

P R O C E E D I N G S

8:03 AM

MS. CALVIN: Before we begin, I would like to remind everyone that you need to sign in, and if you haven't done so, please do so?

Good morning, I am Veronica Calvin, the Executive Secretary for the Hematology and Pathology Devices Panel, and I would like to welcome you to this joint meeting of the Hematology and Pathology Devices Panel and the Immunology Devices Panel.

Before we move into today's agenda, I will read brief summary minutes from the last meeting of each Panel. The last meeting of the Hematology and Pathology Devices Panel was held on January 28, 1998. The Panel discussed the Autopap System manufactured by Neopath Inc., and voted in favor of recommending approval with conditions.

The last meeting of the Immunology Devices Panel was held on February 2, 1998. The Panel voted in favor of recommending approvable with conditions on the Hypertech(?) free PSA assay. They, also, discussed and made suggestions on proposed prescription use labeling for unitized bladder cancer tumor marker assay.

Today, the Committee will discuss, make recommendations and vote on a premarket approval application for an immunohistochemical device indicated for the detection of HER2 over expression in breast cancers.

Attached to your agenda you should find the specific questions to be discussed during the open Committee discussion.

At this time I would like to introduce our Chairman, Dr. Timothy O'Leary. He is the Chairman of the Department of Cellular Pathology at the Armed Forces Institute of Pathology in Washington, DC. He is, also, the Chair of the Hematology and Pathology Devices Panel.

I would like to introduce Dr. Steven Gutman. He is the Director of the Division of Clinical Laboratory Devices in the Office of Device Evaluation.

Now, I would like the Panel members who are here to introduce themselves beginning with Dr. Ladoulis who is the Chair of the Immunology Devices Panel.

DR. LADOULIS: Thank you. I am Chair Charles Ladoulis, formerly Chair of Pathology at Maimonides Medical Center, Associate Professor at Sunny Health Science Center Downstate.

DR. DAVEY: Diane Davey. I am a Panel member for the Hematology and Pathology Panel, and I am at the University of Kentucky in Lexington, Kentucky.

DR. FELIX: Juan Felix, also, a member of the Pathology and Hematology Devices Panel and I am an associate professor of pathology at University of Southern California.

DR. MILLER: I am Carole Miller from Johns Hopkins

Oncology Center. I am a member of the CBER Advisory Panel and served as one of the primary reviewers for the ODAC Herceptin meeting.

DR. FLOYD: Alton Floyd, the industry representative on the Hematology and Pathology Devices Panel.

MS. CALVIN: Thank you. Just as a reminder, please speak into the mike so that they can pick up.

Now, I will read the conflict of interest statement.

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 employers' financial interests.

However, the agency has determined that participation of certain members and consultants, the need for whose services outweighs the potential conflict of interest involved is in the best interests of the government.

A waiver is on file for Ellen Rosenthal for her

financial interest in a firm at issue, and waivers have been granted to Drs. Mary Kemeny, Charles Ladoulis and Harry Homburger for their interests with firms at issue that could potentially be affected by the Committee's deliberations.

The waivers allow these individuals to participate fully in today's deliberations. A copy of these waivers may be obtained from the agency's Freedom of Information Office, Room 12A-25 of the Parklawn Building.

We would like to note for the record that the agency took into consideration certain matters regarding Drs. Juan Carlos Felix and Carole Miller.

These panelists reported past and current involvement in the form of contracts, consulting and speaking engagements with firms at issue on matters not related to what is being discussed today.

Since these interests are not related to the specific issues before the Committee, the agency has determined that they may participate in the Committee's deliberations.

In the event that the discussions involve any other products or firms not already on the agenda for which an FDA participant has a financial interest the participant should excuse him or herself from such involvement, and the exclusion will be noted for the record.

With respect to all other participants we ask in

the interests of fairness that all persons making statements or presentations disclose any current or previous financial involvement with any firm whose product they may wish to comment upon.

Next, I will read the appointment of temporary voting status. Pursuant to the authority granted under the Medical Devices Advisory Committee Charter of the Center for Devices and Radiological Health dated October 27, 1990, as amended April 20, 1995, I appoint Carole B. Miller, MD, as a voting member of the Hematology and Pathology Devices Panel for the duration of the meeting on September 4, 1998. For the record, Dr. Miller is a special government employee and a voting member of the Biological Response Modifiers Advisory Committee of the Center for Biologics Evaluation and Research. She has undergone the customary conflict of interest review and has reviewed the material to be considered for this meeting, signed Michael A. Friedman, MD., Acting Commissioner of Food and Drugs.

Dr. O'Leary?

DR. O'Leary: Thank you very much. At this point we will begin the open public hearing, and the first part of this is the FDA presentation, and Dr. Peter Maxim will begin.

Are there any speakers before that? I am sorry, I am getting confused. Are there any people in the audience

here who would like to address the Panel prior to the FDA presentation?

In that case, we will go on to Peter Maxim.

DR. MAXIM: Good morning, Dr. O'Leary, members of the Hematology Pathology Panel and Immunology Devices Panel. As a means of introduction I would like to talk to you a bit about the PMA that you are going to be reviewing here today before we get into the actual review process and the examination.

On September 2, of this year, 2 days ago, the Oncological Drugs Advisory Committee met to review data associated with the Herceptin clinical trial for a monoclonal antibody indicated for the treatment of women with metastatic breast cancer.

The Panel recommended that this therapeutic product be approved by the FDA and that is currently being further reviewed by our Center for Biologics Evaluation and Research.

The PMA that you are looking at today and will be reviewing today is an immunohistochemical assay to measure over expression of HER2 and to select patients for Herceptin treatment. This is not, however, the assay used to select those patients for the Herceptin trial.

The Herceptin trial was under way and there was quite a bit of logistic issues associated with bringing back

those specimens to be evaluated using this test. The company then performed a study to demonstrate similar performance of this DAKO immunohistochemical assay to the clinical trial assay that was used to select those patients, and that was performed by LabCorp. We refer to both the LabCorp and the clinical trial assay today.

You will have an opportunity today to examine data to look at the concordance of this candidate assay to the clinical trial assay that was used in those studies. You will have an opportunity to evaluate the performance of this assay with respect to reproducibility and other analytical features, a chance to examine the scoring features of the assay and the semiquantitative nature of evaluating breast cancer tissues, as well as the accuracy of the assay as measured against a panel of tissue specimens which have been highly characterized by independent investigators.

As we move through the agenda, Dr. Goldstein from the Center for Biologics Evaluation will provide you with a biological overview of HER2 as an introduction. The manufacturer will then make its presentation regarding the clinical studies that they did perform for this PMA and the Center for Biologics Research reviewer will present to you the review of the Herceptin and a summary of the review and a summary of the ODAC Panel meeting recommendations regarding Herceptin clinical trial.

You, also, have as a resource today available to you the fact that your Chairman, Dr. O'Leary as well as Dr. Miller served on the ODAC Panel on Tuesday and can familiarize you and will serve as the liaison and a link between that particular meeting as well as this one.

Finally, the CDRH reviewer will go over our review of the PMA issues and the list of questions that we would like you to consider during your deliberations later on this afternoon.

Thank you very much and at this time I would like to introduce Dr. Julia Goldstein who will discuss the biology of HER2.

DR. GOLDSTEIN: Thank you, Dr. Maxim.

Good morning, and my name is Julia Goldstein, and I am the Chair of the CBER Committee and the product reviewer for Herceptin. The Center for Biologics and the Center for Devices have been closely working together in the review of these two products. Herceptin has a license application submitted by Genentech, and the proposed indication is for the treatment of patients with metastatic breast cancer whose tumors over express the HER2 receptor.

In parallel the Center for Devices reviews the PMA for the immunohistochemistry kit submitted by DAKO Corporation, and the indication for this kit is to determine patients' eligibility for treatment.

Dr. Susan Jerian will later summarize the outcome of the meeting from Wednesday. During my presentation I will address the following four issues. First, what is the biology of the HER2 receptor because Herceptin is a monoclonal antibody that binds the HER2 receptor and what is the physiological role of this receptor?

Second, what is the pathobiological significance associated with HER2 over expression? Then what is the clinical relevance of HER2 over expression and finally, I will briefly describe what is Herceptin's molecular structure and what is its mechanism of action.

HER2, also, called, ErbB2 belongs to a family of receptors called ErbB family. The family functions by homo and heterodimerizing with each other. They share extensive sequence homology which suggests that they have similar mechanisms of signaling and transduction.

On the right hand side of the slide there are listed some of the ligands known to bind to these receptors. As you can see there is no ligand that has been yet characterized that binds the HER2 receptor and the current view is that HER2 is the preferred dimer partner for all these other receptors, HER1, HER3 and HER4, and acts as a coreceptor for multiple signals and define the signals that are being transduced by the other three.

HER2 is a large protein. It is 185 kilodaltons.

It is a transmembrane glycoprotein. It consists of an extracellular domain with two cysteine rich domains which are picture in pink and throughout this presentation HER2 will be represented in pink. It has a transmembrane domain, single spanning and then it has an intracellular domain with tyrosine kinase(?) activity. Normal HER2 expression has been extensively studied in adult and fetal tissues and it has been found that HER2 expression was found in epithelial cells derived from the three germ layers, in particular in the gastrointestinal, urogenital, respiratory, skin, breast and also placenta.

Furthermore expression more recently has been found in neurons, Schwann's cells and glial cells and mast(?) cells. It has been found by the group of collaborators that HER2 plays a crucial role in cardiac and neural system embryonic development. This group develops in mice that do not express the HER2 gene and what they found is that the mice carrying the non-allele die at the embryonic age of 11 days due to lack of cardiac trabeculation and also, these mice had an altered development of the nodose ganglion(?) sensory ganglia and motor nerves. So, this demonstrated that one of the roles of HER2 is to regulate. It is involving mesenchymal epithelium communications during embryonic development.

As I said before, HER2 participates in an

interactive network of receptor-receptor interactions. Here you see HER2 that has heterodimerized with the HER1 receptor or EGF receptor, with HER3 and HER4. These interactions regulate crucial cellular function such as cell fate, growth and differentiation.

Upon ligand binding to the specific receptor and I want to focus on this side of the slide, for example, the EGF growth, the epidermal growth factor binding to the receptor, the intracellular tyrosine-kinase phosphorylates and now the complex of ligand receptor heterodimerized with HER2 which transphosphorylates. The tyrosine kinase has now become docking sites for multiple adaptor proteins and specific substrates and will ultimately regulate and map kinase activation and finally these little acellular functions of differentiation, survival and proliferation.

So, in summary, HER2 during embryonic development plays a crucial role in mesenchymal epithelial communication and in adult tissues it plays a role as coreceptor for multiple mesenchymal derived growth factors regulating crucial functions of the cells.

Under certain circumstances HER2 becomes over expressed, and I will discuss now what is the biological significance associated with HER2 over expression.

In vitro assays in particular cell transfections with the HER2 gene and in the genetic engineering of the

transgenic animal have shown that HER2 over expression is an important component of neoplastic transformation.

HER2 over expression leads to constitutive activation of the receptor and this leads to an increased proliferation rate and increased resistance to tumor necrosis factor alpha, decreased expression of adhesion molecules. In particular it was shown that ekaherins(?) and alpha 2 integrins(?) when they decrease, increase their incidence of metastasis development and progression and an increase in vascular endothelial growth factor secretion which supports neovasculature formation.

What is the clinical relevance that has been associated with HER2 over expression? Retrospective analysis of clinical data has shown that HER2 over expression is a prognostic indicator. Patients whose tumors over express HER2 have a shorter disease-free interval and a shorter overall survival and their tumors have a more aggressive disease progression regardless of the stage or node status. These tumors are more invasive, have a higher incidence of metastasis and are more resistant to chemotherapy.

I want to emphasize again that this is based on retrospective studies. Finally, what is Herceptin and how does it work? Herceptin is a recombinant humanized monoclonal antibody in which the complementary determining

region in the variable section is here represented in the lines derived from the murine parenteral antibody 45, and you will hear from 45 later, grafted into the backbone of a human IgG1. It contains 6 percent of murine residues and it has been shown by the sponsor that it binds to the extracellular domain of the HER2 with affinity.

The next two slides deal with how Herceptin works and these two slides are derived from part of the sponsor's publications submitted to the BLA and some of them that were not in the BLA but are from the bibliography existing.

So, in vitro studies have demonstrated that Herceptin exerts its effects basically due to two arms. One is the biochemical. That is in this slide and the immunological which is in the next slide. The biochemical arm is pictured inside the ellipse and is due to binding of Herceptin to the HER2 receptor through FABs. Herceptin has been shown to down regulate HER2 and together with the blockade of the heterodimers formation this produces signal transduction blockade. Also, it has been shown that Herceptin has cytostatic effect because it up regulates CDK2 kinase which is an enzyme that regulates cell cycle progression.

In addition Herceptin has shown that it restores expression of hyaluronate(?) molecules, again, ephrins and alpha-2 integrins and sensitizes breast tumor cells in

humans to TNF alpha.

The immunological arm of the response, again it is inside the ellipse is due to FC binding of the antibody to the FC receptor or CD16. In vivo assays have demonstrated that Herceptin mediates antibody dependent cell cytotoxicity and it is postulated that the in vivo mechanism would be the recruitment of CD16 bearing cells to the site of the tumor and this includes anti-cells, macrophages, monocytes and activated T cells.

Other in vitro assays and animal models have shown that Herceptin enhanced chemotherapy induced cytotoxicity. When administered together with cis-platinum it has a synergistic effect and in in vitro assays and animal models with Taxol, methotrexate and vinblastin has an additive effect.

This concludes my presentation, and I think that the sponsor has now the podium.

DR. O'LEARY: Thank you, Dr. Goldstein for this nice short synopsis of a very complex biology.

At this point we will begin the sponsor presentation, and I believe the first speaker is Dr. Gretchen Murray.

DR. MURRAY: Good morning, ladies and gentlemen, members of the combined Panels and the Food and Drug Administration, members of the audience.

As Dr. O'Leary stated, I am Dr. Gretchen Murray, Regulatory Affairs Manager at DAKO Corporation. We are here today to present a diagnostic in a rather unusual situation in this PMA application because DAKO has developed a diagnostic product to meet requirements for the Genentech therapy Herceptin.

Dr. Robert Cohen from Genentech will be sharing our presentation. In addition, Dr. Mads Roepke of DAKO Denmark will be presenting product-specific information including product development and testing.

For those of you who are not familiar with DAKO we are a privately held company that has been producing antibodies for approximately 30 years. We have over 2000 products available for both research and diagnosis, and the majority of which are for immunohistochemistry.

Why is there need for a commercial HER2 assay? We have a critical need at this point for a diagnostic test for the Herceptin therapy for women who have aggressive breast cancer. Herceptin happens to be over expressed in a fraction of people with breast cancer. During the clinical trials Genentech used an immunohistochemistry clinical trial assay to identify the subpopulation of patients who are HER2 over expressors. The diagnostic that is needed needs to be simple, accurate and reproducible in any of the laboratories that will be using it.

The development time line here represents both the Genentech Herceptin development as well as the diagnostic. In April 1995, Genentech initiated Phase III clinical trials for Herceptin. They approached DAKO in December 1996 to develop a commercial assay.

In the middle of March 1997, enrollment in Phase III clinical trials for Herceptin were completed. By September 1997, a prototype had been developed and DAKO and Genentech met with Center for Biologics Evaluation Research and Center for Devices and Radiological Health specifically Division of Clinical Laboratory Devices.

At that time the sponsor presented their plan for the PMA application testing. We proposed to do a concordance study of our assay against the clinical trial assay using the same laboratory that had done all of the testing for the Herceptin trials.

We agreed. We presented that we would do a 75 percent concordance using a one to one comparison of positives and negatives. The 75 percent concordance is the null hypothesis.

The Food and Drug Administration came back and requested that we do analyses of HER2 both genetic, nuclear material and protein over expression to ascertain the similarity between the immunohistochemistry outcome and other tests.

In May 1998, the BLA was filed. Two weeks later the premarket approval application was filed. Two days ago ODAC had their meeting where the Herceptin treatment was evaluated. At this point in time Dr. Cohen from Genentech will present outcome of the ODAC meeting and implications for the diagnostic. Oh, that is right, one more slide, excuse me.

In the premarket approval application we presented this proposed intended use that the DAKO anti-HER2 system is intended to detect HER2 over expression in breast carcinomas to select patients appropriate for treatment.

Dr. Cohen will now do his analysis.

DR. COHEN: Good morning, members of the combined Panels and FDA members and guests. It is a pleasure to be here. We owe a great deal of gratitude to the FDA. Following our meeting a year ago we asked to be here as close as possible with our Herceptin filing and Panel meeting, and we know that was difficult, and we really appreciate the opportunity to be able to have this Panel review DAKO's diagnostic.

My goal here is to just mention some issues that we heard discussed at ODAC, and I believe that Susan Jerian will consider these in much more detail.

At ODAC the pivotal Herceptin studies were presented and in aggregate they showed overall benefit in a

population of patients with 2 and 3 plus levels of HER2 over expression diagnosed using our clinical trials assay which as Gretchen mentioned was an immunohistochemical assay and we will discuss it in much more detail.

In that discussion the issue was raised that exploratory subgroup analysis of the pivotal data suggested that the magnitude of the treatment benefit varied by the degree of HER2 over expression, and we will show you some of that data.

A discussion ensued at ODAC relating to that issue and it was clear, at least to us that physician analysis of the risk/benefit for Herceptin will include considerations of the level of HER2 over expression.

Next slide, please?

The slide just shows some of the issues involved in considering the complex issue of Herceptin's benefit. This is from a comparative study which is called 648 in which patients received for first line treatment of metastatic breast cancer chemotherapy with or without Herceptin treatment. The primary end point in this study is time to disease progression.

This slide shows the time to disease progression on the left panel for a stratum of this study in which the chemotherapy assignment was to taxol, and the two lines here showed the time to progression for taxol treatment alone as

well as the time to progression for taxol treatment in combination with Herceptin. As you can see there is a difference between those two curves. The numbers of patients are small, but on the median there is a 2-month delay in time to progression when Herceptin is added to taxol.

In addition shown on the right panel here is a difference in response rates of approximately twofold with taxol alone producing a response rate of about 11 percent and taxol plus Herceptin approximately doubled that to 21 percent.

Next slide, please?

This slide dramatizes the difference between two plus and three plus patients with regard to the magnitude of the treatment benefit. This, again, is the paclitaxel stratum of the study and shows patients who were three plus by the clinical trials assay for expression of HER2, and this constitutes about three-quarters of the patients enrolled in this study.

Here you see a much more dramatic treatment benefit. As shown on this curve, the time to progression, there is a more than doubling for these patients of the time to progression by the addition to Herceptin to paclitaxel treatment, a very dramatic treatment benefit.

In addition that is also, shown by the response

rates on the right where paclitaxel alone produces a 17 percent response rate and in these patients with three plus over expression the addition of Herceptin produces a 50 percent response rate.

Next slide, please?

I will move more briefly through these slides. This is the other stratum of the study in which patients received anthracyclines in combination with Herceptin. For low-level expression, that is the two-plus patients there is essentially no delay in the time to disease progression following the addition of Herceptin compared to controls and in addition there is no effect on response rates.

Next slide?

For the three-plus patients again the benefit is much more substantial. This is highly significant as well as being clinically relevant with a delay in time to disease progression and an improvement in response rates. These are four different clinical situations that clinicians can expect to encounter in practice, and the discussion that ensued at ODAC suggested that information about the level of HER2 expression in patient samples was going to be relevant to treatment considerations, and I raise this up front in our presentation because I think this is the focus of the kind of information that we think the DAKO kit can provide to practicing physicians and their patients so that they can

make informed treatment decisions.

Next slide, please?

Mindful of the discussion at ODAC we would like to propose a slightly revised intended use here. We first raised this issue in a letter that we sent responding to FDA questions 2 weeks ago but we have now formalized it with this revised intended use statement which I will read.

DAKO anti-HER2 IHC system is intended for laboratory use to identify semiquantitatively by light microscopy HER2 over expression in breast cancer tissues routinely processed for histological evaluation.

The DAKO anti-HER2 IHC system is indicated as an aid in the assessment of patients for whom Herceptin treatment is being considered and that contrasts with the language in the original application in which we said that it was indicated to select patients for treatment. It is now clear, now that we know that there is a modifying effect of the level of HER2 over expression, it is clear that the issue is much more complex, we believe than simply identifying patients as positive or negative, and so we propose this revised intended use statement.

We will cover the following studies in support of the premarket application. Dr. Mads Roepke from Denmark who led the development effort will describe a pilot concordance study, reproducibility studies, as well as two molecular

correlation studies that Gretchen briefly mentioned.

I will then come back and describe what we believe to be a pivotal concordance study in which as Peter Maxim described we compare the performance of the DAKO IHC assay with the clinical trials assay used for the inclusion of the patients in the Herceptin clinical trials.

So, I will now introduce Mads Roepke.

Oh, one more for me. I guess just to highlight the issues and to let you know where we are going with this presentation we will present the evidence to you that suggests that the DAKO anti-HER2 system provides comparable results to the clinical trials assay, and we believe that because of the comparability that the results of this kit will aid substantially in the assessment of patients for whom Herceptin treatment is being considered.

DR. ROEPKE: Good morning, ladies and gentlemen. My name is Mads Roepke. I am come from DAKO in Denmark, and I have been in charge of the development of the DAKO kit. During the next half hour I will go through some background information concerning the kit. I will comment on some development issues, and I will give some details concerning the reproducibility study and comment on some supportive studies.

In this picture you see the kit and its components. In the front you see reagent vials and a slide

holder with control slides and in the back you see the kit box and the kit insert.

The kit contains these nine reagents, and I would like to focus your attention on the primary anti-HER2 antibody, the visualization reagent and the control slides.

When the primary antibody, the anti-HER2 antibody is applied to the section it is followed by the visualization reagent that visualizes the immunological reaction in the tissue. The control slides are important as controls for kit performance as I will return to.

Furthermore, the kit contains buffers, a negative control reagent and substrate buffer and chromogen. The aim of the development has been to develop a commercial immunohistochemical kit for the detection of HER2 protein over expression in carcinomas to select patients appropriate for treatment with Herceptin.

The primary consideration of DAKO during the development of our kit has been as follows: Using the kit results should be obtained that are concordant with the results obtained with the clinical trials assay. The kit should provide reproducible results across laboratories. It has to be suitable for manual as well as for automated staining, and the kit will contain reagents, performance controls and a standard procedure assuring consistent immunostaining.

The DAKO kit utilizes a rapid polyclonal antibody to HER2, code number AO485 and the clinical trials assay has used two monoclonal antibodies, the 4D5 and the CB-11. This clinical trials assay was used to select patients for treatment with Herceptin.

In this slide, you see some background information concerning the HER2 protein and the antibodies reacting with it.

This represents the cell membrane. As you can see the HER2 protein is a transmembrane protein receptor with an extracellular domain and an intracellular domain.

The monoclonal antibody clone 45 reacts with an extracellular epitope whereas the monoclonal antibody CB-11 and the polyclonal code number AO485 react with an intracellular domain epitope near the C terminal of the HER2 protein.

The monoclonal anti-HER2 antibody clone 45 is the murine parent antibody that has been used as a source for development of the humanized antibody used by Genentech for treatment of HER2 positive breast cancer patients.

As the monoclonal antibody clone 45 has been used for development of Herceptin, it would seem reasonable to use this antibody, also, for selection of patients appropriate for treatment with Herceptin.

However, 45 was developed for use in potential

therapy and turned out not to be applicable for use in immunohistochemistry. Further the complexity of the staining protocol using this antibody made it less suitable for use in a commercial assay. It was, also, found that a 45-based assay exhibits a lower sensitivity than the AO485 and CB-11 based assays. As a consequence this antibody was disregarded in the further development.

As shown by Dr. Goldstein the ErbB family of membrane receptors consist of four members, ErbB1 or EGF receptor, ErbB2 or HER2 receptor, the ErbB3 and ErbB4 receptors. As you can see, a certain degree of homology exists between these four members of the ErbB family.

The polyclonal antibody AO485 has been tested for reactivity to other members of the HER2 family and no cross reactivity in immunoblotting techniques has been seen.

I would like to invite your attention to Panel A where the five lanes represent antigens shown on the right side, namely a control transfectant, and EGFR or HER1 expressing cell line or HER2 transfectant, HER3 transfectant and a HER4 transfectant.

Please notice that only Lane C with the HER2 transfectant shows positive reaction whereas no reactivity whatsoever is seen in either A, B, D or E lanes, and this documents the specificity of HER2 antibody AO485

In the clinical trial assay and, also, in all

studies carried out using the DAKO kit the following scoring system has been used. The score will go from zero to one, two and three plus. The scores designated zero and one plus are considered negative for HER2 over expression, whereas the two plus or three plus staining as positive for HER2 over expression.

In the clinical trials assay patients with a positive score for HER2 over expression were eligible for Herceptin treatment. This scoring system was developed during the early phase of the clinical trials and used by Genentech for selection of patients for treatment.

DAKO adhered to this system during the entire phase of our development, and in this slide you see a breast carcinoma exhibiting a typical positive HER2 immunohistochemical staining of the tumor cell membrane and on the next five slides I will demonstrate the staining pattern and describe the reactivity seen giving the actual score. At the same time I would like to put your attention to the booklets passed around with pictures of immunohistochemical staining since the quality of those pictures might be better than the ones you see on the screen.

This slide shows pictures of three cell lines included in the kit as performance controls. The MDA cell line gets a score of zero and is used as a negative control

for the kid. The MDA-175 is a one plus, a weak control. This is an important performance control and negative staining of this particular cell line invalidates the staining run. The SK-BR-3 cell line gives an immunohistochemical score of three and is an example of a strongly positive HER2 staining.

It is important to know that these controls are included to assess the assay performance and are not to be used for quantification of the test results.

This slide shows a breast carcinoma which will be given a score of zero. As you can see no staining is seen in this case. A score of zero will, also, be given if less than 10 percent of the tumor cells are stained.

On this slide a one plus immunohistochemical staining is shown. This is equivalent to faint, barely perceptible membrane staining in more than 10 percent of the tumor cells. Notably the cells are only stained in part of the membrane, and this is an example of the two plus staining of the breast carcinoma. It is a weak to moderate complete membrane staining which is observed in more than 10 percent of the tumor cells. This is to be considered as positive for staining of over expression of the HER2 protein in the clinical trials assay.

Finally, a three plus staining is a moderate to strong complete membrane staining observed in more than 10

percent of the tumor cells, and again, you see a breast carcinoma with a strong staining of all tumor cells, as noted as a brown rimming of the tumor cells.

This slide summarizes some of the differences between the clinical trial assay and the DAKO assay. As mentioned, the clinical trial assay in fact consisted of two separate assays using two different antibodies. It is a complex assay where two different procedures for antigen retrieval are used.

The assay time for the clinical trials assay is overnight. In comparison the DAKO assay is a single assay based on one antibody. It is a simple assay with an assay time of less than 3 hours. The clinical trials assay requires four tissue sections in order to evaluate a patient's HER2 status whereas the DAKO assay only requires two sections to do this.

Lastly, the CTA comprises over 35 steps whereas the DAKO assay only takes 11 steps to complete.

The reduced number of steps is, also, desirable since it increases the robustness and reproducibility of the system and it, also makes the kit more user friendly, and here we have the staining procedure of the DAKO kit.

After deparaffinization of the tissue slides a heat inducted epitope retrieval procedure takes place. This is followed by blocking of endogenous peroxidase and

incubation with the primary antibody.

Then an incubation with a one-step visualization reagent takes place and finally an incubation with substrate chromogen and counter staining and mounting.

I have highlighted these three because those are the parts of the development I would like to comment upon in the next part of my talk.

Now, I will turn to developmental comments. In order to design a reproducible test we set out to make a standard procedure and a simple assay. The kit was to contain ready-to-use reagents and performance controls.

Two primary antibodies were evaluated in a pilot study, the polyclonal anti-HER2 antibody, AO485 and the monoclonal anti-HER2 antibody, clone CB-11.

In the pilot study the two antibodies were compared in 103 mammary carcinomas in comparison with a clinical trials assay and the purpose of the pilot study was to select the optimum antibody for the IVD kit and furthermore to obtain a preliminary estimate of the degree of concordance between the two candidate kits and the clinical trials assay.

The design of the pilot study is outlined in this slide. One hundred and ninety-seven breast tumor samples were obtained from Vanderbilt University and were screened for HER-2 over expression using the clinical trial assay.

Of these 197 cases, 103 samples were selected and tested with both the A0485 and the CB-11 based kits. These 103 samples were selected in order to gain an equal number of negative and positive cases being zero to one plus as negative and two plus and three plus as positive.

Here you see the staining intensity using a clinical trial assay of these cases. You see an equal mix of zero and one plus and two plus, three plus cases are shown.

When comparing the clinical trial assay and the A0485-based assay a very similar distribution is obtained as you can see from this and in a four-by-four table the following results are obtained. You can see the clinical trial assay and the cases giving zero, one plus, two plus and three plus are shown here and the DAKO kit is shown here with the cases representing zero, one plus, two plus and three plus totaling 103 cases. The diagonal is highlighted in yellow showing the perfect agreement between the two assays and this sums up to a total agreement, an overall agreement of 72 percent.

In a two-by-two table where zero and one plus are considered negative and two plus and three plus are considered positive a concordance of 89 percent was obtained. The results from the pilot study demonstrated that a polyclonal antibody A0485 as well as the monoclonal

antibody CB-11 gave results in concordance with the clinical trials assay.

Furthermore the performance of the two antibodies was comparable in the pilot study. The monoclonal antibody clone CB-11, however, occasionally exhibits some unwanted cytoplasmic staining, and the polyclonal antibody is cost effective and reliably manufactured.

As a consequence we chose the polyclonal A0485 for the concordance study. We, also, had to select a visualization system for our kit. Two different visualization systems were compared. The first designated EnVision/HRP was a dextran-based polymer where a secondary antibody is coupled to HRP labeled dextran. The other system is an Avidin-Biotin system.

The structure of the EnVision visualization reagent is outlined in this slide. The EnVision reagent consists of HRP enzyme molecules shown in orange and secondary antibodies shown in green covalently linked to a dextran backbone, and a primary antibody shown in yellow reacts with the HER2 antigen in the tissue.

The visualization reagent is added and the secondary antibody reacts with the primary antibody. In this way the HRP enzyme reporter molecules are linked to the primary antibody in a single-step procedure avoiding the use of the traditional Avidin-Biotin interaction. The HRP

system was chosen since it is cleared by the FDA and has been used by DAKO for more than 5 years. It gives low background since it avoids the detection of endogenous biotin in the tissue sections. It is a simple two-step procedure. It is sensitive, and it provides reproducible results.

A buffer for pretreatment has, also, to be selected and during the development we considered three different buffers for pretreatment of tissue sections. The citrate buffer pH 6 was chosen since this buffer is a very well-established buffer for pretreatment of paraffin-embedded tissue sections, and, also, this buffer provides reproducible results.

In order to obtain an optimal immunostaining the use of heat-induced epitope retrieval is required. The source of heat was, also, evaluated and we compared four different sources of heat, the water bath, the microwave oven the steamer and the pressure cooker.

The water bath was chosen for source of heat since this is an effective method for heat-induced epitope retrieval and, also provides reproducible results and the water bath is widely available in many laboratories.

In our kit insert the use of the water bath as source of heat-induced epitope retrieval is recommended. As mentioned the cell lines were used as controls in the

clinical trial assay and during the development of the DAKO kit. The HER2 protein expression of these cell lines is well characterized.

The three cell lines were obtained from American tissue culture collection ATCC and a working bank is established at DAKO.

During the clinical trial assay and during our entire product development of the DAKO HER2 kit three control lines, cell lines have been utilized. The three lines have been produced and embedded in paraffin. The three cell lines are the MDA-231, MDA-175 and the SK-BR-3.

These lines contain a consistent number of HER-2 receptors on each cell as determined by Scatchard analysis. Furthermore the different IEC scores shown here were given to the cell lines when tested in immunohistochemistry. The MDA-231 has approximately 20,000 HER2 receptors per cell and gives an IEC score of zero whereas the cell line MDA-175 has around 90,000 HER2 receptors per cell and immunohistochemistry gives a one plus score.

Finally, the SK-BR-3 has more than 2 million HER2 receptors per cell and gives an IEC score of three plus. Here, I, again, show you the staining of these three controls, the zero, the one plus and the three plus.

Next slide?

These controls have to be included in every

immunohistochemical staining run to serve as performance controls of all kit reagents. A lack of staining of the one plus control line invalidates the staining run. The use of cell lines provides a secure supply of these controls.

In summary of the product development it was stated that a kit has been developed that contains optimal reagents, performance controls and a standard procedure. Furthermore the kit is appropriate for testing in reproducibility studies, and now, I will turn to the reproducibility studies we have conducted using the HER2 kit.

The plan of the reproducibility studies is as follows. Interlaboratory testing was conducted using both manual staining procedures and automated staining.

Intralaboratory testing was conducted doing within-run tests, run-to-run tests and manual versus automated staining.

The design of the interlaboratory studies was as follows. The goal was to determine the percent agreement of 40 specimens among six laboratories using manual and automated immunostainers. For the manual staining two labs qualified. Presently one more additional study is in progress. Using automated immunostainers three of these laboratories were qualified.

Some of the test laboratories were disqualified

from the reproducibility study and the reason for disqualification was failed staining of the one plus control. It was found that several of the laboratories did not follow the provided procedure and primarily the laboratories performed suboptimal heat-induced epitope retrieval.

This slides represents results from the interlaboratory testing using manual staining and two laboratories have been compared. On this site we have results reported as zero, one plus, two plus and three plus from Site 1, and here we have cases reported as zero, one, two and three plus from Site 6.

The results highlighted in yellow on the diagonal show the cases where perfect agreement between the two sites was obtained. There are a few discrepant cases as you can see but only one case differs from the perfect agreement diagonal by more than one plus.

The overall agreement is 70 percent. Results from the interlaboratory testing on automated immunostainers are shown in this slide. Laboratories are compared as shown in the left side and in a binary system using negative or positive results being zero or one plus, two plus, three plus we have these levels of concordance. Using the zero, one, two and three scoring scale percent agreement between 63 and 64 is obtained.

This is the design of the intralab reproducibility testing. Within run and run-to-run tests have been performed. In the within-run testing a total number of five mammary carcinomas were tested and three sections from each carcinoma were stained in three laboratories.

In the run-to-run testing the same material was stained on two additional days, once by the original technician and once by a different technician.

Further a comparison between results from manual and automated stainings from two qualified laboratories were compared.

This slide shows results from one within-run reproducibility study on five specimens using automated immunostaining and here you can see the three replicated results from the three replicates of the same specimen, as you can see perfect agreement on all five specimens was obtained across the three replicates.

The next slide shows an example of results from a run-to-run reproducibility study on five specimens using automated immunostaining. Two different lab technicians participated, and again, perfect agreement is obtained between the 4 days and between the two technicians participating.

Here we see the results from a study comparing results from automated and manual immunostaining using the

kit. This is conducted at one site. Again, the concordant results lie here in the yellow. Seven cases are discordant in this study, and they all lie within the one plus limit.

Comparing results originating from manual and from automated staining the following results were generated. In a binary system negative or positive a level of agreement between 88 percent and 100 percent was obtained. In using the zero to three plus scale the agreement was between 83 and 92 percent.

In order to compare the results obtained two different readings of the same slides were conducted, an inter-reader and inter-reading reproducibility study. This is an example of such a study where a set of 40 slides was read twice by the same evaluator about 1 or 2 months apart. The two readings were concordant in 68 percent of the cases. Again, the outliers are within the one plus limit.

As a conclusion from the reproducibility studies it can be stated that the use of proper controls, performance controls ensures the reproducibility of the test system, and when the procedure is followed the kit provides reproducible results.

Furthermore the kit performs reproducibly among laboratories and within laboratories, that is between days, between technicians and between automated and manual immunostaining.

In order to gain more information on the specificity of the HER2 kit some supportive studies were conducted as suggested by the FDA. The first study set out to evaluate the HER2 staining intensity by the HER2 kit in a panel of breast tumor samples thoroughly classified with regard to HER2 amplification and expression by the group of Dr. Michael Press.

The HER2 gene amplification in the breast tumor samples was determined by Southern blot using DNA extracted from frozen material and FISH on paraffin-embedded specimens.

The RNA expression was tested by Northern blotting and protein expression by Western blotting and immunohistochemistry on frozen sections.

The HER 2 kit was tested on 168 such cases representing tumors with various levels of HER2 amplification and over expression. The study was conducted in the laboratory of Dr. Michael Press.

The results from this study are outlined in this slide. Dr. Press graded the breast carcinomas in four groups, the 1L group is a normal or single gene copy group considered negative by Michael Press. The 1H group is a single copy DNA but with over expression with other methods of the HER2. The two to five OE or over expression group gives a moderate level of HER2 gene amplification and the

above five over expression group has a high level of HER2 gene amplification.

All these three groups were considered positive by Michael Press. These 168 cases were tested using the DAKO assay and the score was classified as negative, that is the zero and one plus cases and positive, that is the three and four plus cases.

As seen from the results all 99 cases considered negative by the press designation were, also, found to be negative using the DAKO kit. Furthermore 26 of the 30 amplification cases were found to be positive using the DAKO kit.

The 1H and two to five 0 expression groups are borderline groups and they are difficult to classify by any single test. The 1H group has been shown to have a high stromal content which makes them difficult to evaluate and the two to five over expression group can be expected to have displayed a range of protein expression patterns.

In conclusion an 85 percent agreement between the DAKO kit and the Press classification was found. Furthermore no false positives with the DAKO kit were found.

The second supportive study was a comparison between the DAKO assay and a FISH-based assay for demonstration of HER2 gene amplification. The purpose of the study is outlined on this slide.

Forty-two breast carcinoma samples obtained from three sites were tested and each sample was tested for HER2 gene amplification using a FISH assay, and developed and tested as Cytometry Associates methods. The HER2 protein over expression by IHC was done using the HER2 antibody A0485 by DAKO, and here you see the results from the study.

Using a tumor FISH ratio cutoff of 1.3 and a tumor FISH ratio is equivalent to the mean number of HER2 genes per chromosome 17 in the tumor cells, a feasibility study carried out at Cytometry Associates established a cutoff value of 1.3 which is the normal level plus two standard deviations.

Using the cutoff value we ended up with these results. As you can see, using a DAKO score of seronegative being sero one, one plus and positive between two plus, three plus cases we have a concordance of 82 percent, namely 33 cases of the 40.

So, as a conclusion from this study 82 percent concordance rate between the DAKO test and the FISH assay was found.

In summary a kit has been developed that yields reproducible results. It demonstrates agreement with other molecular measures of HER2 status, and it is appropriate for testing in a concordance study and now I will turn over the floor to Dr. Robert Cohen to comment upon the concordance

study.

Thank you.

DR. COHEN: The outline from my presentation of this concordance study is shown here. I want to take you through the study design and the results and then I want to take some pains to extrapolate the results in the sample population that will be used to establish concordance to the expected population of patients with metastatic breast cancer. This issue was briefly touched on at the ODAC, and I think it is an important one that we want to consider carefully.

Finally, we will discuss clinical implications relative to a positive or negative assessment based on staining intensity and ultimately I will end by discussing staining intensity level as perhaps a more reasonable guide to the clinician for the appropriateness of Herceptin treatment.

The goal of the concordance study was to demonstrate that the DAKO assay exhibited performance characteristics comparable to the clinical trials assay in guiding the selection of patients appropriate for treatment with Herceptin.

Next slide?

The source of tissue is important, and Peter Maxim touched on it. The first thing to notice is the second

bullet point there. Tumor blocks were not available from Genentech's Herceptin trials. We examined, in our pivotal trials we treated over 700 patients. They came from hundreds of hospitals. Primary tumor specimens were sent to a central reference laboratory and returned to the referring hospital.

In addition to do a proper evaluation of concordance we would have needed negative cases. Those patients were not enrolled in our studies as they were not suitable for treatment. Consents were not available from those patients and so we didn't have available tissue to do a proper concordance study on patient material.

In addition, as Gretchen Murray told you, it was close to the end when this kit was available, the kit that Mads just described to you. We only had, well, it was after the time that we had enrolled our last patient into the pivotal studies.

So, instead we proposed in September 1997 to the combined meeting of CBER and CDRH, we proposed that we made reference to a tissue bank and the bank that we used was the NCI supported bank called the Cooperative Breast Cancer Tissue Resource or CBCTR. That is a bank of paraffin-embedded breast cancer tissue specimens for which some follow-up and demographic information is available.

There is a total of 9000 cases which are present

at four geographically distributed sites, and we made application to the CBCTR and we are pleased to be able to use their resource.

Next slide?

The study objective is shown here. Our goal was to determine if the level of concordance between the DAKO kit and the clinical trials assay is acceptable and by concordance just to be clear we are talking about the proportion of samples classified as positive by both assays and by positive as Mads described positive means for us appropriate for treatment. It doesn't necessarily mean and I want to avoid confusion here, it doesn't necessarily refer directly to the level of protein over expression. We were meaning positive here in the sense that we would have treated such a patient or offered treatment to that patient, so, proportion of patients classified as either positive by both assays or negative by both assays positive being two plus or three plus, negative zero or one plus.

With regard to acceptable concordance the hurdle that we discussed when we met with the FDA in September 1997 and in a follow-up telecon we discussed a target concordance of 75 percent or better, and importantly that 75 percent predicated on an equal mixture of positive and negative cases.

Next slide?

Why 75? As Mads described to you, the clinical trials assay was a complex assay. It required two independent assays for immunohistochemistry and the algorithm that was then used was the higher staining intensity score was the one that used. Such a complex assay obviously represents a more challenging standard to duplicate with a single streamlined test.

In addition we had ample evidence from our own development work clinical development work with regard to these two assays, 4D5 and CB-11 which were actually the components of the clinical trials assay, and in our hands the level of agreement between those assays which we believe to be credible and which we used as the basis of entering patients onto our studies was 79 percent, and then finally in the concordance study that Mads just described to you we obtained a point estimate of 89 percent in a small study of tissue from a single site. The confidence interval of 82 percent to 94 percent suggested to us from a single site that we should aim slightly lower and in particular, and it may not have been immediately clear the distribution of cases in that concordance study was skewed to the extreme. There were zeroes and three pluses and fewer cases in the middle, and we thought that concordance in a real world population might be lower as there were more cases pushed closer to the cut points.

Next slide?

Why a 1:1 ratio? Why not the real world frequencies? The reason is to minimize overestimation of concordance, and our own data as we will show you suggested that zeroes, for example, constituted close to 60 percent of the patients submitted for study. So, one could just imagine the example here if you had a population that was 90 percent positive and a test which was positive 100 percent of the time, you would have a concordance of 90 percent but you would have an essentially worthless test, and so the most robust test of concordance in a skewed population is to push for an equal mixture of positives and negatives and that is what we did, but having done that I think it is important to, and we will extrapolate that onto a population of the type that we expect to encounter in real world clinical practice, and we had discussed that as well with the FDA a year ago.

So, the sample size and composition is shown here. The 1:1 ratio I just described. Those cases we imagined and did power calculations to show that 600 total samples distributed in a 1:1 mixture would provide 90 percent power to detect a 5 percent improvement over the prespecified unacceptable level of concordance of 75 percent. The 90 percent power was obtained at the 5 percent significance level, and we intended to use a one-sided test because we

were only interested in proportions above 75 percent.

We estimated that the tissue bank might supply samples at the positivity rate of 25 percent or so, and that we would need to screen 1200 samples to ensure that we got close to the 300 positive samples that we required.

Next slide?

This slide shows the contributing centers for the CBCTR. There are four sites, one in Portland, Oregon, one in St. Louis, one in Miami and one in Fox Chase in Philadelphia.

A total of 1198 specimens were screened, and they came from these sites as shown. Slightly fewer came from Miami, but otherwise the balance was pretty good.

Next slide?

The flow of the work is shown here. The staining intensity was measured on all specimens using the clinical trials assay, using both assays which constituted a clinical trials assay, that is 4D5 and CB-11 and then the staining intensity score the higher of the two was derived from that analysis and that was the CTA score.

A subset was then identified which consisted of all the positive cases and a random selection of the negatives to produce the 1:1 mixture of samples. In order to conserve tissue at the CBCTR we didn't obtain sufficient tissue on all 1200 cases but rather we sent the accession

numbers back to CBCTR and obtained additional sections so that we could perform the DAKO kit on the selected subsample and then we assessed the agreement between the clinical trials assay and the DAKO assay, and I will take you through that analysis.

Next slide?

Of the 1198 specimens that were initially received and sent -- oh, by the way, all of this work at the central laboratory LabCorp in North Carolina. LabCorp was our central reference lab for all of our clinical work that Genentech did in support of Herceptin. They screened over 7000 patient samples in the course of that work and are a well-qualified lab to do this study.

They received those 1198 specimens of which 1190 proved to be evaluable by the clinical trials assay. Two hundred and seventy-five were identified as positive and 915 were identified as negative, that is zero or one plus.

Of these 915 negative cases a random sample was obtained to match these 275 positives and that consisted of 278 negative cases.

The accession numbers were then sent back for these 553 cases, and we obtained sufficient tissue to perform the DAKO assay. Of the 553 cases 548 proved to be evaluable by the DAKO kit.

Next slide?

Just a brief word on the non-evaluable cases. In general we were pleased with our ability to screen these cases. In more than 1700 assays we lost only a total of 13, and of those 13, 11 of them were because the H&E stain couldn't find evidence of tumor. Only in two cases was the tissue specimen destroyed or damaged in the processing.

Next slide?

This slide shows the raw numbers for the concordance results which was the prespecified primary end point of the study. Shown above is the results of the clinical trial assay as described by negative or positive.

Shown on the left is the negative or positive result by DAKO and again these are one plus and zeroes in the negative and two plus and three plus in the positives.

Of the 548 samples evaluated by both assays 431 were concordant; 215 were negative by both assays; and 216 were positive by both assays, and the discordant cases were evenly distributed as you see here.

Next slide?

This slide shows the concordance results and the statistical analysis on the point estimate. The raw concordance in the sample population was 79 percent. The statistical testing to show that that was greater than 75 percent was confirmatory with a P value of $P = .0153$. In other words, that point estimate is significantly higher

than the prespecified 75 percent, a level which we regarded as unacceptable. The 95 percent confidence interval on that estimate ranges from 76 to 82 percent.

Next slide?

Now, I want to take you through the extrapolation of these results onto the expected distribution of cases that we expect to encounter as we go forward and screen patients in the real world for Herceptin treatment. The distribution of the cases we used in the sample set I just described to you is shown on the right, and as you see 50 percent of the cases were zero and ones which was by design and 50 percent then were twos and threes.

In addition what you notice here in the green and red is that fully 40 percent of these cases were one and two plus. That is not the distribution that we expect to encounter in the real world, and we have over 7000 cases from LabCorp in support of that. This shows 1000 of those cases, but we have the rest of the data which is essentially not different.

Fifty-eight percent of the patients that we screened in our four Herceptin clinical trials, 58 percent were zero, and this is the clinical trial assay that included both 4D5 and CB-11. Fifty-eight percent were zeroes and 23 percent were three pluses. Nine percent were one pluses and 10 percent were two pluses. So, half the

number of cases were clustered close to the cut point in the real world population distribution compared to the sample that I just described.

Next slide?

So, we can use that information to develop an extrapolated concordance, and this is a straightforward analysis. We take the clinical trials results and simply adjust the numbers to yield a distribution that I just described. So, these columns, relative numbers in each column total to 58 percent in the zeroes, 23 percent in the threes, 9 percent in the ones and 10 percent in the two pluses.

The performance characteristics, that is how the DAKO assay discriminates within each column is determined by the actual results that we obtained in the concordance study on the sample population. So this is a very simple extrapolation.

Shown in color are the samples which agreed with regard to the positive and negative determination which was the primary end point of the study and then shown in white are the samples which were discordant.

Next slide?

This slide shows the extrapolation in another way and allows us to derive a concordance estimate for the real world population. Again, the total percentages are shown

here. This is the clinical trials level of positivity or staining intensity score and as you work through each column for zeroes the percent of cases which would match, that is be negative is 50 percent and the non-matches which would be positive among the zeroes would be eight and so forth, and we then total those up and derive a concordance estimate for this population of about 83 percent which means that 17 percent were discordant.

In addition, this allows you to evaluate the probability of matching by staining intensity score and you can see as you would expect that in the extremes the test is more concordant. Among zeroes the likelihood of getting a negative score by the DAKO assay was 86 percent and among the positives, the strong positives, the three pluses the probability of a match was 94 percent.

In the middle concordance is less good and I think this is exactly analogous to the situation that Mads described with regard to the Press molecular characterization study.

These are more challenging cases and as with any cut point some samples will wind up on either side of the cut point.

Next slide?

This slide simply shows in a two by two grid the data that I just described to you. As I said, 83 percent of

the cases were concordant; 55 percent would be expected to be concordant, 55 percent expected in the negative group, 28 percent of the concordant cases in the positives and along this axis 12 percent of cases would be expected to be positive by the DAKO assay and negative by the clinical trials assay, and 5 percent positive by the clinical trials assay and negative by the DAKO assay, and it is these discrepant cases and their clinical implications that I want to discuss in some detail.

Next slide?

Again, with regard to positive and negative we are using positive and negative to mean two plus and three plus. That is the exact entry criteria that we used for our study. Okay, what about these 12 percent of cases? The clinical implication of being positive by DAKO and negative by the clinical trials assay is that such patients would be selected as appropriate for Herceptin therapy. The potential net benefit of Herceptin has not been assessed in patients like these.

All of our data presented and reviewed by ODAC 2 days ago was in patients who were two and three plus by the clinical trials assay and therefore we can make no estimate of the net benefit that might or might not accrue to these patients.

It is true that the DAKO assay may, in fact,

reflect some level of HER2 over expression status. Mads showed you in the Press study that the rate of false positives was zero among 99 negative cases but with regard to the level of expression sufficient to get into our studies that may be a lower level.

So, the bottom line is we don't have information about these patients and cannot speak to what risks or benefits they might experience.

Next slide?

With regard to the 5 percent of patients who would be negative by the DAKO score but who might be expected to be positive by the clinical trials assay the clinical implication is that these are patients who could potentially benefit from Herceptin and would not be selected for therapy. The further implications are that the magnitude of the expected treatment benefit from Herceptin is potentially large, particularly in patients who turn out to be high level expressors. The good news here with this particular cut point is that relative few appropriate patients would be missed by this cut point.

Next slide?

This slide shows the same extrapolated concordance on the four-by-four matrix. I simply colored now a different cut point, and this was the cut point suggested, I believe to the Committee and highlighted after CDRH review.

This is a cut point that considers two plus as negative and three plus as positive, and you can see the concordant cases in color.

Again, this is the expected population as you see here, the 58 percent negatives and 23 percent positives. I want to draw your attention here to this cell right here, and I don't mean to just pick one cell out of the 16, but we regard this as important. I showed you in the beginning, and now I am speaking as the sponsor of Herceptin, we showed you in the beginning of this presentation that the magnitude of the benefit among patients who are positive at the three plus level is relatively large, and that was commented on explicitly at ODAC.

These 15 patients here are three plus by the clinical trials assay and this cut point would exclude them from treatment. They would not be offered treatment on the basis of this test with this cut point applied. To us as the sponsor of Herceptin we think this would be an unfortunate consequence of this approach to the diagnostic.

Next slide, please?

To summarize, and I showed you this slide initially, but I want to come back to it now in making a proposal for a revised intended use statement, the pivotal studies showed overall benefit in a population of patients who were both two plus and three plus with regard to HER2

over expression using the clinical trials assay.

The exploratory subgroup analysis did suggest that the magnitude of the treatment benefit varied by degree of HER2 over expression, and we are comfortable with that. That actually speaks perhaps to mechanism based activity. So we can at least rationalize the result. It is an exploratory subgroup and those are always dangerous.

In any event because of that information and that is information that we expect to share with prescribing physicians in our package insert I think it is highly likely that that is the direction that the Herceptin discussions will go. We believe that physician and patient analysis of the risks and benefits of Herceptin will include an assessment of the level of HER2 over expression.

Next slide?

We, therefore, propose a slightly different intended use, and we offer this data in support of it. I am now going back to the original concordance on the 50/50 population. So, this is not extrapolated data. I am simply looking at the difference between the clinical trial staining intensity score and the DAKO staining intensity score and plotting it as a frequency histogram. Along the X axis then is the degree of staining mismatch and the direction is shown here. It is the CPA score minus the DAKO score. So, the slightly higher sensitivity of the DAKO test

is indicated here by the higher column in the minus one versus the one.

This is an approximately bell-shaped distribution, and it is centered on zero which is perfect agreement. Fifty-two percent of the cases in that large concordance study of 548 samples showed perfect agreement. In addition as shown in the white shading 89 percent of the cases agreed to within one level of staining intensity and only shown out here in the yellow were disagreements of greater than one level.

We believe that this data speaks to comparative performance characteristics of these two assays. This is not random noise. These assays are both looking at HER2 expression on the cell surface, and we think this is compelling evidence of that fact.

Next slide?

In summary, we believe that the DAKO assay exhibits performance characteristics comparable to the clinical trials assay, and we do believe that it may be substituted for that assay as an aid in the assessment of patients in whom Herceptin treatment is being considered.

Next slide?

We make this proposal which is different from what we proposed in the PMA just to be clear. We believe based on the differential efficacy suggested by the subgroup analysis

and other considerations which were raised at ODAC that it is reasonable to report the HER2 staining intensity score in full, zero, one plus, two plus and three plus to allow accurate assessment of the risks and benefits of Herceptin treatment in a given patient, and we imagine that the Herceptin package insert will contain significant language about our experience with these levels in our pivotal trials, and we think that that is the kind of information that prescribing physicians and patients will want to use in making informed treatment decisions.

This is just to reprise the slides I showed initially. There is treatment benefit again in the two plus cases. I have just extracted the Taxol group. Some patients we imagine would want to be treated who were two plus, and this might be the reason why. I am not saying that we would recommend it but this is the reason why we think the staining intensity score is valuable.

Next slide?

And that is the difference between the two plus and the three plus.

Next slide?

So, again, I will just leave this up. This is the revised intended use statement that we made at the beginning, and I will just read the last paragraph. We believe that the DAKO anti-HER2 IHC system should be

indicated as an aid in the assessment of patients for whom Herceptin treatment is being considered.

Thank you.

DR. O'LEARY: Next, because we have had several Panel members come in since the time we did original introductions, I would like to ask those who have joined us over the course of this morning's presentation to introduce themselves beginning with Dr. Hortin.

DR. HORTIN: Good morning. I am Glen Hortin. I am at the NIH Clinical Center.

DR. O'LEARY: Dr. Kemeny?

DR. KEMENY: I am Mary Kemeny. I am the Chief of Surgical Oncology at North Shore.

DR. O'LEARY: Thank you. Ms. Rosenthal?

MS. ROSENTHAL: I am Ellen Rosenthal. I am an engineer and a free-lance writer.

DR. O'LEARY: Did I catch everybody? Okay, thank you.

Next, I would like to give the Panel members an opportunity to query the manufacturer on the presentation, either DAKO or Genentech on their presentations and items that they might wish to have clarified.

DR. MILLER: First, I would like to thank the sponsors for doing the revisions that I suspect they frantically did yesterday in regard to the ODAC --

DR. O'LEARY: Maybe even today.

DR. MILLER: Right. I appreciate that because it really did help me clarify. I think those are important changes that you have made. Just back to why you cannot get the data from Genentech study, just to explain a little more. If in fact, you screened 7000 patients, I assume they signed consent to be screened. Why then could you not go back and get those? If they signed consent to be screened and there are 7000 positive and negative patients, and you got blocks, and you cut some off and sent the blocks back, why could you not then go back and ask for those blocks again, and if it was, could that help us more than the concordance trial you showed?

It is a very nice study, but the primary data on the patients that are screened, I think would be preferable.

DR. COHEN: This fellow who is shadowing me here is Steve Sack who conducted the clinical trial for Genentech. Let me just answer one part of your question by pointing out the problems inherent in the time line that we were dealing with?

We looked at the power of this, and we would have needed 600 samples to provide a compelling estimate of concordance, and that would have involved 600 negative samples as well, and that would have taken a great deal of logistical support to do that.

In addition with regard to the consent form that patients signed, I believe they signed for consent for entry, for assessing entry into the study. That was done, and they were found to be ineligible. Following that I am not sure that their consent, and we actually got a legal opinion to suggest that their consent would not have allowed them to have their samples reanalyzed for some other purpose, and I think that is very important. Is that what you were going to say, Steve?

DR. MILLER: That is the answer. The consent did not allow it. You couldn't do it. It is a shame. I am surprised that any consent didn't include some flexibility in the ability to use this.

DR. COHEN: My lawyer will note that.

DR. SACK: The consent was signed by the patient after the screening was performed upon the entry into the study. So, that was the consent for the study. There was no consent for the screening.

DR. MILLER: How could you do screening without consent? If the patients' samples got sent to a commercial laboratory with the intent of doing a clinical study, and that was a research test you were doing as part of the clinical study how could they not sign consent to have their samples screened?

DR. SACK: At this point, again, there was no

intent or compelling, there was nothing to compel the patient to enter the study. So, it was to provide information to the patient and the physician as they assessed options, the study being one option, many other studies; these patients considered other studies and many of the patients, in fact, were not positive and therefore entered into either other studies or other treatment protocols.

DR. MILLER: I thought anytime you did something to a patient sample with an experimental procedure that was not directly intended for clinical outcome that that required consent.

DR. SACK: Maybe the FDA can clarify this, but I do not think that is the case.

DR. KEMENY: Patients can ask for HER2/neu on their samples.

DR. MILLER: But not at a central reference lab that was not done at your clinical lab.

DR. KEMENY: In most clinical centers the specimens can be sent out to labs.

DR. O'LEARY: It is still not entirely obvious given the clinical centers that are likely involved because many clinical centers have informed consent which does allow the use of the tissue for research purposes. Now, I realize that in the current environment we may be rapidly changing

here, but at the time that this was done, I am a little bit surprised by that as well.

DR. COHEN: It was a logistical problem, and I think there were legitimate issues with regard to the patients who were not enrolled in our trials. Their obligations to us, I think, were not entirely clear.

DR. O'LEARY: Sure, we understand that.

Yes, Dr. Ladoulis?

DR. LADOULIS: I wanted to move to another question having to do with the concordance data. It seemed that through most of the presentation there seemed to be a strong argument that a valid concordance would be if a two by two analysis were to be done showing concordance between zero, one plus and the ICTA at zero, one plus by DAKO being negative, and yet at the conclusion you have summarized that the preferred scoring that you would claim is that the full scale be adopted for one plus, two plus and three plus. I think the problem with that is on Pages 94 and 99 in your presentation here when you compare them is the numbers of patients who would be classified by DAKO as positive and yet actually are negative by the CTA or would be negative is substantial. That is like 275 patients out of 318 negative patients by CTA would be classified or I am sorry, 117 out of 318 would be classified as positive if you adopted one plus, and then might be qualified for Herceptin treatment.

Out of those 318 negative CTA, maybe it is not 117, maybe it is 133 out of -- we will put it at 50 okay. I think it is 99, that is I guess Page 95 or whatever, but you proposed that --

DR. COHEN: Just for me could you please clarify which patients you are viewing as discrepant here?

DR. LADOULIS: I highlighted it. You proposed that instead of just the three plus DAKO positive patients being determined to be positive or qualified as positive that you also, include patients who are going to be one plus, two plus as well.

DR. COHEN: No, we did not propose to include one plus patients as positive or appropriate for treatment. Let me be very clear on what we are proposing? We had proposed a cut point, and when we proposed the cut point it was consistent with what we had done in our clinical trials, and those trials are positive for the group in aggregate as Sue Jerian I am sure will show you but there is, in fact, and it was the subject of much discussion at ODAC and legitimately so, we believe, that there is a modifier of the treatment effect based on the level of over expression.

What that means, I believe to a practicing physician and if you will permit me I may call on one, I think what it means is that there is an issue about the difference between two plus and three plus in the mind of

both the physician and patient. We don't in providing the one plus information seek to compel anyone or even suggest to patients that they be treated as one plus. We acknowledge that we have no information on such patients and would not represent that they ought to be treated.

The question is one of information to patients, and it is clear that with any cut point if we take it as positive or negative we are actually losing information I think. We have concordance among the whole range, and I showed you a frequency histogram to suggest that.

DR. LADOULIS: I will acknowledge that you would be losing information. The question is in terms of the safety of the use of such a device in a clinical laboratory setting widely disseminated and used. The question is what are the risks of including patients who are one plus, okay?

If the scoring is not positive or negative but just a scale then the likelihood is that patients who are one plus may be elected by a physician to receive Herceptin treatment, and the question is since the ODAC finding was that there is currently known benefit for such patients, the question is what are the risks involved in such treatment either in terms of the treatment itself or for delays or interference with any other therapy that the patients might otherwise be elected to receive. So that is the question as to why not stick to, I think one of the earlier

considerations that is a positive and negative based on the one plus, two plus rather than leaving it to a scoring system by a laboratory which is going to be maybe multiple laboratories reporting to the clinicians and leaving clinicians at risk to make a bad determination for the patient. I think that may be safer to establish the positive/negative borderline as a scoring system.

DR. O'LEARY: After a short comment what I would like to suggest on this area because I think it will be clarified by the FDA presentation by Dr. Miller's experience at the ODAC meeting and so forth that we not explore this in depth until after the FDA presentation, but if you would like to make a short response?

DR. COHEN: Just a brief comment. We are not suggesting that such patients be treated and to the extent that these assays are concordant we fully recognize that we as the sponsor of Herceptin have no experience with patients in the one plus class.

If there were perfect concordance we wouldn't be having this discussion, I don't think, but there isn't. We could imagine not reporting one plus but I think why we would want to go that way is that we do believe that there are differences between two pluses and three pluses and I do want to come back to that issue as would reflect clinical decision making. I think the one pluses may be another issue

entirely.

Dr. Davey?

DR. DAVEY: Yes, I just wanted to find out. I think there is probably not, but do you have any data on molecular FISH characterization and Herceptin response; is there anything at all? I know it was part of your trial but in terms of other ways of over expressing like the FISH and how patients would respond to Herceptin, that is one question, and the other one is I can see, you know, with laboratories we would probably want to try to do some proficiency testing or some sort of interlaboratory comparison. How do the slides, if you do unstained slides, how long do they last because there was something, I think, in our data about deterioration?

DR. COHEN: I will let Mads take the second question in a minute. Your question about FISH. FISH was never used in our pivotal studies as a criterion for treatment. We have no information. I think the molecular studies that we presented suggest that these tests do agree to some extent, and I think there will be overlap and populations identified using standard FISH cutoff criteria, and I have little doubt that in the real world there will soon be a rapid experience gained with FISH. There was a letter to that effect read at ODAC, and we would certainly not stand in the way of that kind of scientific enterprise.

I think that is important information. There are FISH tests available, but the short answer to your question is we as the sponsor of Herceptin have no direct information regarding FISH, and Mads can answer the other one.

DR. ROEPKE: Concerning the stability of the HER2 epitope protein in sectioned material we see some deterioration over time when kept especially at room temperature which is the normal thing to do in many pathology labs. So, we do not advise that tissue sections are kept over time, and if they are to be kept for more than -- are not stained within the same week they are cut, we should advise that they be kept at 2 degrees.

DR. FELIX: For how long?

DR. ROEPKE: We have stability on the controls that we use for 6 months at 2 to 8 degrees after sectioning of the material.

DR. FELIX: So, your kits obviously must be kept refrigerated and your control slides must be kept refrigerated?

DR. ROEPKE: That is right.

DR. FELIX: But you have no data on the length of the agent viability in slides kept at room temperature?

DR. ROEPKE: No, that is right. We conducted a study that is included in the PMA that is the archival study, and we detected its deterioration. Of course, we

could not use the material, and that material was kept for, I believe 6 months.

PARTICIPANT: Six weeks maybe, but you know better, 6 months?

DR. ROEPKE: Six months, at least 6 months at room temperature.

DR. FELIX: So, without data as to when that deterioration occurs, we know that it occurs sometime before 6 months. So, 6 months is unacceptable. Without any data we would have to assume that it can happen at any time between a few days and 6 months. I Mean we have done some work on antigen deterioration but without data you cannot say when this occurs, right?

DR. ROEPKE: I think it is safe to say that this tissue can at least be kept for 2 to 3 weeks at room temperature because that is what we have been doing in the past. Discovering this deterioration problem we have put our tissue bank in the cold room in order to have a better stability.

DR. DAVEY: I am sorry, 2 to 3 weeks it is or is not okay?

DR. ROEPKE: That would be okay.

DR. DAVEY: At room temperature?

DR. ROEPKE: Yes.

DR. O'LEARY: Dr. Miller?

DR. MILLER: Based on the way your clinical program is now, what are you doing with testing samples to get on to the compassionate use of Herceptin?

DR. COHEN: Initially in the compassionate use study we were still doing the same LabCorp kit. We were offering that to sites, but I think in compassionate use, and maybe Steve needs to comment on this as well, I think the -- maybe you should?

DR. SACK: The compassionate use program in place since 1996, allowed HER2 testing by any method that was chosen by the site. It turned out a high frequency of those actually I think had been sent to MPS, a central pathology laboratory which in fact has been using for some time now the same rapid DAKO antibody.

Recently the NCI has taken over the protocol, and the DAKO Corporation has offered to the NCI these DAKO kits for use at sites now I think 40 sites around the country. So, some of the patients now are currently being enrolled.

DR. MILLER: Do you have any data from that to at least preliminarily look at responses in two plus and three plus using this kit?

DR. SACK: No, we don't.

DR. O'LEARY: Do you have any idea when you might have such data?

DR. COHEN: The total scope of the expanded access

program is only a few hundred patients, and as Steve said, few of them have been enrolled using the DAKO test. Depending on our discussions with FDA regarding Herceptin it is likely that a substantial experience may not come until after product approval, but then it could be quite quickly gained.

DR. O'LEARY: Dr. Felix?

DR. FELIX: I want to go back to your selection criteria for your study. As you mentioned you picked an equal number of positive and negatives, and that, of course, was commendable, and your negatives you said that you randomly chose. Is that why most of the cases are at zero plus? More than two-thirds of your negative cases were zero plus. Now, in a way that may bias the correlation of it because non-expression is less likely to be misinterpreted as a positive than a weak expression. Was there any given thought given to actually equating zero pluses and one pluses?

DR. COHEN: Yes. There is a couple of answers to that. The first question was how was it done. So, it was done by the Balooly(?) variable and the SASJAM-P(?) program simply pulling them out of a hat. So, the ratio between zero and one plus should reflect the incident ratio between zeroes and ones that we observed in the whole population. We didn't specifically select for zeroes or against ones. The

ratio between zeroes and ones is the same as the ratio between zeroes and ones among the whole 1200. To your specific question, did we give thought to balancing since there were 25 percent perhaps at each level, if you take the power calculation that we gave you which was the 600 samples needed and you imagine that you could see as few as 9 or 10 percent of cases in a given level it would have increased the number of cases we need to screen substantially from the 1200 that we actually needed. We would have needed more cases, and we were concerned about that, particularly about two plus cases. If we wanted to balance for two pluses at 25 percent our own data were suggesting that that would be a relatively low population.

So, we didn't take any special pains to balance among the four groups. The result was as you saw that 40 percent of the cases were either one plus or two plus. So, were close, but we didn't do precisely what you are asking.

DR. FELIX: Now, when the randomization occurred did you go back, and did you check whether the ratio selected by the randomized procedure reflected the population from which you randomized it?

DR. COHEN: I believe we did. I am pretty sure. Let me just quickly check that?

PARTICIPANT: Actually we did it pretty blinded. We just lumped together all the positives. We had all the

positives and we counted the number of positives, lumped all the negatives and let it represent the ratio in the population and after we had chosen it we did not then go back and check and change the selection. So, we relied on initial randomization.

DR. COHEN: But we have all the listings.

PARTICIPANT: But we have all the listings, yes. We have that information.

DR. COHEN: The likelihood of a significant deviation from there with that number of patient is pretty low.

DR. FELIX: It would be interesting to actually see that.

DR. O'LEARY: Dr. Miller?

DR. MILLER: In the interlaboratory reproducibility studies should we be concerned that four out of the 10 centers that were initially asked to be part of that study were not able to perform the test correctly, considering that these studies, I assume were specially trained and were getting grants to do this, and it is not going to be your run-of-the-mill laboratory that is buying this off the shelf and saying, "Now, I am working with it." I am not a pathologist. So, is that of concern?

DR. COHEN: I am not either. Let me propose two answers for this. One is for Mads to take this question,

but, also, --

(Laughter.)

DR. COHEN: That is the easy one, and my other one is to, also, ask one of my pathology laboratory director consultants to take it, also, and just comment on the applicability to a path environment, if that is acceptable to you?

DR. O'LEARY: That is fine, please?

DR. MURRAY: We weren't supposed to bring this up at the Panel meeting, but what happened was we tried to mimic the real world. We sent the kit out with the instructions, "Here is the kit. Here are the control slides. Run testing on them. After you have done your testing send them back, and we will look at them, and then we will send you the slides randomized for the interlaboratory reproducibility." We didn't go into the laboratory and hold their hands. We didn't stand over them while they were running their run, if they suggest tweaking any of the steps, and so what you are seeing is real world.

DR. MILLER: Since you got four out of 10, did your change your instruction kit so that when you go out into the real world it will be better than 60 percent being able to do the test?

DR. MURRAY: Right. One issue in the labeling was we said that three plus made it a valid run, and we missed.

We should have said that the one plus staining has to be present. In other words, don't evaluate your staining run if you don't see staining in the one plus. So, actually after the staining was completed we went back and looked at the control slides for the one plus staining intensity and invalidated the runs that were done where zero staining was seen.

DR. FELIX: How often did that happen and why?

DR. MURRAY: You can answer that one.

DR. ROEPKE: I just want to put a comment to the one plus control line that Gretchen mentioned, and it further shows how important this control, performance control is because often labs that were indeed able to make a one plus staining of that control, those labs performed very well and the intralab reproducibility in those labs was pretty convincing.

So, what I am suggesting is that when we strengthen our kit insert to the point where the one plus control is considered very important and considered carefully by the technician, then we have a very reproducible kit and that as explained by Gretchen, that failed in certain labs and they were disqualified from the study.

DR. FELIX: What efforts were made, if any, to find out why they failed? Why wasn't that one plus

positive?

DR. MURRAY: Okay, we have two laboratories one of which put all of the slides in the water bath at the same time and turned the timer on to 40 minutes. They then repeated a dummy run of the same volume and found that it took 15 minutes for the water bath to come back to 95 degrees. So, they were only heated for 25 minutes.

One laboratory used a microwave, and we have not - there is a variability in the strength of each microwave, and we don't know what are the requirements in terms in terms of wattage and time to have an equivalent 40-minute heating from the water bath. I cannot answer what the correct microwave requirements are which is why we recommend water bath which is reproducible between laboratories, and that is four of the runs out of the 10 that were done.

DR. O'LEARY: So, if I might follow up on this, of the four methods that were initially considered, also, having steam and pressure cooker which a few laboratories do use, you have no idea of what the performance characteristics would be there? Is that a correct assumption?

DR. ROEPKE: We have tested both as you mentioned, the steamer and the pressure cooker, and we found it to be less reproducible, at least in our hands in our setting. We found the water bath to by far be the most reproducible

system if the instructions are followed.

DR. O'LEARY: And please forgive my imperfect memory but does your proposed package insert specifically suggest that these technologies not be used?

DR. ROEPKE: The proposed insert recommends strongly the use of water bath as a source of heat for pretreatment.

DR. O'LEARY: But I am trying to remember if there is a negative suggestion with regard to other possible sources of heat.

DR. ROEPKE: It does not have at the moment, but what I still want to strengthen is that our performance control, the one plus control line is the important control for the kit, and if an alternative mode of heat is used and the one plus control line performs adequately then the staining run is acceptable.

DR. O'LEARY: Sure. I understand that. The reason I asked the question is because there is a second consideration which is the laboratory looking at clinicians and patients breathing down their necks for the results and so performing something, having it fail, going back and performing it again, it is a different laboratory work flow, and that is just why I wanted to explore that issue.

DR. COHEN: There may be two things, one, not to let a sleeping dog lie, let me just answer Dr. Felix's

question? In the population, the ratio between one plus and zeroes was .453 and in this sample it was .433.

DR. FELIX: Thank you.

DR. COHEN: And maybe it would be worth getting a comment from a laboratory director on the issue that you raised. Would that be okay?

DR. O'LEARY: Sure.

DR. COHEN: Okay. So, I would like to call on Steve Anderson who is the senior technical director at LabCorp who has done over 7000 of these assays and has a great deal of experience, not only with the assay under consideration here but, also, with our clinical trials assay.

DR. O'LEARY: Okay, thank you.

DR. ANDERSON: As Bob said, I am Steve Anderson. I am senior technical director at LabCorp. I am in charge of the laboratory that performed both the clinical trials immunohistochemistry assays and, also, the DAKO assays.

Now, when we were asked to perform the DAKO assay we followed their protocol as was recommended. We didn't see fit to change the heat, the mechanism of antigen retrieval, and so, I think that is important.

The other issue with regard to the one plus control of how often you might expect it to fail, we have done hundreds of runs of this assay, and it does fail on

occasion, but I don't have the exact numbers. It is not a frequent event, at least in our hands.

DR. O'LEARY: Okay, so, when used with the water bath heat source we know this is pretty good?

DR. ANDERSON: Yes.

DR. O'LEARY: Thank you.

Yes, Dr. Felix?

DR. FELIX: I had another question somewhat unrelated, but regarding the sample size of your correlation with molecular studies, the FISH study, how did you determine sample size there? It seemed like a fairly small sample size.

DR. COHEN: How did we get to 43?

DR. FELIX: Yes.

DR. COHEN: It was a prime number.

(Laughter.)

DR. COHEN: The study was originally proposed actually and done at Genentech as a feasibility study, and then in our discussions with FDA following our September 1997 meeting we had some conversations about correlating the results of the DAKO assay with other means of determining HER2 over expression, and we meant it, and they requested it as an exploratory analysis. So, we had that study done, and we simply finished it and wrote it up and filed it with the PMA. The Press study, the 168 cases, that is a well-known

multi-block that Michael Press as you know has used many times and that is how that number was chosen. So, these were not powered to detect anything, just to simply explore whether there was a general correlation.

DR. O'LEARY: Okay, Dr. Davey?

DR. DAVEY: Again, just to make sure that I understand on the stability of the slides, if a test were you know, to be used, I am sure that a lot of smaller hospitals still wouldn't want to run it. They would probably refer it. So, at this point the recommendation would be to send the block and not unstained slides? Is that correct, for testing? That is what you did for all the testing done?

DR. COHEN: No, actually not. I mean we allowed them to do, in this setting of the concordance study, actually what was shipped were freshly cut sections.

DR. DAVEY: And those were shipped --

DR. COHEN: They were shipped in days. It was very quick.

DR. DAVEY: Was it shipped at room temperature though or --

DR. COHEN: I am not sure of that.

DR. ROEPKE: They were shipped at room temperature and tested within 2 weeks from sectioning.

DR. DAVEY: Okay.

DR. COHEN: And when there are clinical trial

studies which I think, also, may be relevant we allowed both blocks or sections to be shipped. In most cases it was blocks that were cut centrally at LabCorp but in some cases it was sections, but they were turned around very, very quickly.

DR. O'LEARY: Dr. Miller, did you have something?

DR. MILLER: I guess I was getting back to the stability because if the test is indicated for the decision of whether or not Herceptin should be used that is used at the time a patient has metastatic disease, and so, you have to be able to determine -- not every patient who has a lymph node dissection or has a primary pathology will get, in my understanding will get HER2/neu tested. I think a lot of places will want to know the information from a prognostic standpoint, but that is not really what this test is indicated for. It is indicated for at the time somebody develops metastatic disease that you go back and get the blocks, take the blocks; you cut it, and you do the test. That is my understanding. Is that correct?

DR. COHEN: That is correct. Does your question then relate to the stability of the antigen epitope within a block?

DR. MILLER: I mean it could be 10 years between the time the patient gets diagnosed with primary tumor and they have metastatic disease. Metastatic disease is not

always rebiopsied, and there are some discussions about it might be interesting to have them rebiopsied because there is a discussion that the HER2/neu expression will be higher in metastatic disease than in the primary disease, but my understanding is that you may be doing this on blocks that are 10 years old if we are lucky enough. Now, with HER2/neu over expression it usually will be much quicker, but is that going to work?

DR. COHEN: So, just to clarify, the situation that you described is the situation that we anticipate. We are in agreement. Not only that, that is actually the situation which we just described here. These are archive specimens, many of them many years old, and the same is true for our clinical trials. It is our general impression that the epitope is stable within formalin-fixed paraffin-embedded blocks, but Mads could probably comment even more directly on that.

DR. ROEPKE: The general idea is that all sorts of protein targets or antigens are stable within the paraffin block and only upon sectioning there is the deterioration. So, they are stable.

DR. O'LEARY: This seems to raise this question further. Has anybody looked at the distribution of staining intensities in a random sample from, say, 10 years ago and compared it to the distribution in the current sample to see

whether there is variation? One might expect to see antigen not stable, to see a significant change in that intensity distribution, numbers of cases by intensity.

DR. COHEN: It is a nice experiment, and we haven't done it. One thing though I would point out. I think it is at least my impression having looked at the demographics that there are confounders of the distribution with regard to tumor size, stage at presentation and --

DR. O'LEARY: I understand these are smaller tumors now that are presenting.

DR. COHEN: So, there would be some challenges in actually getting a comparable population, but it is an interesting study, and there are 9000 cases available in the CBCTR.

DR. O'LEARY: We will take a couple more questions.

Yes?

DR. DAVEY: Diane Davey, again. Just one more question. I am thinking since I do a lot of cytology we are often asked to aspirate the recurrences. Now, we try to do in general when we do ERPRs and other things we try to do a cell block, but what about smears? I mean I realize that that is not the primary focus of your -- but people are going to start using the thing lots of different ways, and I just wondered if you had any information?

DR. ROEPKE: We have tested the kit on cytopins of the controls and find that it performs as it should, giving the accurate score.

DR. DAVEY: What kind of fixation?

DR. ROEPKE: That is fixed in acetone 2 minutes, that is cytopins of the cell line controls.

DR. COHEN: Steve, also, I think had some experience.

DR. ANDERSON: Yes, and again, we have some experience with fine needle aspirates with immunohistochemistry for HER2/neu, and again as Mads has suggested by using fixation like acetone fixation you can detect HER2/neu over expression in fine needle aspirates. I don't have a global answer to that other than the fact that we know that it can be detected in fine needle aspirates.

DR. FELIX: How about alcohol fixed fine needle aspirates?

DR. ANDERSON: I would have to go back and look at our data again. I know that fixation is and can be an important parameter in making sure that the epitope is stable in a frozen or an aspirate slide, and again I would have to go back and look what our experience is but I believe it is in acetone fixed aspirates that we have seen that overexpression is detectable.

DR. O'LEARY: Thank you.

Dr. Miller?

DR. MILLER: This is a general question. Given that the HER2 expression is a very important prognostic factor for primary breast cancer as well, and I understand the reason that this test is being evaluated in this way is because there is a drug that -- why not look and do you have plans to actually look at the big picture about whether this test should be used and the prognostic significance of this test in primary breast cancer because I have a feeling that when it gets out there that people aren't going to wait until the person relapses to go back and get the blocks. People will want, will ask for it on primary tissue and it would be -- do you intend to help us and get that information?

DR. COHEN: It is a question for DAKO, and I will leave it to them, but I would make the general observation in response to your general question that having a treatment now ups the ante. Now you have a reason to do this test based on treatment, and I think that is going to provide a lot of information and a lot of teeth with regard to the prognostic information.

We believe on the basis of the molecular characterization studies that we have reported here that this test is looking at HER2 on the cell surface, and I think it is quite likely that prognostic conclusions can be

drawn. I will let Gretchen respond with regard to where DAKO would like to take it, but certainly we would support from our perspective at Genentech, we would support that thoroughly.

DR. MILLER: The best chance to cure breast cancer is early.

DR. MURRAY: Yes, we do have plans to do further testing for prognosis, and the test we are planning to run which we are getting organized now is to test the same tissues for ER, PR and HER2 and that way we will have a global picture of the breast cancer status.

In addition, it may be further down the line, we may go for other cancers, and again, that will be prognostic.

DR. O'LEARY: I have a question. Again, this is about looking at the tissues as they come out. I will preface it by saying that we get tissues in consultation. We find laboratories that seem to fix things in saline and process them on heaters, and --

(Laughter.)

DR. O'LEARY: -- so, we see that there are institutions that have blocks that we cannot stain with anything, and what -- are you proposing, any particular internal tissue control for determining that the tissue itself is suitable for immunohistochemical staining because

garden variety breast cancer unless it is being looked at for ERPR immunohistochemically probably won't be assessed with a panel. It is not generally required for the diagnosis.

DR. MURRAY: The current package insert which is obviously going to be amended does contain information that a positive tissue control should be run. However, running the tissue, using another marker, for example, to determine how well it has been fixed is not the current practice, and we don't have recommendations for that at this time.

DR. O'LEARY: Thank you.

Dr. Ladoulis?

DR. LADOULIS: This is Charles Ladoulis. Can I amplify and extend the remark that was just made by Dr. O'Leary? For example, some teaching institutions in the past used not formalin but Buen's(?) or other fixatives for even their routine histopathology because it at the time seemed to be an appropriate thing to do.

Now, for some purposes such as flow cytometry some of these fixatives render the specimen useless. What evidence do you have that other fixatives that might be in common use in some hospitals different from neutral buffered formalin might invalidate this DAKO assay?

DR. ROEPKE: We have tested the kit on formalin of course, and, also, on Buen's, on the Buen fixed cell lines

and find it to work as good as formalin on formalin-fixed material.

DR. MURRAY: And then we will request Steve Anderson back up here again because he has received all of the tissues in whatever form they arrived.

DR. ANDERSON: The tissues in addition to neutral buffered formalin, we have looked at Buen's fixed tissue and can detect virtually no over expression in Buen's fixed tissue, and we have looked at tissues that are fixed in a non-formalin fixative like Prefer(?) and that, also, is, HER2/neu over expression is detectable. However, as a reference lab you need to know that up front because you need to treat those samples differently than you would one fixed in neutral buffered formalin.

If the submitting clinician or account does not provide that information, that, in fact can compromise the results, but you can detect it, but that information needs to be known.

DR. LADOULIS: I am sorry, then the package insert should allow for some specification for laboratories who are performing this test that they should adopt a different procedure.

DR. ANDERSON: The experience that we have is with the clinical trial assay because we did receive samples in a couple of alternative fixatives.

DR. O'LEARY: Dr. Felix?

DR. FELIX: Now, you have just mentioned that if you knew the alternate fixative, you would modify what you did?

DR. ANDERSON: You have to treat the sample differently. For example, if a tissue is not fixed in formalin, if it is fixed in Prefer, it will not be cross linked. So, protease digestion is not necessarily appropriate for that.

DR. FELIX: But in the particular assay that we are evaluating today there is no protease digestion at all.

DR. ANDERSON: And all I am doing is addressing it from a clinical trials assay perspective because that is where our experience is.

DR. FELIX: What we are asking is whether the current assay has evaluated other fixatives and I mean all you have to do is to say that we can only use formalin fixed until other data come out.

DR. COHEN: We understand. I think Mads commented on the breadth and depth of our experience with cell lines. That is right.

DR. O'LEARY: Dr. Hortin?

DR. HORTIN: The assay uses a polyclonal antibody. Do you have adequate stocks of that to use the same antibody for a period of time or will the performance change when the

plus with either the 4D5 or the CB-11 antibodies.

Next slide, please?

This is a breakdown of enrollment on the study. The first two rows provide you with the randomization arms. The last four rows, the breakdown is by treatment, chemotherapy treatment and randomization. So you can see that about 40 percent of the patients received Taxol, and there is good distribution between the subgroups.

Next slide, please?

The only mention I am going to make about demographics in the pivotal study is that there is an imbalance if you compare the patients who received anthracycline therapy versus those who received Taxane(?) therapy, and what we see is an increase in the number of baseline poor prognostic factors and an increase in the amount of prior therapy. There is almost double the number who had positive nodes who received Taxol, a higher incidence of mastectomy, double the number who had adjuvant chemotherapy and there were no patients who had transplant in the AC group and AC Herceptin groups compared to 18 percent in the Taxol-Herceptin groups. So, this is an important factor to keep in mind when looking at the results of the efficacy data because the Taxol patients were expected to do worse as compared to the AC patients.

Next slide, please?

rabbit dies and you have to get a new one or kind of how frequent will that be? Do you have a stock that will last a few years or will the performance change fairly frequently?

DR. COHEN: Mads can answer that.

DR. ROEPKE: We have at the moment a stock that will last for more than 15 years' production, and we are, of course, with new HER2 rabbits, with ongoing immunizations. So, we are working on the next, but we have 15 years of the current batch.

DR. HORTIN: Also, I was wondering, is there a specific procedure for high-altitude use? I am not sure that the water bath procedures would get to the same temperature at say, Denver or Colorado Springs.

DR. ROEPKE: I think that should be taken care of by running controls prior to going along. If you have extreme scenarios like that running kit control, performance controls and timing the heat-induced epitope retrieval to adequately stain the one plus control, so we have a performance that we prescribe.

DR. HORTIN: Are your controls paraffin embedded?

DR. ROEPKE: Yes, they are.

DR. O'LEARY: I would like to ask if Dr. Floyd or Ms. Rosenthal have any questions.

MS. ROSENTHAL: Actually I had two pages of questions but they were really very well covered in the

presentation. My greatest concern had to do with the reproducibility and within the laboratory and let me just find the reference, okay?

If you want to go, I will come back to this.

DR. O'LEARY: Dr. Floyd?

DR. FLOYD: Al Floyd. I would like to make a couple of comments on a number of remarks that have been made about techniques, technology. We are dealing with some very practical matters here. The fixation issue is a very critical one as Dr. O'Leary understands. We have had this discussion many times in the past, but the issue of fixation is being complicated by a number of other factors at the moment.

The comment that is made about institutions where things may be put in saline and cooked has been a significant issue discussed in the histotechnology community recently simply because inspectors from various certifying agencies have gone in and scared the pants off of surgeons and operating room personnel and insisted that they never have formalin anywhere around.

The issues about the way in which one fixes tissue are important as has been stressed here primarily because of the way the fixatives act. The cross-linking fixatives which is probably a misnomer even though every textbook says so, formalin is a two-stage fixative. It causes cross links

after 8 hours of exposure. It only happens in routine hospital environment over the weekends and the routine activity you get an additive fixative effect.

A fixative such as the preferred fixative that is mentioned is glyoxal based. It is, also, an additive fixative, but it does change the dynamics of epitope retrieval and the heating process.

It is important that any institution understand all of this because otherwise your material is not reliable, and any institution doing this has to validate their internal procedures to be in compliance with HCFA and CAP and all of their inspections.

One other comment I would like to make is about the degree of aging of epitopes on cut sections. This has been known for a long time, for a very specific number of epitopes. We know that some of them are very, very susceptible. There are two factors involved that have been clearly identified, not necessarily for HER2/neu but for a variety of other epitopes. One of those is the mechanism by which the laboratory puts the section onto the slide. Excess heat at that time destroys certain epitopes. Others are time related and appear to be some type of oxidation that takes place on the surface of the slide itself.

This can be prevented by putting a barrier, an air barrier over the slide, and in fact for old blocks, old

blocks that have been available for many years can be demonstrated to be quite active for most epitopes as long as you get away from an unsealed surface and so as a point of managing one's blocks, one should always seal surfaces of blocks. We used to do it in the old days, but in modern practice people don't

Thank you.

DR. O'LEARY: Thanks and back to Ms. Rosenthal?

MS. ROSENTHAL: I am referring to the PMA supplement to Page 3 where in answer to a question you said that in the interlaboratory reproducibility study report which has been amended in this PMA, in the amended report site 3 data were omitted from the statistical analysis since both staining procedures and staining runs failed. The manual staining data from UCLA was, also, omitted. Site 4 remains a protocol validation. So those data are omitted, and I think Dr. Miller referred to anxious patients breathing down your neck.

DR. O'LEARY: That was me.

MS. ROSENTHAL: That was you? Okay. I really have grave concerns about how this is going to look out in the public arena, you know, and if this is portentous of the kind of results if a laboratory might not violate and say, "Well, let us go on with this because we need results." How are you going to enforce restrictions?

DR. COHEN: I think this is an issue for DAKO and not for me, and I will let them comment. I do think just as a general principle as a practicing physician myself I think to some extent we have to trust people to do the right thing. The situation that you just described is wrong. It is fundamentally wrong, and I would hope that that doesn't occur.

DR. O'LEARY: This will be the last comment before break.

DR. MURRAY: I am actually going to ask one of the other users of our kit to come up and address this, also. What we did in the reproducibility is this is basically the first run, and they were given only one run to do their staining.

If we had had more sections off of these tissues, and they said themselves that they invalidated the run we sent them back which would be normal laboratory practice. Then we would be able to have results from these laboratories, and in fact, we are going to have UCLA do their manual staining because we need three laboratories doing manual and three laboratories doing automated.

Dr. Susan Gammond is here from Impath(?) who did one of the sites, and I would like her to speak to the issue of what happens with the staining run because they actually repeated half of their automated staining because the run

failed.

Dr. Gammond?

DR. GAMMOND: Thank you, Susan Gammond, Vice President of Product Development for Impath. We were one of the participants of the laboratories involved in the intra and interlaboratory reproducibility studies by DAKO.

We ran three automated staining procedures with the kit. The second procedure came out with a result which was described earlier which was that the one plus controls were not observed. As a result of that as we would do with any control that doesn't work we dropped that run. We said, "This didn't work," and we ran the whole procedure again, and it worked fine.

So, I think the bottom line is that any good lab running this when you have controls that don't work, you don't accept the results, and that is just part of performing any scientific or technical experiment.

DR. MURRAY: Did that answer your question?

MS. ROSENTHAL: Yes, I guess the answer would be in the labeling to be very specific and strong about the way this kit is used.

DR. O'LEARY: Okay, for the moment because we need to get Panel members to get their lunch orders in and others that are going to do things of that sort, I am going to go ahead and stop. We will begin again at precisely quarter

of. The standard reference clock is on the wall back there.

(Brief recess.)

DR. O'LEARY: Thank you very much. Now, we will begin the FDA presentation which will start with Dr. Susan Jerian.

DR. JERIAN: I am happy to be here. I am going to give you an abbreviated version of the data that was presented 2 days ago at the Herceptin meeting with the Oncology Drugs Advisory Committee, and I am, also, going to provide you with the outcome of their discussion and their votes from that meeting.

Next slide, please?

The two clinical studies which were the focus of the license application for the therapeutic Herceptin were the 649 and 648 studies. Six forty-nine was a Phase II open label study of Herceptin as a single agent and enrolled 222 women.

Six forty-eight was a Phase III randomized study, open label comparing chemotherapy alone to chemotherapy with the addition of Herceptin, and that enrolled 469 patients.

There were additional studies submitted to the BLA but they are not presented with this data at this time.

Next slide, please?

The Phase II study design, as I mentioned was single arm Herceptin alone. This was a multi-center study

at 54 sites internationally with a target enrollment of 200.

Patients were to have had metastatic breast cancer with measurable disease and by immunohistochemistry staining using the clinical trial assay, not the DAKO kit were to be two plus or three plus and they must have progressed after having received one or two prior chemotherapy regimens for their metastatic disease.

Next slide, please?

The primary end point of the Phase II study was overall response rate. This was defined as the sum of the complete and partial responses sustained for at least 4 weeks and verified by the Response Evaluation Committee which I have abbreviated as REC.

The secondary end points were median duration of response, median time to progression, time to treatment failure and survival.

Next slide, please?

Just for your information, the Response Evaluation Committee was established as a separate institution with a separate charter. It had clear guidelines, and in my estimation the Committee actually followed these guidelines quite well.

Patients were evaluated by the REC at the time that they progressed, and part of the reason for having this Committee is that investigator tumor measurements were not

collected, and it was an open label study, and the end point of time to progression can be a somewhat subjective assessment.

There were some limitations in the charter in how it dealt with specific elements of response evaluation and in fact what the FDA did is go over all of the case report forms from all of the patients to evaluate patients, keeping these limitations in mind.

Next slide, please?

The results of the evaluation for response rate in the single arm study demonstrate that there was a 14 percent complete and partial response rate with a median duration of response of 9 months. Three percent of these patients achieved CR, and the median duration has not been reached yet because follow-up is not complete. Not all those patients have progressed to this time point.

Next slide, please?

Here this slide just provides you with an idea of the distribution of the duration of response and in the CR row I have listed the individual links of duration and the ones with the asterisks are those patients who have not yet progressed.

Next slide, please?

Median time to progression was 3.1 months and the median survival was 12.8 months.

Next slide, please?

This is the Kaplan-Myer(?) survival plot for patients treated with Herceptin alone. The aqua curves are the 95 percent confidence intervals.

Next slide, please?

So, in summary of the Phase II efficacy data, the overall response rate is 14 percent, median duration of response 9 months and median survival 12.8 months.

Next slide, please?

The pivotal study or Phase III study was the two-armed design Herceptin plus chemo versus chemo alone, and that is how the randomization was conducted. Patients could have received either anthracycline plus cyclophosphamide chemotherapy or paclitaxel which is, also, called Taxol.

If they had prior anthracycline therapy in the adjuvant setting they received paclitaxel. Again, this was an international study and the target enrollment was 450.

The stratification was by geographic region, metastatic site and prior anthracycline therapy.

Next slide, please?

These patients were to have had metastatic breast cancer with measurable disease, be two or three plus by immunohistochemistry, have had no prior chemotherapy for their metastatic disease, so a slightly different population than the single agent study, the Phase II study that I just

showed you, and the entry criteria for laboratories and other clinical assessments was sort of all included in a general statement that patients were to be suitable candidates for receiving concomitant cytotoxic chemotherapy as evidenced by screening, lab assessments, hematologic, renal, hepatic and metabolic functions. The protocol was amended part way through to eliminate specific cut-off points.

Next slide, please?

Let me just go back a minute? Part of the reason for doing that is that enrollment was very slow on these studies. Then it was felt that if the investigator had a little more discretion, he could enroll patients in an easier fashion.

The primary end point of the study was time to tumor progression. Secondary end points were overall response rate, duration of response, time to treatment, failure of survival and quality of life.

Next slide, please?

As I mentioned, the Phase III study protocol was amended a couple of times during the conduct of the study. I want to point out for you the element listed as IHC or immunohistochemistry. Initially samples were screened with only the 4D5 antibody. Subsequent to the amendment they could have been positive, positive being two plus or three

The primary end point as I mentioned was median time to progression. What you see here is the Kaplan-Myer plot. The green is the control or the chemotherapy alone arm, and the yellow is the Herceptin plus chemotherapy arm, and these are for the two randomized groups, and as you can see the curves separate early on. They remain separate throughout, and they are significantly different, with a P value of less than .001.

If we then break out on the next slide the patients who were only treated with anthracycline with or without Herceptin we see a similar curve. Again, it separates early and remains separate throughout with good follow-up and P value remains less than .001.

Next slide, please?

The effect appears even stronger when you separate out the patients treated with paclitaxel and again the same basic pattern and the P value is less than .001.

Next slide, please?

The specific numbers for median time to progression are presented on the next two slides. For the two randomized groups the Herceptin plus chemotherapy arm had a median time of 7.3 months which has improved over the chemotherapy alone arm where it was 4.5 months.

Next slide, please?

Then if we look at the subgroups we continue to

see improvement within the subgroups although it is more impressive in the Taxol subgroup, 6.7 versus 2.5 as compared to the AC subgroups, 8.1 versus 6.0.

Next slide, please?

The secondary efficacy end points were overall response rate, duration of response and I am not going to deal with time to treatment failure but I will look at survival.

Next slide, please?

The overall response rate in the two randomized arms was improved by the addition of Herceptin to chemotherapy and you see the results here of 43 percent compared to 28 percent with chemotherapy alone. The CR rate was fairly comparable, and the P value on the bottom applies to the overall response rate here.

Next slide, please?

If we look at the subgroups for response rate, again, we see a bigger effect in the Taxol subgroups, 36 versus 15 percent, and then compare that to AC 48 percent versus 38 percent, but again, the direction is in each case in favor of the addition of Herceptin to chemotherapy. You can, also, see that the CR rate remains comparable throughout.

Next slide, please?

Looking at median duration of response there was,

also, an improvement there, 9.3 months compared to 5.9 months.

Next slide, please?

And looking at survival overall again this is a Kaplan-Myer plot of the proportion alive over time. The data after 1 year are very immature as you can see by all the tick marks here and all the censoring that went on and so it is difficult to comment on how that part of the curve will appear in the future as time goes on.

I think we can say that at 1 year there appears to be some improvement in the short-term survival although if you do go by this curve the way it appears now the median survival is the same, and the difference at 1 year is significant.

Next slide, please?

So, in summary of the Phase III results for efficacy there is an improvement in median time to progression when compared to control such that the Herceptin plus chemotherapy arm adds another 2.8 months. AC plus Herceptin, now, if you separate the subgroups adds 2.1 months, and Taxol-Herceptin 4.2 months, the effect being bigger in the Taxol subgroup.

Next slide, please?

The Herceptin plus chemotherapy arm had a higher response rate, 43 percent versus 28 percent and longer

duration of response, 9.3 versus 5.9 months and for the subgroups the improvement in response rate was significant for the Taxol-Herceptin group but not for the AC-Herceptin group.

Next slide, please?

The ability to make conclusions about the median overall survival is limited because the data are not mature at this time. The 1-year overall survival is improved in the Herceptin arm both overall and in the subgroups.

Next slide, please?

Now, I am going to take the efficacy data and divide it up by two plus and three plus over expression using the clinical trial assay.

Next slide?

As you already know the 4D5 is the parent antibody for Herceptin and it binds to the extracellular domain. CB-11 binds to the intracellular domain and the polyclonal used in the test kit filed in the PMA binds to the intracellular domain and you are more familiar probably than I am even with the method of assessment.

Next slide, please?

If we look at the distribution, the incidence of patients who were two plus and three plus of all those who were enrolled on these two studies we see a great deal of consistency such that about one-quarter of the patients are

going to be two plus and three-quarters were three plus.

Next slide, please?

Now, if we go back to the Phase II study where Herceptin was used as a single agent in patients who were treated for their metastatic disease with prior therapy so fairly refractory patients we see that two of the 50 patients who were two plus were responders.

Of the three plus patients the response rate was higher, 17 percent compared to that 4 percent.

Next slide, please?

Then let us look at response rate from the Phase III pivotal study. If we look at two plus patients and compare the two randomized arms there is an equivalent response rate, 32 percent versus 33 percent.

If we look at the three plus patients there is a significant increase in response rate for three plus when Herceptin is added to chemotherapy, 47 percent versus 27 percent.

Next slide, please?

And here I have broken out the subgroups for you for response rates. You can see if we look first at the AC subgroups with two plus the response rate for AC-Herceptin was 40 percent and for AC was 40 percent. If we then take the AC patients and look at the three plus there is a significant improvement. Adding Herceptin the response rate

is 50 percent, AC alone 36 percent.

Now, if we look at the Taxol subgroups who were two plus it is 21 percent versus 16 percent. This difference is not significant, and if we look at the three plus patients 41 percent versus 14 percent; this is significant.

Next slide, please?

Now, let us look at median time to progression separating out two plus versus three plus, and I have included the randomized arms. I provided you with that data here. The sponsor provided you with the Taxol and the AC subgroup data.

Actually when you do divide it out the subgroup the numbers get a little bit small, and it is difficult to do the estimates but here you can see the curves are superimposable for patients in the two arms and the P value is .56.

Next slide, please?

Then when we look at the three plus patients this is very reminiscent of the curves I showed you earlier, and the P value here is .001. If you take the difference between these two curves and the difference between the prior two curves that is significant, and is evidence of an interaction.

Next slide, please?

The same assessment but this time looking at

survival. These are the two plus patients.

Next slide, please?

These are the three plus patients.

Next slide, please?

So, in summary of the data related to immunohistochemistry there is a higher response rate among the two plus patients as compared to two plus patients treated with Herceptin alone as second or third line therapy.

Patients with tumors scored as three plus had significantly improved response rates when Herceptin was added to chemotherapy compared to patients with tumors scored as two plus.

Next slide, please?

The addition of Herceptin to chemotherapy significantly improved the time to progression and survival among three plus patients. That should be short-term survival. The addition of Herceptin to chemotherapy did not improve time to progression or survival for two plus patients, and there is a significant interaction between the level of over expression and the effect of Herceptin on time to progression.

Now, I am going to go through some of the safety data, and I think that will get to a question that was asked earlier about what risks the patient may experience with the

addition of Herceptin to chemotherapy or the use of Herceptin alone.

Next slide, please?

Often with antibody therapies we see what is called infusional toxicity and we saw that in this case here. Nearly one-half of patients experienced this, and it primarily occurred with the first infusion, and this is in both studies the Phase II and Phase III, and the symptom complex consists of chills, fever, pain, occasionally pain at the tumor site, asthenia, nausea, vomiting and headache, and that is listed in order of frequency. Rarely hypotension will occur, and these symptoms were self-limited and treated easily with standard medication.

Next slide, please?

The most concerning toxicity that we saw increased in the Herceptin groups was the cardiotoxicity, and I am going to dwell on this for just a moment. The nature of this toxicity manifested as congestive heart failure and what we did and what the sponsor did was to evaluate patients based on the New York Heart Association classification system which is a one-to-four system where three and four patients are the most severe, four being quite limited in activity and symptomatic at rest.

Class I and II patients were less severe, Class I patients being asymptomatic. What this slide shows you is

the incidence of cardiotoxicity in each of the subgroups in the pivotal study and in the last column here. These are the results from the single arm Phase II study.

The black bar is the total incidence of patients who had Class III and IV congestive heart failure. Added on top of that is the red bar, those who were Class I and II and the total is the incidence in that arm.

So, for the AC-Herceptin arm it was 28 percent. For AC alone it was 7 percent. For Taxol-Herceptin it was 11 percent. For Taxol alone it was 1 percent and for Herceptin alone it was 7 percent.

Let me just comment that the Herceptin patients again were, Herceptin alone patients were sicker patients and had other problems going on. I don't think you can necessarily directly compare this group to these groups, but this is the incidence and how it breaks down.

Also, as you can see from here the percentage of patients with the very severe events were really much higher in the AC-Herceptin arm and some of these patients actually required dopamine, dobutamine and had severe complications.

There were two deaths in the AC-Herceptin arm and two deaths in the AC arm that were completely or in part due to the cardiac toxicity.

Next slide, please?

I am just going to briefly mention that other

adverse events were increased in the Herceptin arms of the pivotal study. That is when you add Herceptin to the chemotherapy, either AC or Taxol you see increases in particular adverse events of concern, particularly leukopenia, anemia, diarrhea which actually was almost doubled, abdominal pain and infections were increased.

Next slide, please?

So, in summary of the safety data when Herceptin is used alone it produces an infusional toxicity, cardiac toxicity and GI toxicity. When Herceptin is added to chemotherapy, again, there is the infusional toxicity, but there were increases of cardiac, gastrointestinal, hematologic and infectious toxicities.

Next slide, please?

Now, I am going to provide you with the questions that were posed to the ODAC Committee and the results of their vote on those questions.

Next slide, please?

The first question was does Herceptin as a single agent provide net clinical benefit for patients with metastatic breast cancer when used as second or third line therapy, and the Committee unanimously voted yes, 11 to 0 that it does provide net clinical benefit when used alone.

Next slide, please?

Second question reads as follows: When compared

to Paclitaxel alone does the efficacy profile of Taxol-Herceptin provide sufficient additional clinical benefit to outweigh the increased incidence of toxicities specifically infusional, cardiac, GI, hematologic and infection?

The vote here again was unanimous, 11 yes and 0 no.

Next slide, please?

Now, going to the issue of Herceptin added to anthracycline-cyclophosphamide therapy the question reads as follows: When compared to AC alone does the efficacy of AC Herceptin provide sufficient additional clinical benefit to outweigh the increased incidence and severity of cardiotoxicity and other adverse events?

The vote here was two yes and nine no.

Next slide, please?

We then asked the Committee to discuss the question of two plus versus three plus, and actually they didn't vote on this question. There was a very broad spectrum of opinion and I am going to provide you here with what each of those opinions was.

So, not any one opinion am I trying to represent as the opinion, but there was a spectrum.

Next slide, please?

One opinion was that its use should be restricted to three plus only.

Next slide, please?

One opinion is that it should be restricted to three plus or two plus.

Next slide, please?

Another opinion was that patients should be treated who were strongly positive and that relevant data should be included in the labeling for Herceptin about the two plus, three plus data that I showed you earlier, those comparisons.

Next slide, please?

A final recommendation was that there should be a Phase IV commitment to test for direct relationship between the test kit results and clinical outcome of patients treated with Herceptin, and that concludes my presentation.

I will be happy to answer any questions you have.

DR. O'LEARY: Yes, we will start with Dr. Kemeny.

DR. KEMENY: Do you know about the two and three plus data in the Phase II study? You didn't give us any.

DR. JERIAN: Yes the response rate for the two plus patients was 4 percent and for the three plus patients was 17 percent.

DR. KEMENY: Do you have data on how many of the patients in the Phase III study, I think you said somewhere had previous adjuvant chemotherapy and did they have Adriamycin adjuvant chemotherapy?

DR. JERIAN: Yes, I have that data. I can pull it up on a slide for you, but it will take me a minute.

DR. KEMENY: Is there anything significant in that?

DR. JERIAN: No, the balance was very good. Oh, for the two plus versus three plus?

DR. KEMENY: Yes.

DR. JERIAN: No, I am sorry. I don't have that data comparing the two plus and three plus.

DR. O'LEARY: Dr. Felix?

DR. FELIX: You mentioned that two antibodies were used during the Herceptin trial and that at first only one of the two antibodies was utilized and then another one was added. I am sure the data exist but were the data analyzed as to the efficacy of the Herceptin therapy as to which antibody was used?

DR. JERIAN: Yes, we looked at that, and there was no difference.

DR. FELIX: There was no difference?

DR. JERIAN: We looked at time to progression, and we looked overall and we looked at the subgroups, and there was no difference between CB-11 and 4D5.

Now, there were only 169 patients for whom the CB-11 was done of the 469.

DR. O'LEARY: Dr. Felix?