a perfect world, and what we see is at least compelling evidence to continue.

The policy specific issues, certainly I am not the expert in that area, but is does seem to be available to us to look at those cancers and what circumstances with what drugs are best pursued under this coverage recommendation, should we make it. It seems to me the processes are in place to do that, so I don't fear that it can't be done. And certainly, the devil will be in the details of that kind of process.

But after everyone has said their piece, I guess if any of us in this room were to face the terrible news from our clinician tomorrow that we had cancer, who among us would say no, this assay is not for me. And I believe if it's for us, it is for Medicare beneficiaries as well. So that's my comment.

DR. FERGUSON: Thank you. Before asking some more of the panel, I have one clarification, Grant, or anybody else. As I understand it, our job is really to answer these questions, which can be posed in the form of motions. For us to add some more questions like should Medicare cover this, yes or no, for our panel, or under what circumstances for our panel, I did not think was our job. Has that been changed?

DR. BAGLEY: No. We spent a long time with these questions as staff. We spent a great deal of time toiling over them. And the reason is that the answers to these questions we think are particularly relevant in giving us some guidance in developing a policy. So I think we are particularly interested, and the number one goal should be, you know, to go through those questions and give us not only answers, and they aren't really yes or no questions necessarily, but to give us some discussion and to give us

some rationale. And that's one of the biggest reasons for having this transcribed word for word, so that discussion around those questions can, we can use as guidance in trying to develop some policy.

So I think once these questions have been dealt with, and we can have the discussion and the committee's thinking around these questions, the reason for number six is that the committee can then entertain some other issues, and I think we can hear the other concerns from the committee or from individual committee

00401

members. But I think at this point, it is particularly important and relevant to deal with the questions presented, because we felt those were necessary for us to have the direction for policy.

DR. FERGUSON: I was going to ask that we handle question five last and put six before it, because it seemed that the fifth question was sort of the bottom line question. Is that --

DR. BAGLEY: That's not inappropriate, as long as, before lunch time, we deal with question five.

MR. MINTZ: I have a question. Grant's comments just prompted a question. I had the privilege of serving on the myeloma panel, and we went beyond the questions certainly to the very direct question of coverage, you know, by virtue of motions that were made during the course of discussion. I'm interested in your comments about that, and whether you felt the panel sort of exceeded its charge in that regard, or whether that was just something that happened in due course and was appropriate.

DR. BAGLEY: Our lesson learned from the myeloma panel and again, the reason for

00402

toiling over those questions as we did, our lesson learned was, number one, the committee discussion was far too short. We had lots of presentations from the outside, the committee

didn't have much discussion, so when we took it all home to try to make sense of it, we thought, I wonder what this committee member was thinking, I wonder what they were thinking. So that was number one lesson; we wanted to know why the committee is making the decisions they are, not just do this, do that.

Lesson number two is, it's very easy for the committed to sit here and say, and we've heard a sense of that this morning, you know, this sounds pretty reasonable and I think I would want it, so other people ought to have it available. On the other hand, it needs to be fit, we need to look at the science, the policy has to be restricted and figure out where it fits, and maybe gather some data, so we think HCFA ought to pay for it, but only when reasonable and necessary, and they ought to gather some data. We knew that coming in, so our sense is that that's the policy direction we need to go, and we need to hear from the committee.

00403

Not that the committee can say, cover it for this ICD-9 code and this one, but we need some policy directions on what we need to look for.

So, the lesson learned from myeloma is -- and that's why we fashioned the questions, so I don't think it's totally inappropriate for the committee to make some expressions, but the myeloma panel left us to cover this, but I really think it's necessary, and we did have good people, and the right kind of people and figured it out. That's not being that critical, but you know what I mean.

DR. MINTZ: But it did direct, it did have a specific vote on coverage. It went through cytogenetics and it went through each of the clinical conditions and voted 11 times, I think, to cover under a variety of circumstances.

DR. BAGLEY: That's very true, and we're cognizant of that. But I think our goal in fashioning these questions, at least as far as threshold issues, were to give us the information

that we need for policy.
DR. MINTZ: So it is a different direction in that sense?
DR. BAGLEY: Yeah, I think it is. And

00404

I think it also makes it clear that HCFA, after all, cannot abrogate the coverage decision making to the panel. The panel is advisory and the panel needs to give us advice so that we can make the coverage decisions, and simply having battling coverage decisions doesn't help the information we need to make the decisions.
MR. MINTZ: Just for this panel's information, the myeloma panel voted to cover everything but refractory relapse, and went through a whole series of votes to do that. So it is very different from what we experienced.
DR. BROOKS: Dr. Bagley, I guess your response to one of the other questions and whether or not, you know, clinical trials are over when Medicare decides to make a decision is like very very crucial to me, because we have data on certain areas that I think we all agree at least has clinical utility. But what we actually don't have in each individual, patient by patient, whether they got the sensitive drugs or not in all those studies. And so, I still do look upon it as a potential willy-nilly thing unless we collect data as we go, and maybe reevaluate in a few years as to what -- because

00405

you know, we have that 20 percent problem there, and that was brought up in the negative predictive value.
So, I would really like to know whether or not in a coverage decision, you can mandate a trial that is, you know, it may not be the usual clinical trial that everyone thinks about, but at least, that you can say, I'm going to have the NCI or the CDC or somebody, collect data as we go here, that if people are not using the test to give the sensitive drugs after a few years, then what's the sense of having the test? Or if

everybody's just ignoring the results and so forth, or in fact, if when they get a negative result and people went ahead to treat with the negative drugs and they responded 25 or 30 percent of the time, which is about the 30 percent of the time they respond to almost anything, so I would like to really know that. And so, what's your answer to that question as to whether you can require in a coverage decision that at least ongoing, much more than your claim? I mean, I don't believe you can evaluate this by your claim data, you know, who got what.

DR. BAGLEY: Well, it's -- it would be

00406

long and time consuming for us to try to tie a coverage decision to what NCI was doing, and it would probably, we would be here years from now talking about the same issue. That's one possibility. But the other one is that normally if it requires a clinical trial, we don't cover it, and if it's gone beyond that and then diffuses, we do cover it. Now as I mentioned, there are ways in which we can cover things and we can gather information, and by law we can only gather what's necessary to process the claims. But it's surprising, the amount of clinical information that that makes available to us. It is a fairly rich data resource.

And I mentioned, for example, PET scans for single pulmonary nodules. Now what that policy says is that, it says that the purpose of this is as an alternative to open biopsy, and that we would not expect to see open biopsies after single pulmonary nodules. That's what it says in the narrative of the policy. Does that mean that we would never pay for an open biopsy after a negative PET scan for a single pulmonary nodule? No, of course not, because there are going to be clinical circumstances in which it's

00407

necessary. But it does say, and it's an expression, and what we did, the way we described it is that we said, you know, when you approach a

PET scan for a single pulmonary nodule, you have reached a fork in the road.

And we would look at sensitive testing, or resistance testing, as a fork in the road. We have a test. We're going to do a test, and that test is a fork in the road for us, and depending on the result, we're going to go down path A or path B. Now, I think that a physician and a patient should have a full discussion of that fork in the road before they go down a path. And to do the test and say, well, this test was a fork in the road but don't like the answer, let's go the other direction, makes you wonder whether the test was necessary or whether the direction is necessary.

And so what we would do, hypothetically for example, is that the policy would say that we would ordinarily consider it not medically necessary, or not reasonable and necessary to use chemotherapeutic agents shown to be resistant or not sensitive by this testing system. That means that an oncologist ordering the test, and then

00408

embarking on a course of therapy which was not shown to be beneficial by the testing would be very likely to having a carrier medical director say, I think these claims are not reasonable and necessary, I'm going to deny them. And then they could go practice their medicine in front of an administrative law judge.

And I think that would not be an unreasonable kind of policy. That would allow us to gather some kinds of information, and we would see where it goes. But to say we are going to put a clinical trial together as part of our coverage, I mean when we start paying for it, the ability to collect information, and the impetus out there, and the stimulus for people to do clinical trials goes away. And we've seen it time and time again. And bone marrow transplant for bone cancer, I think is a perfect example. When payments started, the clinical trials stopped. And it was only after many many years

of accumulated data that we then found out that this was not a wise course.

DR. FERGUSON: Dr. Hausner?

DR. HAUSNER: This is a question for Dr. Bagley, so that I can understand the process. The test itself would be relatively expensive, say compared to a CBC, in other words, whatever the price is, and I know we're not talking about price, but let's establish that. A concern that I have, and this is, I practice pathology in a world probably quite similar to the world that Dr. Fischer practices his family practice, and the question of proper utilization is very much on my mind. Would the decision to pay for the test eventually reflect itself in a hospital DRG? In other words, would the test be so expensive that the DRG for say CLL have to be modified in order to incorporate the test?

The second question that I have, and it's a concern particularly for CLL, which is a malignancy that sometimes isn't even really treated in an older patient necessarily, would be how to avoid this becoming a standard beginning of disease test that might or might not be used? Is there any way to monitor, and I know you can't say, well, by God, if you do this test, you'd better get some chemotherapy. I am not quite advocating that. But you know, do you understand the flavor of what I'm asking? Because from my point of view, I want to do the right thing for the patients, but at the same time, I don't want to inadvertently give a group the key to the treasury.

DR. BAGLEY: Well, nothing in Medicare is simple. Let me start out by saying that. It's always complicated. And the one, well, one of my adages in Medicare is that however complicated it seems to be, it will be more complicated by the time you get to the end. So how this would be paid for and in what fashion is, I think is something for the future to

decide. Obviously a patient in the hospital who undergoes surgery or biopsy and the tissue is harvested in the hospital, then the test itself is going to become part of the DRG. Now if that means that the cost of many malignancies is going to go up, there is a way, and the mechanism is that the DRG over time can reflect increased costs. But it's not rapid, it's a slow process. Would all of the costs of this be part of the DRG? No, it wouldn't be, because some things are not included in the DRG. You know, physician services are not included in the DRG. So to the extent that some of these are physician services, pathology services, consultation service, you

00411

know, and a certain portion of these tests are evaluation and billed as consultation services, not as laboratory services, then they wouldn't be in the DRG, but they would be in the part B physician fee schedule payable out of part B, and they would become simply payments from that mechanism. So you know, it's going to be fairly complicated, but I think it's -- you bring up the point that the inpatient part of this is going to be in the DRG, and it's going to impact on how hospitals choose to have oncology patients taken care of and their admission status.

DR. FERGUSON: Kathy, and then Dr. Fischer.

DR. HELZLSOUER: This is Kathy Helzlsouer. I just want to raise for discussion among the panel how we are going to define clinical utility, because I think that's the crux of it and really the first question brings this out. If utility is going to define as, is this a marker of response rates, I think that's what the literature has been designed to show. I think the appeal that Dr. Fischer did, that gee, if I could get oncologists to stop treating when it's appropriate and not have all the side effects,

00412

that's great, but we don't have any evidence that this would change clinical practice, that those

oncologists that you don't refer to because they say never never never quit, they are probably still going to say never never never quit, they may not even do the tests and if they do it, they probably won't always use the results. And we don't have any evidence that it's -- would it change practice or should it, because of this issue of false positives, false negatives. And negative predictive values, we learned yesterday, is highly dependent on the prevalence of the population under study. So we have to be careful of the literature that we're looking at in using negative predictive value or positive predictive value to guide us.

So I think, like I said, it seems reasonable to me, it gives them added information. If that's the criteria, then that's fine. But if we really want more than that, and we expect this to change practice, I guess I'm not convinced that it would or should. And we don't have -- the issue of the added value, I think comes up, because in the literature that we have here, with the exception of the CLL study,

00413

which I am just looking through this morning, is what is the added benefit beyond what we already know, which is what Dr. Brooks is raising. Do we have all the other markers that are now being used and accepted, not just histology, but other prognostic markers really define who's going to ultimately respond and not respond. Does this add something to that information or doesn't it? We don't have from the studies that have been done to date anything to say what the added benefit of this test truly is, with the exception of perhaps CLL.

So I agree with you. I think that as I said, a huge benefit would be if you could eliminate toxicity from agents that would not be used, and that may be the case, at least the first line. But you're talking about a specific situation too, which is the other overriding issue. You're talking about metastatic end stage

disease when you were jumping to chemotherapy. So that's what I'm struggling with. I think if the diagnostic tests say it's not in the patient, there's obviously no harm. The only harm would be, though, it it's used to guide therapy and there's a significant rate of people that would

00414

not get therapy that would be potentially beneficial, with 20 percent response rates in the face of a test that says resistant. So if the policy is tied to eliminating most chemotherapy, I think there is the potential for harm here in how the tests would be used.

DR. FERGUSON: Dr. Fischer?

DR. FISCHER: Yeah. We talked yesterday about where the bar should be set for the level of evidence. I would hope that the bar wouldn't be set at the level where the people who were coming up with this test have to demonstrate that somehow oncologist's behavior has changed for the better. That's much higher than we need.

DR. HELZLSOUER: That's your argument for using the test, though.

DR. FISCHER: But to borrow Greenspan's term, I'm not worried about irrational exuberance with this test. I mean, the fact is, most people don't believe it, and I think that it's a much smaller job for the champions of this technology to convince us than it is to convince all the oncologists of the world. And I think that, you know, that will happen. The evidence that we saw from '99 that has been published is a lot better

00415

than was published in 97, and it looks like we are on to some good things here.

This is very different than orthopedists and MRIs, you know, where you have an orthopedist and an MRI, you've got somebody getting a test. This is not something that's going to overwhelm the oncology community overnight because most people don't believe it. And I think the champions are going to be expected to publish data that will be disease and

drug specific, and to the extent that evidence is persuasive, people will change their behaviors. So I don't want Medicare doing any studies, to be quite honest with you, and I wouldn't trust them to do it. I think that the scientific community is going to answer these questions over time. The question is whether Medicare should pay for it given what we know about it at this point in time and you know, I think the date is pretty persuasive. I am hopeful that that will change the practice of oncology, because I see a very different part of that practice than what the average oncologist does. And they're the patients and their family who have failed chemotherapy with very terrible results.

And I just want to repeat this one more time. In the total practice, in my total practice, there has been as much harm as benefit from oncology.

DR. FERGUSON: Yes?

DR. BROOKS: Well, I don't personally believe that, because over the course of 25 years, as has been published in a number of publications, the survival rate for cancer is going up, precisely because of all the clinical trials and so forth, so I just do not believe the American public is being harmed by oncology.

DR. FISCHER: I didn't mean oncology, I meant chemotherapy, excuse me. Oncology is different.

DR. BROOKS: But it is chemotherapy.

DR. FISCHER: No, no. It's a lot of things, including detection of cancers earlier. But in terms of somebody now has to get chemotherapy and you're looking at their benefits versus the cost, it's on balance a very hard call to me, to say that my patient practice has benefitted more than they have been harmed by chemotherapy.

DR. FERGUSON: All right. Miss Kraft, and then we're going to have a break.

MS. KRAFT: I'd just like to echo Dr. Fischer's comments. I mean, the statistics prove that only one in every four patients actually respond from treatment of cancer. And I'd also like to comment on Dr. Bagley's comment that bone marrow treatment for breast cancer patients, deciding to pay for that was a bad decision. Well, I don't necessarily look at it that way. I mean, it was determined that it would be funded, and did not the testing and all that was done following the approval of that show the American public that yes, bone marrow transplants shouldn't be used in breast cancer?

And then to comment with Dr. Brooks's comment about if this is approved, will clinical trials stop and no testing will be done, again, I echo Dr. Fischer's comment with the fact that I don't think there is any laboratory test out there that has been approved and been funded by Medicare where as soon as it's approved and funded that the testing in that laboratory, test in any trial has stopped in modern medicine, for the last 20 years. I think after a test is funded by Medicare, that the trials and -- not

00418

necessarily the trials, shall we say, but the testing done on patients increases 10 to 100-fold. So I think after approval the test will probably find out a lot more information than we have so far here today.

DR. FERGUSON: Thank you. We will convene back here at 10:15.

(Recess taken from 10:00 a.m. until 10:20 a.m.)

DR. FERGUSON: Can we come to order? Dr. Bagley has one point of order. Dr. Bagley?

DR. BAGLEY: Now we're getting to the end of this, and I've given you an explanation of why we presented the questions. And we need to go through these questions, and Dr. Ferguson is going to go through them in an orderly process and we're going to get through them by noon. People have travel plans and we are going to be

out of here at noon. So we have to keep to a schedule.

Before we start voting on these questions, I want to clarify one issue, or at least one thing that might be a misperception. The new process, and this is only the second time we have done an advisory committee, but the new

process which we put together, which I think is a real step forward, is a process in which we do these things in the open, we do them in a way that we get input from all parties, and that's why the committee is supposed to represent all parties, and that's why we've got the opportunity for public dialogue here. But it is a new process, and I think we are feeling our way along, and it's one in which I think the medical profession itself is trying to find their proper role in where they ought to be here, and what they ought to be saying.

And this advisory committee process is not for every issue. I mean, if we had every issue before a committee like this, we would only handle 10, 11 or 12 issues a year, and we would not adopt many new things for Medicare. So it's only for selected issues. Now, how do we select them? Well, we've announced our reasoning is if it has a high impact for the Medicare program, an important advance, something that affects lots of Medicare beneficiaries, something that is very important for Medicare and its beneficiaries. If it's something that is controversial in the scientific community. If it is something in

which there is a varying coverage, which appears to be quite disparate without good explanation across the country. Well, I think you've heard in a day and a half that this issue clearly meets all three of those criteria. We though it did, and that's why we presented it to this committee.

But think on those for a minute. Significant controversy within the scientific

community. We heard that and we've tried to bring that to the fore and present it. And I -- the reason for this is I want to explain the fact that I don't think you should read anything into the fact that we made a staff presentation on this issue, in which we pointed out what we thought were many problems with the literature and many gaps. That shouldn't be looked at as prejudicial. That shouldn't be looked at as the fact that we have a position or we have anything else. It was in terms of balance, because as you've noted, there are many proponents of this procedure who are pointing out the favorable aspects of the literature and the promise of this procedure, and we are relying on them to make that presentation. But we think in terms of

00421

balance, we wouldn't have brought it here if it was one sided, if it was easy. And we want the committee to hear both sides. And if we have presented, at least been the devil's advocate in saying there are some problems in this literature, this is an early technology, this is controversy in the medical community, and if we as a staff, or if I have as the HCFA member here, appeared to point out problems, if that appeared to be negative, I want to correct that misperception.

If we had our mind made up, we wouldn't need to come to this panel. If we knew the answers, we wouldn't need to come to this panel. If they were easy questions, we wouldn't need to come to this panel. So when it's been necessary, we have been the devil's advocate and brought up the contrary arguments, because there has been adequate representation of the pros, we've tried to bring up the cons, only for the purpose of presenting this to the panel and having presented it to the panel and having made it clear that we don't have our minds made up, that we are after all neutral, and waiting the panel's deliberations also. So having made that

00422

hopefully clear, that there should be no bias at this point, let the deliberations begin.

DR. FERGUSON: Thank you. Before we start, are there any other points that the panel would like to make before we start on these?

DR. LOY: I would just like to point out that this discussion has been compared to many other diagnostic tests and I really don't, at least in my mind, I don't see this as a diagnostic test. The diagnosis has already been made. Clinically probably has already been confirmed via other laboratory tests or other modalities. I see this as a test which gives probability into the likely outcome in terms of response from a specific therapeutic modality, so I think it's a little different from a diagnostic test, and I don't think it should be compared to other diagnostic tests.

DR. FERGUSON: Thank you. Yes, Dr. Brooks?

DR. BROOKS: Yes. I guess I want to ask my former question a different way of Dr. Bagley, which is to say that, can Medicare in its rule making or whatever coverage decision it's called, put a requirement not to things in

00423

general, but if there are to be new clinical trials of new agents, new agents now, or a new agent combined with various other old agents on various clinical arms by the government supported NCI system with the C Comps, everything else, that those new trials employ this new technology, so that data at least may be gathered by that process?

DR. BAGLEY: Well, real quickly, no, we can't tell NCI what to do, we can't tell oncology groups what to do, and we can't dictate clinical trials.

On the other hand, oncology is a unique area in which most of oncology is a clinical trial. Most of the drugs we pay for are investigational. In fact, if it becomes too settled, then the next generation is

investigational, and that becomes standard practice in oncology. So oncology is a little unique that way. So much of what we do pay for in standard daily practice is in fact a Phase III clinical trial, a protocol, or even a Phase II clinical trial in some cases. So, there will likely be in the future, you know, should the promise of this technology hold true, there will

00424

likely be in the future an approach by Medicare for some hypothetical new chemotherapy agent which is very expensive, very controversial, and can be shown to be associated with a high predictive value. There is hypothetically a Medicare coverage policy in the wings which will say we are not going to pay for this until it's shown to be effective. Maybe this isn't the test that's going to do it, but I think the day will be there when we will select chemotherapy in a more rational way and that will happen. So yes, I mean, that approach can be put into policy, but not directly in the way of we're directing clinical trials, but when we are paying for the drug, we can say this is how we'll pay for it.

DR. FERGUSON: All right. Yes, Dr. Loy?

DR. LOY: I'd like to make one other comment and that is, in answering these questions that we have been presented with, I would hope again that the panel would try to be very specific about these. I feel like we're, again, flying around 35,000 feet, and I think there is a real need for trying to hone in on which specific clinical indications, when, why and how often

00425

these tests should be performed, and that should be part of these answers to these questions, in my opinion.

DR. FERGUSON: Well, I'm not sure we can do that, but I don't say that it's not important. That may be the crux. However, we will do what the possible is, and see. Now what we're going to do is the following: I have asked

several members of the committee to propose a motion, these questions, individual questions as a form of a motion, and then we will have somebody second that, whoever chooses, and then we will discuss these individual things for about eight minutes or so, and then we will have a vote on that motion. And it will be a simultaneous vote of this panel, not a sequential vote.

And the voting members, you may have on your list, but the chair is not a voting member unless there is a tie. Interestingly, we have 11 voting members, so I'll see. Supreme Court.

Now, Dr. Murray has volunteered to make a motion for the first question.

Motions, Discussions and Recommendations

DR. MURRAY: I guess that's my reward for saying that we have to get to the questions

00426

in my comments earlier. So with regard to the first question, the issue of scientific evidence, it's my belief that scientific knowledge is best advanced in a free and open arena, not one that comes with administrative preconditions, and accordingly:

I move that the advisory committee recommend that the clinical response as well as survival rates be accepted as appropriate measures of clinical utility.

Just repeating that, I move that the advisory committee recommend in favor of clinical response, not to rule out, but that clinical response not further defined be acceptable as an appropriate measure of clinical utility.

DR. MINTZ: I'll second that.

DR. FERGUSON: Moved and seconded. Now, is there some discussion on this point? Dr. Klee.

DR. KLEE: Does the term clinical response include quality of life, or is that inferred or explicitly --

DR. HELZLSOUER: May I make a comment? This is Kathy Helzlsouer. I don't think it should include quality of life, because we have

00427
absolutely no data to judge quality of life issues here. So I think it's clinical response is defined responders, non-responders, as we have in the data. That would be my opinion.

DR. MURRAY: I did not intend to include quality of life as a clinical response. However, it goes without saying that many measurable clinical criteria affect quality of life. So I think it's a bit semantic, but to answer your question, no, quality of life as a specific criterion is not intended to be included.

DR. FERGUSON: Is there further discussion on this?

DR. BROOKS: So Dr. Murray, I would gather then that your motion would be to say both to the second part of that, what outcome measures should we rely on to best assess clinical utility, both clinical response and survival rates?

DR. MURRAY: That's correct. That would be an equivalent way of stating the motion.

DR. BAGLEY: What we're talking about is trying to be very procedural, and as we go

00428
through each one of these questions, I think some of them will be quicker, some will take a little longer. We want to afford an opportunity if there's a relevant public comment somebody needs to make, whether we should do all the questions first or go each question, and I think what we're going to do, if they can be very brief, but again, if they're relevant, take a public comment before we vote on each question, if it's necessary.

DR. FERGUSON: Is there further discussion from the panel regarding this motion? All right. Is there a comment or some point from the floor regarding this motion?

We have our protocol lady again coming to tell us. My script is missing.

Would anybody like to make any lasting remarks? This is for posterity. Speak now or forever hold your piece. Okay.

Can I have the recorder read the question, or read the motion back?

THE REPORTER: It will take me a second to find it, but yes.

DR. FERGUSON: I thought I was trying to be very precise about this. If it's going to

00429

take a while, maybe you could just read it again.

DR. MURRAY: The motion is that the advisory committee recommend that clinical response as well as survival rates be accepted as appropriate measures of clinical utility.

DR. FERGUSON: Okay. All in favor, raise your right hand. Okay.

All opposed?

For the record, it was unanimous.

Do I have to read their names? Okay. So, unanimously approved by Robert Murray, David Sundwall, George Klee, Paul Mintz, Richard Hausner, Mary Kass, Cheryl Kraft, Neysa Simmers, John Brooks, Paul Fischer, and Kathy Helzlsouer.

Now, we will move on to question number two. I have asked Dr. Kass to make a motion regarding question number two.

DR. KASS: I'm not certain how to put this in the form of a motion. But I would submit that evidence was presented yesterday showing that tests have been done with combinations of drugs, and continues to be done, so I think that evidence was presented supporting this. I'm not quite sure how to state that in the form of a motion.

00430

I suppose I could move that that has been presented to this panel.

DR. LOY: Second.

DR. FERGUSON: So it has been moved and seconded, that evidence was presented to the panel supporting these tests with combinations of drugs. Is that it?

DR. KASS: Yes, sir.
DR. FERGUSON: Is there discussion on this point?
DR. BROOKS: Yeah, I think it depends what we're talking about, because the question reads: The assay described in the literature test responses to single drugs. What is the evidence reporting test results in combination chemotherapies? Combination chemotherapies were tested in the tests, but I think the question gets at whether there were responses to combinations. Is that correct, or maybe I am misreading that, because certainly I agree, there were.
DR. KASS: I was reading that as test responses.
DR. FERGUSON: Dr. Sundwall?
DR. SUNDWALL: Just a comment. I'm not

00431

sure it's useful for us to revisit this, but I did hear yesterday that these tests are customarily done on single agents, and that provides the utility to the oncologist for the decision making. However, it is possible to do the combination treatments and they are sometimes done but were not considered that useful. And I'm not sure how -- it's not for me to decide, or possibly this panel, but I think we ought to be clear if we make this a consensus of the committee that in fact it's customarily done on a single chemotherapeutic agent, and if any of our experts want to clarify that, they can.
DR. FERGUSON: Well, I asked this question specifically yesterday, because it seemed to me that most of the stuff we were presented was with single agent testing. And I was told by a number of the presenters that in fact routinely they do test two agents together on a single test, and that that is, although I don't know how much of the publications indicated that, but apparently some did.
So, is there further discussion about this?

Is there a public comment on it?

Okay. Dr. Weisenthal, and Dr. Nagourney too. Briefly.

DR. WEISENTHAL: Yes. This is Larry Weisenthal. The data that you heard yesterday, there were three separate types of data which dealt with that issue.

Firstly, the majority of the correlations published in the literature of the 2,000 some correlation targets, actually if you take all technologies, 4,000 some, 2,000 cell death and 2,000 cell proliferation, of all those reported in the literature, the majority of those patients in fact were treated with combination chemotherapy.

DR. FERGUSON: We are talking about -- this is referring to testing the drugs.

DR. WEISENTHAL: Yes, in 60 seconds. The majority of those patients were treated with combination chemotherapy. In some of the studies, the activity of the best single agent was used to predict for the activity of the combination, and that works very well, and the biologic explanation for that is that most chemotherapy is additive and not synergistic, so normally and frequently, it's the activity of the

most active agent that determines how the combination is going to do.

Many of the published correlations, however, were actually combinations were tested, and patient response to combination therapy was compared with the testing in combination.

Thirdly, it's very important in only a few situations, but in some situations it's very important to test the drugs in combination. A good example is Mr. Stein, who testified so eloquently. His tests showed resistance to all of the single agents but a unique synergy in that one combination. And also the patient that Dr. Nalick presented with the failure of the bone marrow transplant, she was resistant to all

single agents, but to one unique combination.
So in special unique cases where we actually get synergy, it makes sense to do that and it is done.
DR. FERGUSON: Thank you. Dr. Nagourney?
DR. NAGOURNEY: Basically my comment is just to second that very issue, that there are profound synergies in some combinations, some of which can actually rescue patients from failure

00434

of both single agents. I think that biological validity is forthcoming in publication. So I think that combination sometimes can be uniquely interactive and do have biologic validity in the test tube that directly applies to clinical outcomes.
DR. FERGUSON: Thank you. Mitch, did you want to say something?
DR. BURKEN: Yes. Dr. Mitch Burken. I'd just like to clarify a couple of things from my presentation yesterday that relates to this question. The issue that I think was intended by this wording was what is available in the literature to show the applicability of single agent testing to combination regimens, as opposed to what may be going on in the lab under special cases. So I think this relates specifically to the articles in the bibliography and those presented to the panel about making inferences from single drug to combination therapy, making that jump.
DR. FERGUSON: Other comments from the audience? From the panel? Do you want to read this?
(Portion of record read.)

00435

DR. FERGUSON: Okay. Call for the vote. All in favor of this motion? Okay. It looks again unanimous, and this vote was supported by -- do I read these names again? I don't have to? I only have to read them once; if it's unanimous, I only read them once. Sorry.

I'm on a learning curve here too.

Now, we move on to question number three, and I've asked Dr. Fischer to make a motion regarding this question.

DR. FISCHER: This gets to the issue of the 64 bins that we talked about yesterday. In particular, which tumors have been shown to be useful to test, and I'd like to change the question to the following position, that I move that the panel accept:

In considering the presented evidence, the advisory panel believes that HTASs demonstrate a clear clinical benefit for directing treatment of CLL, and promise for other solid and hematologic tumors.

DR. FERGUSON: Okay. Is there a second?

DR. SUNDWALL: Second.

DR. FERGUSON: Dr. Sundwall seconds.

00436

Is there some discussion? You're saying --

DR. MINTZ: I would be more comfortable with the motion if the word clear was removed, and simply said demonstrate clinical benefit, only because of the uncertainties.

DR. FISCHER: I'd accept that.

DR. MINTZ: So would you amend the motion to remove clear?

DR. FISCHER: Yes.

DR. FERGUSON: Okay. Kathy?

DR. HELZLSOUER: I guess I'm -- this is Kathy Helzlsouer -- concerned a little bit about actually the term clinical benefit, how we define that. I'm very comfortable with these that are markers of response, and can be used clinically, but I think even in this trial, I haven't had time to look at the one article that we gave thoroughly, but to state that this has clear clinical benefit is a pretty strong statement. I think it shows clear benefit in directing some therapy there, I might be much more comfortable with that as amended as such.

DR. SUNDWALL: Suggestion. For

consistency's sake, I wonder if we should substitute clinical utility as used above, with

00437

that definition.

DR. HELZLSOUER: I like that, and would be much more comfortable.

DR. FISCHER: I'd accept that.

DR. FERGUSON: Okay so instead of clear clinical benefit, clinical utility. Take off clear. Okay? Dr. Fischer, can you restate that please?

DR. FISCHER: Yes. In considering the presented evidence, the advisory panel believes that HTASs demonstrate clinical utility for directing treatment of CLL, and promise for other solid and hematologic tumors.

DR. HELZLSOUER: I second that motion.

DR. FERGUSON: Okay. Now, is there further discussion on this amended motion? Yes, Dr. Loy?

DR. LOY: I would only bring up the point that has been brought up before, and that is that I don't know that we really defined in the course of disease where that clinical utility might be. There are certainly CLLs that are treated differently and diagnoses that are treated, as compared to the way they are treated at end stage disease or Richter's syndrome.

00438

DR. FERGUSON: Are there discussions around that point, or others?

DR. BROOKS: I just guess, just that there is nothing in the questions that, in other words, the way it's written, and I very much prefer to the way it is on paper, there is nothing about stage of disease, which diseases, et cetera, so I think we just have to consider it as a whole.

DR. FERGUSON: I don't want to upstage HCFA, but I think that to ask the questions in a way that would prescribe specific therapies too tightly would not be what HCFA would want. I mean, in other words, to write a prescription for

the practice of medicine is not what we're trying to do. Am I right?

DR. BAGLEY: Well, we get accused of it all the time, but no, I think as I said, while the questions in some ways are specific and some ways are general, we worked pretty hard on it to give us some direction.

DR. KLEE: It's my sense that you're not going to be able to use tissue type to determine which patient you're going to pay for, given what we know at this point in time, and

00439

that's why this is a little more vague.

DR. MINTZ: Yes, but I think, Grant, you did ask for direction, and I feel that though we have modified the language to say clinical utility, I feel strongly that this motion should be supported. Dr. Bosanquet's article in the British Journal of Hematology, I think speaks to my sense that HCFA should cover this test for CLL, and since you want a sense of the committee, I think that's it.

DR. FISCHER: But not exclude it for other things.

DR. MINTZ: Agreed, yes.

DR. FERGUSON: Is there some -- yes, Dr. Bosanquet, or others from the audience?

DR. BOSANQUET: Some of the panel members have said some nice things about the work that comes out of my laboratory, which is perhaps the smallest laboratory represented here. The motion, and I'm sorry, I am going to inject an added level of slight problem here, but you're talking as a committee about the test, or the methodology, and what I want to bring out is that for hematologics, the extreme drug resistance assays are not the -- some of the incorporation

00440

assays are not relevant for hematologics, it is not anything used to test the hematologic neoplasms. The DiSC assay and the MTT assay are very similar procedures, and are relevant to the hematologics. So I would caution you to say that

hematologics, or CLL in particular, can be tested by this procedure, because there are actually two different procedures. One is the drug resistance assay, much of which has been presented to you. The other is a drug sensitivity assay. Because hematologics have a higher probability response and you can't do a sensitivity assay there, I think the added value of a drug sensitivity assay is also higher, because you are no longer only excluding the bad drugs, you're proposing good drugs. Randy Stein, I would put to you, wouldn't have been here on a drug resistance assay.

DR. FERGUSON: Thank you. I think my previous comment might hold, that we're not going to get that specific. Okay, Dr. Fruehauf.

DR. FRUEHAUF: Just trying to understand the motion, in terms of how narrow or broad it tends to be, but I wanted to emphasize from my own experience in the field that resistance is important, and that's not to

00441

minimize the value of sensitivity, but if sensitivity is the only word used in the motion, I think that detracts from where actually half of the data are in the field.

DR. FISCHER: Yeah, but there is nothing in the motion about either of those. I mean, the term that I thought I had used was HTAs, assuming that that incorporated all the assays. Is that not true?

DR. FRUEHAUF: Well, yes. Okay. So when you say directing therapy, because of my long-standing experience in the field, you think of that as meaning the selection of drugs rather than the deselection of drugs, and so I'm just asking for a clarification, that what you're saying is these assays are applicable to solid tumors and hematologics.

DR. FISCHER: Yes.

DR. FRUEHAUF: And you wanted to emphasize that --

DR. FISCHER: The evidence is very good for CLL.

DR. FRUEHAUF: The evidence is very good for CLL.
DR. FISCHER: That's precisely my point.
DR. FRUEHAUF: Resistance, solid tumors; I just wanted to ask for clarification.
DR. FERGUSON: I think the language can stand where it does.
DR. BAGLEY: The fact that we are considering these methodologies together, and while there may be differences between them, there may be different approaches, we chose to put them together, because we didn't want to use this panel process for single products. And I'm also mindful of the fact that previously extreme drug resistance testing was considered at the negotiated rule making for laboratories and Dr. Weisenthal showed up and said, wait a minute. Extreme drug resistance testing and drug sensitivity testing are really variants of the same technology. There are some differences, there are some nuances, but we are testing susceptibility of tumor cells to drugs, and whether we approach it from the right side or the left, it's the same technology. And we took that to heart and we're looking at all these technologies together.
DR. FERGUSON: Thank you. Okay. I guess I will call for a vote. All in favor of this motion, raise your right hand. And again, we have a unanimous vote, and I have learned I don't have to repeat people's names.
Now, question number four. I asked Dr. Mintz to makes a motion.
DR. MINTZ: I move that if a human tumor assay result indicates that a neoplasm is resistant to a particular drug, that that should not preclude the use of that drug during the course of treatment for that neoplasm.
DR. HELZLSOUER: Second the motion.
DR. FERGUSON: Is there some discussion

on that point?
DR. BAGLEY: Well, I would start simply by saying that I see this as a guidance, and as a tool, that whether we consider it a diagnostic test or a prognostic test, you know, nothing comes to mind that has 100 percent positive or negative predictive value, and I think this is providing a piece of information to the clinician.
DR. FERGUSON: I think that my sense, I am not supposed to vote and I won't, but that a laboratory test is, I think the clinician's view

00444

of that patient considering all things that they have in their hands has to take precedence over a single test. We've come across that in neurology, on whether to turn off the respirator if somebody's somatosensory evoke responses are negative. They get measured, you know, and people say, well, they're negative, so we can unplug this patient. I believe that's not the way to practice good medicine. At any rate, is there further discussion about this point?
DR. BAGLEY: I don't like to ask the question, because I would like to get some discussion from the panel about where we are headed in this thing, and to say, gee, a test which shows the drug is resistant shouldn't preclude using that drug, and I'm curious as to why people think that the test would have utility, which we voted that it did, if you wouldn't use the result to guide therapy, if you said well, it's just prognostic. And I would think that prognosis ought to guide therapy. And so, I mean I realize that physicians are uniquely able to ignore information they don't like. I mean, that's why the aircraft accident rate among physician pilots is like six times the normal.

00445

DR. FERGUSON: Grant's a pilot, I might add.
DR. BAGLEY: Because they use weather reports the way they use laboratory tests.

DR. HAUSNER: Let me try to take a crack at it. Use the analogy of a patient with a urinary tract infection, that we start an antibody, say Keflex, and we do our sensitivity testing on the organism. Then we find out a day and a half later that the drug, that the organism on the plate is resistant to Keflex. Go back to the patient, back pain has gone away, urine has cleared up, fever is gone, and no one, any insurance company said well, you need to stop that Keflex immediately, because it was resistant. And it's an easy question in the context of say a urinary tract infection, because we're dealing with shorter time frames, less is at stake, less money, and it's something that we don't even think about. I don't think anyone here would vote to stop the use of an antibiotic if it's resistant on the plate and the patient is doing better.

In clinical oncology, which I don't practice, we're dealing with a little different

00446

time frame and more is at stake, so I could envision a situation, though maybe I'm stupid, someone could correct me, in which a patient comes in say, with a big meaty Steinal mass, has some form of lymphatic lymphoma, their trachea is being compressed, they've got to get on some chemo. Meaty stenoscopy is done. Tissue is harvested, chemotherapy is started. And then two weeks later, ta da, the organism, the tumor is not apparently, not sensitive to the chemotherapy. Patient course is breathing better and doing well. What do you do then?

So I think that as illogical as it may seem to people who aren't in the field, there is really no other way to answer that question at the current time than the way the motion has been stated. There just isn't any way that I could walk out of here and do it any other way, as illogical as it seems to appear.

DR. BAGLEY: Well, let me follow up on your example, because it's a good one. You make

a good point about clinical response as opposed to the testing. But let's take your urinary infection patient and you put him on Keflex, or let's say you put him on Bacterin, you do a

00447

culture and sensitivity, and the culture comes back three or four days later and it says this is resistant to Bacterin, and they're doing better, and they're clinically better. I don't think most even managed care plans would ask for the money back for the Bacterin. But on the other hand, say they didn't respond, and they weren't doing better on Bacterin, and you looked at the culture and sensitive, and it said this is sensitive to ampicillin and resistant to Keflex. Now I don't know many managed care plans that would say sure, go ahead and give him Keflex. It's more expensive and it's resistant to it, but go ahead and give it to him anyway. I think that's the flavor of the question, is if we don't have clinical response but we've got a prognostic test, should the test guide therapy to the extent of saying we should or shouldn't treat it? That's the direction the question is going.

DR. FERGUSON: Yes, Dr. Mintz?

DR. MINTZ: I will take it from a different angle briefly. Yes, it should guide therapy. But if I had CLL and I were resistant to fludaribine, don't start me on that. But if it becomes fulminant, and if the other therapy

00448

has not worked, by all means, try it. You know, I think it's a guideline, I think it's a tool, but I don't think that because it can't be expected to have a hundred percent positive or negative predictive value, that it should preclude the use of an agent when other apparently beneficial agents aren't working.

DR. FERGUSON: Dr. Brooks?

DR. BROOKS: Yeah. I think I would follow up on both of them and agree with them. And I think the rational answer to your question is it's not a perfect test. If it was a perfect

test, then I would agree, we should preclude whatever it is, insulin in the example I gave, but since it's not a perfect test, then you shouldn't. And secondly, along the line of Dr. Mintz talking, this may be the last drug in the series. I mean, people have cancer, and they may have seen four or five or six drugs, and now you have a result that says it's resistant to Taxol and yet, as we know, there might be a 20 percent chance this patient may respond to Taxol. I'm going to have trouble as a clinician explaining to that patient that, you know, I can't give that because that result's resistant.

I know there is a 20 percent chance, but I'm not going to give that to you. So, I think for the combination of those two reasons.

DR. BAGLEY: But even to the point of -- I mean, you said it shouldn't preclude therapy, but then you said well, gee, if I had CLL and fludaribine showed resistant, I wouldn't give it but I'd give it eventually if nothing else worked. I mean, in a way, there you're saying that we wouldn't give it as primary therapy, but we might reserve it, and that becomes a qualified answer, and I think that's an important nuance.

DR. BROOKS: Well, that's because it's a qualified test. It's not a perfect test.

DR. BAGLEY: But you're saying that if it showed if it was resistant, you wouldn't use that drug, until you had exhausted the ones that didn't show the same thing. I think that's an important distinction.

DR. FERGUSON: Two more points down at the end, and then we're going to call for a vote.

MS. SIMMERS: I think to clarify it a little bit, what you're trying to do is find the drug that has the most probability of working and using it first, and then move down the line. When you get to 20 percent, you're pretty desperate, but should you exclude it from the

possible clinical modalities to take care of this patient, and I think the answer is clearly no.

DR. FERGUSON: Dr. Sundwall?

DR. SUNDWALL: It's clear to me that this issue is huge, and what we're talking about is futile care, we're talking about explicit rationing, and I really sympathize with HCFA because I understand that with all our best efforts to rationalize payment, everyone wants that caveat, but yeah, in my case, in case I'm dying, pull out all the stops. And as long as we understand that, as much as we'd might like to, Grant, we are not going to resolve these issues, because it would be tantamount to explicit rationing, and I don't think we are prepared to do that.

DR. FERGUSON: Dr. Helzlsouer?

DR. HELZLSOUER: Yeah. I think it should guide but not dictate care. I don't think we can use the test to dictate care, and there would be lots of reasons in addition to the fact that you might have a situation, since 20 percent

00451

would respond even if they were resistant on this assay, according to the literature we have, and that's based on sensitivity response. You could have a situation where somebody, you still have a 20 percent chance, and in combination you might choose a less toxic drug rather than a more toxic to which they are sensitive, because you're using it in combination. There are a variety of scenarios you can come up with that this test alone should not be your sole, to dictate therapy alone, and there has to be a combination of other factors along with this test result.

DR. FERGUSON: Dr. Murray?

DR. MURRAY: I'm a little uncomfortable with the motion as stated, because it seems to undermine the value of these tests, that -- I agree with what has been said, that there are extraordinary circumstances, there are primary failures, when it is appropriate to overrule as it were, the results of the laboratory tests.

But the motion as stated seemed to indicate that the test need not be given any weight, and yet what I hear various panelists stating, yes, it guides therapy, yes, you use it for your first line choice, and I would like to see the motion

00452

amended to reflect that, perhaps with a clause, in the absence of extraordinary circumstances, or unless primary modalities have failed.

DR. HELZLSOUER: Well, I guess my concern with your point is that if we had had the evidence to say there was clinical benefit, which is what we took out of the other one, we would probably be having a little different view, and be more willing to be restrictive perhaps, but we don't have that evidence. We just have evidence that this can mark response to certain therapies and even at that, it's not a perfect test. So in the absence of knowing that it really has a benefit in terms of clinical outcomes, and we don't have the evidence for that, I don't see how we can be more restrictive.

DR. BROOKS: And we had the situation earlier, and yesterday, as to how high to raise the clinical bar to approve a test. Now if we raised it to perfection, then I think there is something to be said for that, but we didn't. And in fact in the prior questions, we don't expect anywhere near perfection. And we have the 20 percent issue for example. So I think as long as that's the bar, that's the results of these

00453

tests, then it's going to be hard to tie somebody's hands. And all it means, I know from Medicare's point of view is one thing, tying the doctor's hands, but what you're doing is tying the patient's choices.

DR. MINTZ: I concur with those comments, and not to repeat them, the motion was really intended to be neutral in that regard. The motion was intended simply to say that it doesn't preclude payment for the use of that test. And I saw it as not a heavy handed motion,

but rather as a very lightweight motion, in that it really puts this in the hands of the clinician, and that was the intent of the motion.

DR. FERGUSON: Okay. Can you read that back?

DR. MINTZ: I can do it.

If a human tumor assay test result indicates that a neoplasm is resistant to a particular drug, that that does not preclude the use of that drug during the course of that treatment for that neoplasm.

DR. SUNDWALL: Before we vote, can we modify that to say may not? That's a little softer, because it would leave an open window,

00454

but give a little more weight to the test.

DR. MINTZ: I accept that. Of course I didn't say shall not.

DR. FERGUSON: So you said -- read it again now.

DR. MINTZ: Okay. I hope I'm reading the same thing. If a human tumor assay test result indicates that a neoplasm is resistant to a particular drug, that this may not preclude the use of that drug during the course of treatment for that neoplasm.

DR. FERGUSON: Okay.

DR. BAGLEY: You know, the questions as they are presented, I think the discussion we had, the discussion I tried to provoke, and I think successfully did, about what does it mean, doesn't preclude is useful. And that's one of the lessons we learned in the multiple myeloma panel. We have to have a rich enough discussion around the questions so -- we aren't bound by any recommendation that, exactly the way it is worded, but the discussion around it, which we will have record of, as we try to interpret what the sense of the committee was, I think is fleshed out. So, by trying to provoke that

00455

little bit of a question about what does it mean to use something resistant, I wanted to stimulate

that discussion so we would have that kind of a rich record so that we could interpret what this vote means, and I think we have done that successfully.

DR. FERGUSON: Looks like you did. It seems to me the easiest question took the longest. I call for the vote. All in favor of this motion?

Oh, I'm sorry; what? Public comment? Thanks. Does the public have a brief comment? Three milliseconds?

UNIDENTIFIED SPEAKER: One brief comment. I'm a gynecological oncologist. Out of about 500 patients that I have taken care of using assays over a period of about 15 and about -- over -- using about 300 patients with these third generation assays, I can say that I've seen two that I recalled, and possibly three patients who have responded to a drug that was read as EDR on the assay, three out of maybe 300 something.

DR. FERGUSON: Thank you. All right. So that was again a unanimous vote, unfortunately taken before the public comment. I guess the

00456

woodshed isn't too far away.

Now, I have asked Dr. Klee to formulate question number five in the form of a motion.

DR. KLEE: Number five. I guess I will just have to take it the way it's written here and make a motion, in that:

I move that the advisory committee recommend that there is not sufficient scientific evidence to demonstrate the clinical utility of HTASs in selecting appropriate cancer chemotherapy.

DR. FERGUSON: Okay. Is there a second to that?

DR. MURRAY: Second.

DR. FERGUSON: All right. It has been moved and seconded that there is not sufficient evidence for these tests. Is there some discussion on that point?

DR. SUNDWALL: I'm surprised. I

thought that the discussion so far would indicate there is sufficient scientific evidence to demonstrate clinical utility in the selection of an appropriate chemotherapeutic agent, and inserting not in there surprises me.

DR. KLEE: The reason I was putting it

00457

that way is that this is a very comprehensive statement and if we look at it in all disease states, we haven't seen data, so there isn't sufficient information in that. If we target it to one specific one, we have already said that up in the earlier ones, where we looked at CLL. So I think as it's stated, I don't think there is sufficient scientific evidence to recommend this across the board.

DR. FERGUSON: So you're in effect saying it's a bit too broad. Yes, Dr. Kass?

DR. KASS: My problem with the motion as stated is that if I'm being confused by it after sitting here for a day and a half and listening to all the discussions, I'm afraid that when the Medicare coverage policy is written that it's going to be confusing to the people in HCFA as to what our intention was. I would like to see a motion that clarified exactly the point that you're trying to make.

DR. HAUSNER: I would like to have a crack at just that. To, if you would consider this as I don't know, an amendment or a revision, adding something to the effect that there is sufficient scientific evidence, et cetera, in

00458

certain cases, and you can add in that in other cases, there have not been. And we can use the example of CLL if you want as the poster malignancy for which perhaps there is, or just leave that out. But rather than -- because what's implicit in your motion is, and I understand what you're saying, you're saying that if we said it just, there is sufficient scientific evidence that demonstrates the clinical utility, et cetera, that that's far too

broad. Right?
DR. KLEE: Yes.
DR. HAUSNER: And so what I'm saying is, your motion is far too broad the other way, it's too much the other way.
DR. KLEE: Right.
DR. HAUSNER: But what you really meant and what you were trying to reflect, which I agree with, is that it is not yet a closed book. But in order to be consistent with everything else that we said, I propose that you revise your motion something along the lines that I said about saying that there is for certain malignancies scientific evidence that demonstrates the clinical utility of HTASs,

00459

something along those lines.
DR. FERGUSON: Kathy, and then Dr. Kass.
DR. HELZLSOUER: This is Kathy Helzlsouer. I think the confusion is that in number three we changed clinical benefit to clinical utility, and so we all think we voted on five, which says clinical utility, which says clinical utility. Since we weren't comfortable with the term clinical benefit, and amended that motions, so it's almost now, five is similar to what we did in three, and maybe we need some clarification from Grant as to if you want something else addressed in this.
DR. FERGUSON: Dr. Kass?
DR. KASS: I agree absolutely with that, and perhaps if someone could read to us what we voted on specifically in number three, I think it would become apparent that it was very clearly stated in that what you're trying to get at.
DR. BROOKS: I think it stated promise, so that if we change five to include promise, I think it would be equivalent to three.
DR. FERGUSON: We said clinical utility

00460

for hematologic cancers and promise for solid

tumors; is that correct?
DR. HELZLSOUER: CLL specifically.
DR. FERGUSON: Did we say CLL specifically? Dr. Fischer?
DR. FISCHER: Yeah. I don't think we're going to add much by doing anything with five. I think we should just drop it. The sentiment in the discussion around this issue was done under three, and I think the semantics are just going to confuse everyone, so I move that we drop five.
DR. FERGUSON: Just a minute. We have a motion on the table, that's been moved and seconded and you know, we have to -- Roger's rules, is it? No, Robert's.
DR. SUNDWALL: The motion wasn't seconded.
DR. FERGUSON: It was seconded. It's been moved and seconded.
DR. HAUSNER: Call the question. And my point would be that if it's defeated. Then we have a clean slate. I think quite honestly that Dr. Fischer's idea about quashing it -- I just want to ask Dr. Bagley, is this written in stone

00461

that we have to do anything with these questions? The answer is no?
DR. BAGLEY: No, they are written in stone, and -- well, soft stone. But I mean, the purpose of these questions was to generate the discussion and to get the sense of the committee around these issues. And I think again, the way three was modified, addresses much of the issue, I think the discussion around it discusses much of the issue, and I sense a reluctance in the committee to take a definitive vote on question number five in a definitely broad or definitely proscriptive form, and if the committee decides to not deal with that issue and not take a vote on that, that is an acceptable alternative.
DR. HAUSNER: I'd like to call the question on the motion.
DR. KLEE: Or can I withdraw my

motion?
DR. BAGLEY: I mean there's no reason, because of it having been made, there is no reason that it has to be put to a definitive vote at this time and put people in an uncomfortable position of voting on something they didn't mean to vote on.

00462

DR. FERGUSON: Just a minute now. The question has been called.
DR. HAUSNER: Well, unless he withdraws.
DR. KLEE: I was just withdrawing the motion.
DR. FERGUSON: Okay. I guess we can do that.
DR. MURRAY: I withdraw my second.
DR. FERGUSON: Okay. The question has been withdrawn. Not even tabled, I guess. Withdrawn.
DR. HAUSNER: To nail it down, may I make a motion that the committee not consider question number five, just to nail it down?
DR. FERGUSON: You can make that motion.
DR. HAUSNER: I make a motion that question number five not be considered by the committee at this time.
DR. HELZLSOUER: Well, we already did consider it actually. We considered it in number three,.
DR. FERGUSON: Well, I mean, do I -- has it been seconded? Is there a second to not

00463

considering question number five? It's been moved that we not consider question number five. Is there a second?
DR. KLEE: I second it.
DR. FERGUSON: Okay. There's a second. Now, is there discussion?
DR. MURRAY: I'm a little puzzled by the problem, because we have come very close to number five. I have a question for Dr. Klee. In

your original now withdrawn motion, when you said selecting as it's written here, in selecting an appropriate cancer therapy, what exactly did you mean by selecting? Did it specifically include selecting and excluding? Because I do have a problem with -- I supported your motion to find that there is not sufficient scientific evidence for selection, but there is sufficient scientific evidence for excluding, so what exactly did you mean by selecting?

DR. KLEE: I was just reading it literally, so selecting was rule in, was predominantly, but I also had concerns about the rule out. I don't think there was sufficient scientific evidence for many of the disease entities or subgroups thereof to make a statement

00464

like that, so it was across the board that I had concerns. But I think it has been addressed as it has already been discussed in issue number three where we said there is promise, and we have one case where it looks like there is some clinical utility. So I, that was the basis of withdrawing this motion, is because it looks like we can't go further than what we have already said with issue number three.

DR. FERGUSON: All right. It's been moved and -- yes, go ahead.

DR. BROOKS: It almost gets to whether we want to say any negative. In other words, if we want to use five, not as being very similar to three, but whether we want to change it in such a way as to state that we don't think these have proven value in every cancer, because --

DR. KLEE: Is that not captured in the discussion?

DR. FERGUSON: Okay. Is there any further discussion about removing number five? All right. It has been moved and seconded.

Is there any discussion from the group, the audience, presenters about removing question number five?

00465

DR. NAGOURNEY: Robert Nagourney. And I think both three and five speak to an issue that Dr. Bosanquet raised, and which confronts me directly. We have in one course of discussion looked over different technologies, different end points, different utilities for end points. What I'm concerned by is that my work, which is specifically designated on the basis of what I believe to be a better scientific understanding of tumor biology, the concept of cell death, the measurement of cell death as being a robust predictor of response, my concern here is that HCFA will make a decision that these assays are all the same, and that the measurement of tumor biology can all pretty well be determined.

And to use Dr. Weisenthal's analogy where one finds the person on the roadside and in determining whether they're alive or dead, they can do a core temperature, EEG or EKG, or check for pulse or check or response to stimulus, one does not do a sperm count. You are not looking for proliferative capacity to assess viability. The assay end points that we have sort of skirted over are distinct. Some measure cell viability, and those have been extremely compellingly argued

00466

in favor of by much of the data, if you really dissect the data. Most of what you heard, which convinced you to these remarkable unanimous decisions has been Randy Stein, who was not determined to have been improved in his outcome by eliminating every other possible combination of drug resistant phenomenon, but in fact by identifying an active treatment.

Or Dr. Nalick, who eloquently argued in favor of how well the cells can pick treatments. Pick treatments. And what I'm afraid of here as a clinician who comes under HCFA guidelines, and who practices medicine, whose father has cancer, you could make a decision that you will approve all these tests and they're all really great, and although I know you are not here to determine reimbursement issues, I will find myself

constrained with a difficult and arduous assay which requires larger numbers of drugs under different conditions for prolonged periods of time with subjective and labor intensive tests, to make meaningful selections of cancer treatments. And I will be reimbursed by HCFA, or my patients will be covered by HCFA at a level that covers the lowest common denominator,

00467

eliminate a drug that has a five or ten percent chance. And I will be effectively unable to provide the best test to my patients. And HCFA stipulations say that you either accept HCFA, Medicare reimbursements for an approved test in every situation, or sign off HCFA for two years. What this effectively means is that you reimburse these all the same, and the cheapest assay becomes the assay that's reimbursed, then I write a prescription for my father if I don't get this test approved in a way that I can afford to do it.

So I think that number three and number five speak to issues that there are different tests here, and when you send your message to the next committee, there is going to have to be some distinction between the fact that some tests are difficult and give information to select treatments, and some tests are easier and give more limited amounts of information. And that's sort of be skirted over, and it concerns me gravely.

DR. FERGUSON: Thank you. Is there further discussion or comment on this removing this question.

00468

MR. STRINGER: I'm Jerry Stringer. I'm a consultant, although I am here on my own today. Just in terms of the committee guiding the development of the coverage policy, I actually think it would be important to make a statement -- you made a statement that it's -- basically that it's reasonable and necessary for this test on some occasions. I think as the

experts, it would be nice to know whether you felt that question five, I think says, are there occasions where use of this test would change which chemotherapy agent a patient would get. So I think that's basically all it says; if you do the test, is there a chance that the treatment would change.

Another level of it, does this test have the possibility, or is there scientific evidence that improves patient outcomes in terms of quality of life, and then ultimately the question is, does this test improve patient longevity? Is there scientific evidence on each of those three steps? And I think those questions being answered by the experts will help the coverage policy makers in formulating when the test should be covered, and under what

00469

circumstances.

DR. FERGUSON: Dr. Hausner?

DR. HAUSNER: I guess I didn't really reveal my full plan. If question number five is deleted at this time, my plan was to add number five as question number six, what additional concerns, questions or would the committee like addressed and basically in a rather clumsy way be, table it in that fashion. That was what I was going to do if it were still an open motion, if it were to be defeated.

The other comment that I've just got to say, talking about the Randy Stein case somehow or another influencing my opinion, that is a remarkable story, just that. I don't know what happened there. That could be explained by somebody trying for sainthood. I mean, Mother Theresa might have had some effect on that case as much as anything that we were told about. So that had no influence, although it's a very gratifying story.

DR. FERGUSON: Dr. Fischer?

DR. FISCHER: You know, I feel like I'm dealing with my kids here. I think, you know, I think the committee went as far as it could,

00470
given the science that it was presented, and I feel we are getting beat up on right now, and I'd give you the same recommendation I'd give my kids, settle down and wait a while.

DR. FERGUSON: Mr. Kiesner?

MR. KIESNER: Yes. I think when I look at this question, it is very broad, and I think that the general tenor of what I have heard here today is that there has been a wealth of scientific evidence which compares very favorably to other diagnostic tests, and the panel believes that there is appropriate clinical application of this, but we have not given you, nor have you had the time nor maybe is it appropriate for you to try to comprehend all of the clinical settings in which these types of tests can be used. I think that it is appropriate for this committee to say that there is sufficient scientific evidence for human tumor assay systems to be used in relation to selecting or deselecting a given drug. And then I think it has to go one step further in terms of the policy at some further point in time, and by an entity other than this panel in order to define that specificity. And I would feel that an answer to number five in that sense,

00471
holds that there has been scientific evidence, that there has been clinical utility, which would parallel the answer to question number three, and that, some indication that this should be used by HCFA as the sentiment of the committee, to look in more depth at the clinical setting, and I think that would be the most appropriate way to handle this.

DR. FERGUSON: Thank you. We're going to call this -- go ahead. One more.

DR. BROOKS: Yeah. Just a quick comment. I mean, I kind of agree with Dr. Fischer. You know, we are kind of being boxed around the corner here a little bit, because on the one hand you would like it to say that is of utility in selecting and deselecting the

chemotherapy. And I believe that, you know, with my father being a lawyer, if we say that sort of stuff, then we just voted on we wouldn't preclude therapy based on the assay. So I think it's gets too multiple on their questions. And if you're saying that you think there is clinical benefit as opposed to utility, then we come back to the other thing, and we certainly could, and I am not proposing any motion, but you know, then we could

00472

have a motion based on benefit, so I think, you know, there is various issues in this question.

DR. FERGUSON: I am going to call this question. All in favor of this removing number five? I believe that it's unanimous. Okay.

Now, does somebody want to -- I mean, there are what additional concerns, questions or issues? I haven't asked for a motion on that, but yes?

DR. HAUSNER: My motion is, I would like to make a motion that number five be incorporated as an additional concern for future consideration. I am a little -- when it says would the committee like addressed by who, I assume it's not by us, but I think that number five is still an open issue for the future as this story continues to develop. So assuming that it's not us, I make the motion that the committee recognize that the question number five, is there sufficient scientific evidence, et cetera, be addressed at a later date.

DR. FERGUSON: Is there a second? Dr. Sundwall?

DR. SUNDWALL: Could I amend that before I second it?

00473

DR. FERGUSON: Sure.

DR. SUNDWALL: The discussion to me is either or, which I don't quite understand. I think the problem word is sufficient, and I would support your issue to be on the table for further consideration if it read something like there is scientific evidence demonstrating the clinical

utility of ST assays; however, more research needs to be done to document their utility, particularly in solid tumors.

DR. HAUSNER: I accept that, and maybe you made the motion and I'll second it; okay?

DR. FERGUSON: It's been moved and seconded, I guess. Dr. Sundwall, do you want to read it?

DR. SUNDWALL: There is scientific evidence to demonstrate the clinical utility of STASs; however, more research needs to be done, particularly in documenting their utility in solid tumors.

DR. FERGUSON: Okay. And it has been seconded. Now, is there some discussion on that motion? Yes, Dr. Fischer.

DR. FISCHER: You know, it sounds like the answer is in on hematologic tumors, which it

00474

certainly isn't. You know, I think lots of questions come from this, particular tumors, particular assays, particular drugs, when does it and when doesn't it work. We don't know. I think we have really been pushed as far as the committee is going to, and so, I feel quite at piece about where we are at.

DR. FERGUSON: Yes?

MS. SIMMERS: It seems to me that for question six, what really needed is sort of a laundry list of those concerns and questions that we have remaining, but we're not going to come to a conclusion about making a motion about them, but that we want HCFA to know that they are concerns of ours. And I think this whole issue of clinical trials and their continuation or further research, whichever you way you want to state that, is one of the concerns that has been expressed several times. And I think if it makes the list, there is not really a need for a more specific motion, but just the sense of that, to be registered with HCFA.

DR. FERGUSON: Okay. Dr. Brooks?

DR. BROOKS: Yeah. I just wanted to

say that I would agree with the previous speaker
00475
that, you know, rather than have another motion, although we certainly could have that one motion, but I would not want that one motion to preclude giving the additional concerns or whatever that we may have, that we may want to voice.

DR. FERGUSON: Okay. Kathy?

DR. HELZLSOUER: Yeah. I guess the issue for me, that motion, sounds similar to what we already voted on, so I don't see, I guess the utility, if you will, of rephrasing what we already voted on. Think the issue that should be reflected is where we changed that was the clinical benefit. I agree with Dr. Fischer, that we've gone as far as we can with the evidence provided, and my concern is that we don't have the evidence of clinical benefit and that's what still needs to be shown, in whatever ways, and whatever trials, so that's where I have the concern.

DR. FERGUSON: Okay. Do you want to --

DR. SUNDWALL: Yeah, I would like to withdraw. I have to look at our FDA and see. If I can withdraw my motion, I think that we probably all listed some things we think are issues, and I wonder if maybe the committee needs
00476
to discuss that, or because we are duly appointed committee members, we couldn't in fact provide for you those issues.

DR. FERGUSON: Right. There is a sense of, maybe somebody could itemize these things. There is a sense of the committee that there are some issues that require addressing for which patients is this, are these the best tests, when, when should they be given, what tests, when along their treatment protocols. I mean, all kind of things of that nature and others, I'm sure. Yes?

MS. KRAFT: I think that's what Dr. Nagourney was getting at is he wants us to define some of our concerns, because all of us

that have dealt with Medicare and Medicare reimbursements are concerned with defining what will we be reimbursed for when we order HTA assay tests, and then, will Medicare take the flying leap forward and then define, unbeknownst to us, maybe what tests they will pay for and what they won't. So one concern of mine is that they, in defining what they're going to reimburse, that they contact some of the scientists and physicians in the audience that are doing this

00477

research, that they find out what is the cost of producing the test and get some real life cost data, so when they set what they are going to pay the physicians for doing these tests, that they have realistic up to date direct costs.

DR. FERGUSON: Maybe we could just, since we're doing pretty well on time, we have 15 more minutes, just put some of our concerns on the table for HCFA's consideration, as sort of our final. Yes, please?

MS. SIMMERS: I have three on my list and I'm sure there are going to be many others I agree with. One, I think this whole issue of continued research, and I believe the stimulus is there to do it, because as Dr. Bagley pointed out, oncology is much different, and I believe in order to convince those that are the gatekeepers of ordering these tests so that it opens up to Medicare beneficiaries, the research will have to support the use of that technology, so it should happen, but it is a continued concern that we get better evidence of the utility and benefit of these tests.

I continue to be concerned that the industry work on and continually be cognizant of

00478

accessibility of all Medicare beneficiaries who are facing a cancer diagnosis, and just not be some limited accessibility wise, and they look at ways to address that.

And certainly the policy development, for those of us who have dealt with carriers on a

daily basis, and for their side of the equation, the policy does need to be more specific. I don't think this is the forum where that happens, because there are processes in place that HCFA has used before to develop those kinds of policies, and I certainly want to see that kind of process go on, so that reasonable and specific policies are set forth.

Those were the tree three that I was concerned about.

DR. FERGUSON: Dr. Sunderwall, did you have some?

DR. SUNDERWALL: My only contribution at this time is that I think this particular group of tests under this rubric, whatever STAs, lends itself very well to a national coverage policy. We have experience from negotiated rule making where in fact this would be, could be done with the right expertise, and I would strongly

00479

recommend that be the next step from HCFA. I think it would address most of the concerns people have about appropriate application and whether it should be paid for.

And I would just second what Cheryl just said about appropriate reimbursement, because I do think that it would be a shame to give a green light to add this to the armamentarium of oncologists and physicians, and then find out that it's so underpaid that it's not being used.

DR. FERGUSON: Okay, thank you. Dr. Klee?

DR. KLEE: I had three different things that I'd like to see brought up. One is this question of monitoring the effectiveness of this program if it's put in place, and perhaps even having a sunset clause and review after a certain period of time, to say, did it really meet the expectations that we had hoped for for this length of time?

The second would be to further delineate this question of which tests are

appropriate for which type of tumors. You know, which ones are proliferative, which ones do we

00480

want to have apoptosis markers and such in there.

The other is a further delineation in terms of which types of patients are appropriate for testing. There are certain tumors that are going to have universally good response, or fairly good response, and it doesn't seem like this would be appropriate for that group of patients. And on the other end of the distribution, you've got some that there is no appropriate therapy, or responsive therapy that's going to be coming in, and therefore, treating may not be dependent upon this testing also. So I think it's along the line of the presentation we had yesterday, that we are looking in the middle part of the distribution rather than the extremes. We need to define what those extremes are, or what the middle is in terms of disease therapy indications for this particular technology.

DR. FERGUSON: Dr. Fischer?

DR. FISCHER: No.

DR. FERGUSON: Dr. Brooks?

DR. BROOKS: Yeah, a few things. One, I would go back to ASCO's position and so HCFA

00481

may, I wouldn't require it, but they may wish to have further clarification on how ASCO views these tests, just as additional information for the record, I suppose.

What I would like to say and make almost a recommendation for is that just as with certain testing that's done in clinical labs all the time, whether it be for HIV, hepatitis C, et cetera, you know, there is a requirement that we keep certain data, that if approved for coverage and payment, that there be a requirement of those who were ordering or doing the tests, that they keep certain data available, and that data be open and available to external groups, as our

data is now.

And finally, I think coverage, as mentioned by Miss Kraft a little earlier, or perhaps yesterday, I kind of agree with her, that coverage may actually unable further research to go forward. There will be some type of payment, no matter on what level, and that may shake out which test is better. It may actually foster further research to enhance the test and allow these tests to be used in clinical trials by the oncology groups, so I think that may well be the

00482

case.

DR. BAGLEY: I think I gave my concerns in the beginning.

MR. BARNES: Just very quickly, I would like to encourage in conjunction with the comment about looking at the true cost of the test, that the work at HCFA go further to look at the net cost to Medicare, that the economic analysis and quality of life analysis, which you heard a little bit about, be taken into consideration.

DR. FERGUSON: Okay. My concerns are mostly which tests for which patients, and when in the course of the disease, which I think need to be still looked into.

DR. MURRAY: I think that my comment perhaps duplicates Dr. Brooks, but I know that it's common practice in many, perhaps all genetics laboratories, cytogenetic laboratories that do prenatal testing, it's common practice for them to follow up with outcomes and to correlate their test result with the fetal outcome. And I would encourage the laboratories that do this type of testing to make that effort. Of course you can't demand it as a condition of testing, but our experience in

00483

genetic testing is that obstetricians are very cooperative and I would expect that the oncologists would be equally cooperative. While that doesn't constitute research and perhaps may not be publishable to the extent possible, that

should be available for review.
DR. FERGUSON: Dr. Loy?
DR. LOY: Based on Dr. Nagourney's comments, I hope some attention is give to elaborating on the differences between different testing modalities and when there may be appropriate use of each one of those modalities.
Then I also have an interest in addressing the appropriate frequency of testing, how many times you're going to allow this as reasonable and necessary, over the course of treating patients. And then finally, some attention to who is responsible for choosing the drug of choice for testing. If there was never the intent to use a specific chemotherapeutic agent in the regimen to begin with, then it would seem inappropriate to me that the oncologist should have, the treating oncologist should have some say so about that to begin with.
And then finally, I hope that there is

00484

some consideration given to, from a carrier standpoint, of the documentation requirement.
DR. FERGUSON: Miss Snow?
MS. SNOW: My only concern is that we keep in mind the assessability and affordability for the beneficiaries.
DR. FERGUSON: Thank you. Dr. Kass, no? Dr. Hausner, no. No?
DR. MINTZ: My concerns have been already stated by others, but I want to use this opportunity to state that I think the sense of the committee was best expressed in motion number three, and that these tests show promise for clinical utility, and that motion deliberately did not state, distinguish between sensitivity and resistance testing, so I think the sense of the committee reflects that it is supportive of both sets of testing.
And I would only add that I also hope the coverage is adequate to permit this technology to be used.
DR. FERGUSON: Dr. Bagley.

DR. BAGLEY: I want to do one little bit of parliamentary cleanup work, since in the frenzy of doing the right thing, we may have

00485

gotten ourselves cross wise with Ferguson's, or Robert's Rules of Order. We had a motion on number five, which was seconded. I believe someone reminded me that the questions was called. We can go back and look at the record, but I believe the question was called, and then there was this withdrawal. And actually I'm unclear as to whether that's allowable, but I think we could get ourselves, have a clean record if we consider the fact that motion number five was, the original motion which was, there is not scientific evidence was made, seconded, question called. If we vote on that and it's voted down, and the committee already then went on to vote, saying their sense on number five was that it was dealt with in number six, I think the record will clearly reflect it, but I think perhaps it would be worthwhile to clean up that issue and vote on that original at motion number five,.

DR. FERGUSON: He withdrew the motion.

DR. BAGLEY: Well, there's a question as to whether that's an allowable procedure after the question's been called, so I think if we voted on it, the committee voted on it, if they vote it down, they could then make a motion and

00486

say see question three in the discussion, that's our sense.

DR. FERGUSON: Okay. I guess we should vote on Dr. Klee's original motion. Restate the motion.

DR. BAGLEY: That there is not scientific, that there is not sufficient scientific evidence to demonstrate the clinical utility in selecting appropriate therapy.

DR. FERGUSON: All right. So I am going to call for the vote on that. All in favor of that? One vote in favor. I guess I have to read. That was Dr. Klee that voted in favor.

Do I have to read the -- no. All against? And I guess there is an abstainer or two. Wait a second. So everybody else voted against, is that correct? All against, please raise your hands. Dr. Mintz, you're not raising your hand; does that mean you're abstaining?

DR. MINTZ: Yes.

DR. BROOKS: And so am I.

DR. FERGUSON: So we have two abstainers, and I need to read who abstained? Boy, you guys really -- let's see. Dr. Mintz abstained, Dr. Brooks abstained. Did anybody

00487

else abstain? And all the rest voted against. You want me to restate that?

One voted for this motion, two abstained, and the rest voted against it. Okay. All we all right with Roger's, Robert's? Do I get by my badge for going to Congress.

DR. SUNDWALL: Before people leave, could I call to the attention something that people may or may not be aware of, that Dr. Bagley won't be with us anymore in this capacity. Dr. Bagley is leaving government, and all of us who've worked with him I think owe him a debt of gratitude for his fairness, his intellect and his perseverance.

DR. FERGUSON: The meeting is adjourned.

(The meeting adjourned at 11:55 a.m., November 16, 1999.)