ajh

AT

DEPARTMENT OF HEALTH AND HUMAN SERVICES

FOOD AND DRUG ADMINISTRATION

CENTER FOR DRUG EVALUATION AND RESEARCH

ONCOLOGIC DRUGS ADVISORY COMMITTEE

55TH MEETING

Volume II

Friday, December 19, 1997

8:40 a.m.

Holiday Inn Bethesda
Versailles I,II,III
8120 Wisconsin Avenue

MILLER REPORTING COMPANY, INC.
507 C Street, N.E.
Washington, D.C. 20002
(202) 546-6666

Bethesda, Maryland

C O N T E N T S

BLA Supplement 97-0501 Proleukin (aldesleukin)
Chiron Corporation

Sandra Swain, M.D., Chairperson
LT Jannette O'Neill-Gonzalez, MHS, Executive Secretary

MEMBERS
E. Carolyn Beaman, M.H.S. (Consumer Rep)
James Krook, M.D.
David H. Johnson, M.D.
Robert Ozols, M.D, Ph.D.
Derek Raghavan, M.D., Ph.D.
Victor M. Santana, M.D.
Richard M. Simon, D. Sci

GUEST EXPERT
Julie M. Vose, M.D.

PATIENT REPRESENTATIVE
Kenneth McDonough

FDA
Patricia Keegan, M.D.
Stephen Litwin, M.D.
Jay Siegel, M.D.

C O N T E N T S (Continued)

NDA 20-806 Neomark (broxuridine for injection)
NeoPharm Inc

Janice Dutcher, M.D., Chairperson
LT Jannette O'Neill-Gonzalez, MHS, Executive Secretary

MEMBERS
E. Carolyn Beaman, M.H.S. (Consumer Rep)
James Krook, M.D.
David H. Johnson, M.D
Robert Ozols, M.D., Ph.D.
Derek Raghavan, M.D., Ph.D.
Victor M. Santana, M.D.
Richard M. Simon, D. Sci
Sandra Swain, M.D.

PATIENT REPRESENTATIVE
Wilma Carroll

FDA
Julie Beitz, M.D.
Karen Johnson, M.D.
Robert DeLap, M.D., Ph.D.
Robert Justice, M.D.
Robert Temple, M.D.

P R O C E E D I N G S

Opening Remarks

DR. SWAIN: I would like to call this meeting to order. I am Dr. Sandra Swain and I will be chairing this morning's session. First, we would like to hear the conflict of interest statement.

Conflict of Interest Statement

LT O'NEILL GONZALEZ: Good morning.

The following announcement addresses the issue of conflict of interest with regard to this meeting and is made a part of the record to preclude even the appearance of such at this meeting.

Based on the submitted agenda and information provided by the participants, the Agency has determined that all reported interests in firms regulated by the Center for Drug Evaluation and Research present no potential for a conflict of interest at this meeting with the following exceptions. We would like to disclose for the record that Drs. Janice Dutcher and Kim Margolin have current and past involvement with Proleukin. Because of this involvement, Drs. Dutcher and Margolin will be excluded from participating in the committee's discussion and deliberations concerning Chiron's Proleukin (aldesleukin).

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 participants are aware of the need to exclude themselves from such involvement and their exclusion will be noted for the record.

With respect to all other participants, we ask in the interest of fairness that they address any current or previous financial involvement with any firm whose products they may wish to comment upon.

Thank you.

DR. SWAIN: Next, I would like to go around and have everyone introduce themselves starting with Dr. Litwin.

DR. LITWIN: Stephen Litwin, CBER.

DR. SIEGEL: I am Jay Siegel, Office of Therapeutics, CBER.

DR. VOSE: I am Julie Vose, University of Nebraska Medical Center.

MR. McDONOUGH: Ken McDonough, Patient Representative.

MS. BEAMAN: Carolyn Beaman, Consumer Representative.

DR. KROOK: Jim Krook from the Duluth Clinic.

DR. RAGHAVAN: Derek Raghavan, University of Southern California.

DR. OZOLS: Bob Ozols, Fox Chase in Philadelphia.

LT O'NEILL-GONZALEZ: Jannette O'Neill-Gonzalez, Executive Secretary.

DR. SWAIN: Sandra Swain, Medical Oncology, Washington, D.C.

DR. SIMON: Richard Simon, Biometric Research Branch, National Cancer Institute.

DR. JOHNSON: I am David Johnson from Vanderbilt University.

DR. SANTANA: Victor Santana from St. Jude's Children's Research Hospital.

Open Public Hearing

DR. SWAIN: We have had no requests for anyone to speak. If there is anyone in the audience that would like to make a statement, this would be the time to do it.

[No response.]

DR. SWAIN: If there is no one who would like to speak, we would like to start with the sponsor's presentation by Chiron Corporation. Mary O'Hara.

BLA Supplement 97-0501 Proleukin (aldesleukin)

Chiron Corporation

Applicant's Presentation

MS. O'HARA: Thank you, Dr. Swain.

Members of the Advisory Committee, representatives from the FDA, ladies and gentlemen: Good morning. My name is Mary O'Hara.

[Slide.]

On behalf of the Chiron Corporation, we appreciate the opportunity to present data to support a second cancer indication for Proleukin, a recombinant interleukin-2. The data that we will present today will demonstrate Proleukin's safety and effectiveness in patients with metastatic melanoma.

[Slide.]

As a brief introduction, I would like to go over the important milestones that have been achieved during the development of Proleukin.

Proleukin was cloned in 1983 and in less than a year later, it was introduced into clinical trials in patients with metastatic disease. In 1989, Proleukin became the first product approved to treat patients with metastatic renal cell cancer.

In 1992, the Biologic Response Modifiers Advisory Committee recommended that Proleukin be approved for the same indication, and on May 5th, 1992, the FDA issued a license for Proleukin to treat patients with metastatic renal cell cancer.

[Slide.]

The original product license application consisted of a primary efficacy database of 255 patients treated with Proleukin as a single agent. Proleukin was given at a dose level of 600,000 international units per kilogram every 8 hours as a short I.V. infusion for up to 14 doses. Following 9 days of rest, the cycle was repeated.

[Slide.]

In the original product license application, efficacy was defined as both response rate and duration of response. As a commitment to the FDA, Chiron agreed to follow up all patients who responded to therapy and who were alive at the last contact. Follow-up in this cohort of patients have demonstrated the durability of this response.

[Slide.]

This slide identifies data from the latest follow-up which was submitted to the FDA. Again, this is renal cell patients.

For the 20 patients who achieved a partial response, the median duration was 20 months with a range of 3 to over 97 months. The plus sign indicates that the response was ongoing at the time of last contact.

For the 17 patients who received a complete response, the median duration of response has not yet been

observed, but will be at least 54 months.

[Slide.]

To further highlight this, this slide identifies all of the patients who had metastatic renal cell cancer and who achieved a complete response. What is important to note about this slide is that the shortest duration is 23 months. It was because of these impressive durations of response that Chiron initiated a retrospective analysis of patients treated with the same regimen.

[Slide.]

The largest cohort of patients identified were the 270 metastatic melanoma patients that we will discuss today. These patients were treated on eight clinical protocols. To the best of our knowledge, we have identified all of the protocols that enrolled metastatic melanoma patients treated with this regimen.

Chiron discussed the supplemental application with the FDA in a presubmission meeting. At that time it was agreed that the study selection process, the manner in which the data were collected, audited, and presented would support the review of the application.

On April 10th, 1997, Chiron submitted the supplemental application to the FDA and the FDA granted this application a priority review status.

[Slide.]

Today, we are seeking the Advisory Committee's recommendation for approval for the use of Proleukin in patients with metastatic melanoma. The basis of discussion will be the data generated from 270 patients treated in a multicenter environment.

All responders were followed for at least three years, providing strong evidence of Proleukin's ability to produce durable clinical responses in patients with a poor prognosis. The safety data generated from the trials suggests that the toxicities encountered during Proleukin therapy are predictable, manageable, and generally reversible upon completion of therapy.

[Slide.]

Today's proposed agenda is as follows. I will conclude my introduction by introducing Dr. Michael Atkins, who will be presenting a overview of metastatic melanoma. Following Dr. Atkins, Dr. Lori Kunkel will be presenting the safety and efficacy of Proleukin in patients with metastatic melanoma. We will have a brief conclusion and then we can address the committee's questions.

[Slide.]

I would like now to introduce Dr. Michael Atkins. Dr. Atkins is an Associate Professor of Medicine at the

Harvard Medical School and he is currently the Director of Melanoma and Biologic Therapy at the Beth Israel Deaconess Medical Center.

Dr. Atkins.

Overview of Metastatic Melanoma

DR. ATKINS: Thank you, Mary. Good morning to the panel and ladies and gentlemen. I am very pleased to have the opportunity to be here today and present data related to background data related to national history and the results of conventional treatment options for metastatic melanoma.

[Slide.]

My esteemed colleague from across the town has been George Canellos, who has been quoted as saying that metastatic melanoma is a disease that gives cancer a bad name, and there is a lot of truth to this. These are usually young patients, median age of 46, and therapy is suboptimal with median survival of 6 to 9 months in most series, and 2 to 3 percent of patients surviving long term.

[Slide.]

There will be 40,000 patients diagnosed with melanoma in the United States in 1997, and about 7,500 patients will die of metastatic disease. This represents about 3 percent of all cancers and about 1.5 percent of all cancer deaths. So, we do slightly better with melanoma than

we do with other cancers, but the striking importance is the 10-fold increase in the incidence of melanoma since 1935 with the current lifetime risk being about 1 in 90, and estimated to be approximately 1 in 75 by the year 2000.

With this increasing incidence, there is likely to be an increase in incidence of metastatic melanoma, as well.

[Slide.]

Stage IV metastatic melanoma presents in multiple ways. This is the staging system for this disease. There are patients who have more than one lymph node station involved, a single lymph node that is greater than 5 cm or fixed, greater than 5 in transit metastases, or involvement of skin or soft tissue beyond the site of the primary tumor, or visceral metastases.

These top three areas are areas that involve regional spread of disease and therefore might potentially be amenable to local regional therapy, while this group of patients, which involve spread of disease through hematogenous means, are those that more likely need to be treated with systemic therapy.

[Slide.]

The importance of hematogenous spread of disease as a risk factor for poor outcome is shown here in this database by Balch, et al., where the two most important

single factors related to poor prognosis were the number of metastatic sites and visceral sites of disease. A lot of other factors were analyzed and felt to be not significant factors for poor prognosis.

[Slide.]

These two factors are highlighted graphically on these two slides where you can see that patients with one site of disease do significantly, although not in a major way, better than patients who have two or more sites of disease. In this group of patients, there may be a few patients who are long-term survivors, and this probably results from surgical cure.

[Slide.]

Looking at visceral disease, you can see that patients with visceral disease or visceral disease and non-visceral disease do significantly worse than patients with non-visceral disease, and there are few long-term survivors in patients with visceral metastases.

[Slide.]

In the multivariate analysis by Balch, et al., in addition to number of metastatic sites and visceral metastases, remission duration or the time interval between presentation of the primary lesion and development of metastatic disease was also found to be significant.

[Slide.]

If you look at multiple data series, you can see that the same factors keep coming up. The three that I mentioned for the University of Alabama series are also seen in a larger series from UCLA, John Wayne, where organ site of metastases, visceral metastases greater than one metastatic site, and short disease-free interval were associated with poor prognosis.

In the Southwest Oncology database recently reviewed by Larry Flaherty, they substitute liver metastases for visceral metastases, and added performance status as a poor prognostic factor, and in the ECOG database, they added CNS to liver metastases, and had performance status and male gender as poor prognostic indicators.

So, there is pretty much agreement on the type of patients who will do poorly.

[Slide.]

If you analyze these various series through the literature by the prevalence of these various factors, you can see that the type of patients who were often presented to medical oncologists as candidates for systemic therapy, as reviewed in the SWOG and ECOG database, have about 70 percent of patients with greater than one metastatic site and a 1 to 2 percent five year survival.

The type of patients who may be presented first for surgical therapy, such as represented by the University of Alabama or the John Wayne database, may have between 50 and 85 percent of patients with single metastatic sites, and although the median survival is not much better, this single factor may account for the slightly larger five year overall survival in some of these series.

[Slide.]

What are the treatment options for metastatic melanoma? They include surgery in selected patients, chemotherapy, immunotherapy with interferons, interleukins, some experimental vaccines, monoclonal antibodies, or combinations of the above.

[Slide.]

In the large surgical series that have been reviewed in the literature, between 11 and 33 percent of patients presenting to surgeons with metastatic melanoma are deemed resectable and in those patients who are resected, somewhere between 13 and 22 percent of patients will survive long term. So, somewhere around 5 percent of the total population presented to surgeons will have long-term five year survival.

However, when looking at these surgical databases, one must realize that there is bias involved, and these

biases include the fact that there is a heavier emphasis of patients with single-site, limited lesion, therefore, more surgically resectable disease.

These databases include patients with local recurrences, in-transit metastases, or multiple nodal sites who may be more amenable to surgical therapy, and these patients often have received unspecified other therapies. All of these factors tend to overemphasize the value of surgery in this disease.

[Slide.]

What about chemotherapy? Well, a number of chemotherapy agents have been looked at and found to have modest activity with response rates in the 12 to 20 percent range with the most active agent and one most commonly used being dacarbazine.

[Slide.]

Dr. Hill and Dr. Houghton reviewed large series of patients treated with dacarbazine alone, and they showed that dacarbazine produces about a 5 percent complete response rate, about a 19 percent overall response rate with about a four-month median duration of response, and only 1 to 2 percent of patients being alive at six years. So, dacarbazine leaves a lot of room for improvement.

[Slide.]

There are a number of newer cytotoxic agents which have been investigated, and all of these are derivatives of the agents that I mentioned two slides ago. They include temozolomide, which is a derivative of DTIC, fotemustine, a nitrosourea used in Europe, carboplatin, vindesine, taxotere.

All of these have similar response rates and where they are reported, median survival and two-year survival are not dramatically different than their parent compounds.

[Slide.]

What about cytokine therapy? Well, interferon alpha and interleukin-2 are the cytokines which are most commonly used. Interleukin-4, interleukin-6 are essentially inactive, and interleukin-12 is still under investigation. There is very little five-year survival data for cytokine therapy for metastatic melanoma, with the data that you will hear later today for Proleukin, with the 14 percent estimated five-year survival being an exception.

[Slide.]

Alpha interferon, as most of you know, has been approved for use in the high-risk adjuvant setting, and in metastatic disease produces responses in about 16 percent of patients with about a 4 percent complete response rate, and response is higher in patients with small tumor burdens with

the majority of patients, at least some series who are responding, having their largest lesion being less than 1.5 cm. Responses are rare in patients with liver or bony metastases.

You can see why interferon may have value in the adjuvant setting, but it has limited value in metastatic disease.

[Slide.]

A number of combination chemotherapy regimens have been looked at, and in Phase II studies have been reported to have higher response rates than one would expect from DTIC alone. These include the BHD regimen of BCNU, hydroxyurea, dacarbazine; the BOLD regimen of bleomycin, oncovin, lomustine, and DTIC, the addition of platinum to DTIC or platinum and vinblastine to DTIC, or the Dartmouth regimen which includes platinum, DTIC, BCNU, and tamoxifen, which have been reported to have as much as 55 percent response rate in Phase II studies, but five-year data has not routinely been reported.

[Slide.]

The Southwest Oncology Group recently investigated the value of the Dartmouth regimen, which is used around the country in metastatic melanoma, and in 79 patients reported by Margolin, et al., at the last ASCO meeting, had a 15

percent overall response rate, 6 percent complete response rate, and a median duration of response of eight months.

[Slide.]

There have been several randomized studies which compared combination chemotherapy to DTIC alone, this done by Southwest Oncology Group, looked at BHD versus DTIC, and M.D. Anderson looked at CVD versus DTIC, and although in some of the studies it looks like there is a slightly higher response rate, the median survival has not been significantly different.

Eastern Cooperative Oncology Group recently completed a study comparing the Dartmouth regimen to DTIC, and although this data is not yet available, if the Dartmouth regimen performs for ECOG the way it did for SWOG, it is unlikely that it is going to be significantly better than DTIC.

So, in summary, there is no data that supports that a combination chemotherapy regimen is superior to DTIC alone in metastatic melanoma.

[Slide.]

What about the addition to tamoxifen? Well, an Italian group reported in The New England Journal about five years ago that the addition of tamoxifen to DTIC produced significantly improved response rate and median survival

compared to DTIC alone.

Several other groups, such as the NCI Canada, which looked at Dartmouth plus or minus tamoxifen, M.D. Anderson looking at CVD interferon plus or minus tamoxifen, or Pittsburgh Cancer Group looking at carboplatin and DTIC plus or minus tamoxifen, showed no benefit for the addition of tamoxifen.

The Eastern Cooperative Oncology Group recently completed a study which looked at DTIC plus or minus tamoxifen or DTIC interferon plus or minus tamoxifen, with this arm being identical to the Italian group study.

[Slide.]

The data for the addition of tamoxifen is shown here on this slide, and as you can see, in about 250 patients, there is no difference in response rate, complete response rate, time to treatment failure, and median survival.

In summary, there is no convincing evidence that tamoxifen adds anything to DTIC in this disease.

[Slide.]

The Falkson Group from South Africa looked at interferon added to DTIC in a small randomized study and reported higher response rate, higher CR rate, and higher response duration using a schedule that involved a high-dose

intravenous induction of interferon together with interferon given with dacarbazine.

Three other randomized studies, however, have been performed and showed no difference although they did use different schedules of interferon.

[Slide.]

As part of the ECOG 3690 study, also examined the role of interferon using the same schedule as was used by the South African Group. As you can see here, when you look at the benefit of interferon, there is no benefit in terms of overall response, complete response, time to treatment failure, or median survival.

So, at the moment there is no evidence that interferon adds to the value of chemotherapy in metastatic melanoma.

[Slide.]

In summary, we are able to identify the metastatic pattern that is associated with poor clinical outcome in this disease, and that includes patients with multiple metastatic sites and visceral metastases. Surgery produces five-year disease free survival in approximately 5 percent of patients, but these are usually patients with single metastatic sites or single lesion metastases involving skin, lymph node, and lung occasionally.

[Slide.]

Single agent chemotherapy produces a five year survival of about 1 to 2 percent.

Interferon produces responses in about 16 percent of patients, but the responses are largely confined to patients with small volume disease.

Combination chemotherapy or the addition of tamoxifen or interferon to chemotherapy has not yet been proven superior to DTIC alone.

[Slide.]

In conclusion, Dr. Canellos was right, metastatic melanoma is a bad disease. Responses to conventional therapy are usually short and five year survival is rare, and I hope you will agree that additional therapeutic options are necessary.

Thank you very much.

MS. O'HARA: Thank you.

[Slide.]

I would like to now introduce Dr. Lori Kunkel. Dr. Kunkel is our Associate Director of Clinical Development for Proleukin cancer. Lori will be discussing the safety and efficacy of Proleukin in patients with metastatic melanoma.

Lori.

Efficacy and Safety of Proleukin in Patients with Metastatic Melanoma

DR. KUNKEL: Good morning. I am Dr. Kunkel and I am going to present the data today on the safety and efficacy of treating metastatic melanoma patients with Proleukin.

[Slide.]

As you have already heard this morning, the hallmark of Proleukin therapy has been the durable responses, and as was demonstrated in patients with metastatic renal cell, we can now demonstrate the same durable responses in metastatic melanoma patients.

In fact, what you will see is that when a patient achieves a complete response, they have nearly a 50 percent chance of remaining disease free at five years. Thus, we feel that Proleukin offers an important therapeutic option to patients with this disease.

[Slide.]

I am going to begin with an overview of the clinical study design. Chiron had identified a cohort of patients with metastatic melanoma who had received single agent Proleukin therapy administered by Q8H regimen. The patients were enrolled between 1985 and 1993 and a retrospective review of the data was conducted.

However, Chiron did specify prospectively what the definition of complete responses, partial responses, and response durations would be in order to have consistency across all studies.

[Slide.]

The studies were conducted on seven NCI-sponsored protocols. Four were intramural studies, three were extramural, and in addition, there was one Chiron-sponsored study. The studies were conducted over 22 investigational sites. We now have a median follow-up for responders at 62 months, and we are thus able to bring to you a mature database.

[Slide.]

The study objectives were to determine the efficacy of Proleukin therapy with respect to response rates, response duration, progression free survival, and survival.

In addition, we wish to identify the safety profile in the patients with metastatic melanoma.

[Slide.]

The patients were assigned or randomized and must have received Proleukin administered as a single agent. These patients all had measurable disease, an ECOG performance status from zero to 2, and a signed informed

consent.

[Slide.]

In addition, the institutions had established study criteria which included cardiac and pulmonary function screening, liver, kidney, and hematologic parameters, and patients with CNS metastases, those patients with active infections, or use of concomitant steroid therapy were excluded from the protocol.

[Slide.]

This defines what a course of Proleukin therapy was on these protocols. The patients received Proleukin every 8 hours by a short I.V. infusion. They received up to a maximum of 14 doses administered over 5 to 6 days.

These patients were treated to maximum tolerated toxicity. Doses could be withheld, however, there were no dose reductions allowed on the protocol. There was a rest period of 7 to 9 days, and the patients went on to receive cycle 2.

At the end of the first course of therapy, the patients were reevaluated, and only patients who had stable disease or who were responding to therapy were eligible to receive additional courses of Proleukin treatment.

Within a very short time, both the patients and the physicians knew whether or not there was a response to

Proleukin therapy. If the patient wasn't responding appropriately, they were then eligible to receive other therapies.

The majority of patients entered on these protocols received one or two courses of Proleukin.

[Slide.]

This slide summarizes the eight efficacy protocols that the patients were enrolled on. There were 140 patients enrolled on the intramural studies, 118 on the extramural, and 5 patients on the Chiron-sponsored study.

[Slide.]

Now, there was some variability on the dose that was planned to be administered, variability between the intramural and the extramural side, however, it is important to note that all patients received a dose-intensive regimen. If you look at the median cumulative dose received in course 1, it was essentially the same.

[Slide.]

For example, if we look at the patients on the intramural studies who were scheduled to receive 720,000 IU/kg, they tolerated fewer doses per course than those patients who received 600,000, thus again, the median cumulative dose per course was equivalent across studies.

Now, for simplicity, there are 5 patients on the

Chiron-sponsored studies that will be included in the extramural in further discussion.

[Slide.]

When we look at the patient characteristics on this study, of the 270 patients who met the eligibility criteria, the median age was 42 years. Sixty-four percent of these patients were male and 71 percent of patients had ECOG performance status of zero.

[Slide.]

We also looked at the patient characteristics with respect to prior treatment. Now, all patients had received prior surgery for resection of their primary disease. When we looked at treatments for metastatic disease, we could see that nearly 50 percent of patients had received one form or another of prior systemic therapy and had progressed on that prior to receiving Proleukin.

Fourteen percent of patients had received chemotherapy, 19 percent immunotherapy, a few patients hormonal, and 12 percent of patients had received one or more at the modalities above.

[Slide.]

The other patient characteristics that we looked at included the patterns of metastatic sites of disease, and as you can see here, that 71 percent of patients had two or

greater sites of metastatic organ involvement at time of Proleukin treatment.

In addition, 69 percent of patients had at least one site of visceral involvement at the time of Proleukin treatment. These patients had multiple lesions within each of these sites. Thus, these patients would be a group that would have been predicted by what Dr. Atkins has presented, of patients with poor clinical outcome.

[Slide.]

I am now going to move on to discuss the clinical endpoints of the study. The overall response rate on study to Proleukin treatment was 16 percent with a 95 percent confidence interval from 12 to 21 percent. Forty-three of the 270 patients had objective response.

What we can see is that the response rates were very similar on the intramural and extramural sites with respect to complete responses, as well as partial responses. Thus, we could demonstrate that response to Proleukin could be attained in a multicenter environment.

[Slide.]

The duration of best response has been the hallmark of Proleukin treatment, and it is defined on the studies as being the time from the best objective tumor response until the time that the patient progresses. This

is important to remember because especially in the complete responding patients, they may have had partial responses for several months before achieving their complete response, however, their complete response is calculated from the time they achieve the complete response.

The overall median duration of best response is 8.9 months for the entire group, but most importantly, when we look at the 17 patients who were complete responding patients, we can see that their median duration of response has not yet been observed, but is exceeding 40 months at the time of submission in the fall of 1996.

[Slide.]

This slide summarizes those patients with durable responses. They all exceed two years, and one patient is approaching nine years currently. The positive sign indicates that these are durable responses without any intervention.

[Slide.]

Progression free survival on this study was defined as the time from initial dose of Proleukin therapy until the patients progressed. The overall median progression free survival is 13.1 months. Again the progression free survival has not been reached for those 17 patients who achieved complete remission. For the partial

remission patients, it is 8.3 months.

We know that for the complete remissions that their median duration will be in excess of 54 months.

[Slide.]

This is a Kaplan-Meier plot of progression free survival in all responding patients. As you can see here, that 50 percent of the patients remain progression free for at least a year, and you can also note here that we have not seen any relapses occurring after 30 months. Thus, we begin to see the curve plateau very nicely.

[Slide.]

This is our Kaplan-Meier plot of survival of all patients. The median overall survival was 11.4 months for all patients enrolled on this study. Again, you begin to see the curve plateau and the overall survival at five years is projected to be 14 percent.

We have 30 patients overall surviving and 20 of those patients are non-responders.

[Slide.]

We committed to looking at long term follow-up of all the surviving patients, and when we look at our complete responders, we can see that 10 of the 17 have ongoing complete responses without any further treatment, and all their responses are exceeding 24 months. Seven of the

complete responding patients have relapsed, and 2 of these are long term survivors, and these patients are of particular interest to us because Proleukin had induced multiple regressions at multiple sites of disease including visceral sites. When these patients subsequently relapsed, they often relapsed at a single site, which then allowed them to be treated with local therapy and become long term survivors.

[Slide.]

This is an example of one of the complete responding patients, and the patient presented, as you can see here, at the time of Proleukin treatment with over 50 percent of involvement of her liver with metastatic disease and in addition, she had lung and lymph node involvement.

This is your baseline. This is after one course of Proleukin. You can see a marked reduction in the size of her masses, and actually resolutions of several masses. This is the next course of Proleukin treatment, again continuing response, and at this point, approximately a year after her first treatment, she has some scan abnormalities that, upon biopsy, showed no evidence of disease. So, this patient was coded as a complete responding patient. She did subsequently relapse at a local site, but was again salvaged and went on for a second remission.

[Slide.]

When we look at the follow-up of our partial responders, we can see that 2 patients who were classified as partial responders have ongoing responses exceeding 54 and 91 months, but had no further intervention. These patients were conservatively classified as partial responders because they had persistent scan abnormalities at the completion of the Proleukin treatment. However, they have had no further intervention and are clinically essentially in complete remission.

Twenty-four of our partial responding patients have subsequently progressed, and we have 6 patients who are long term survivors. Again, these patients are of interest because Proleukin had induced remissions in multiple sites of disease in these patients and, in fact, complete remissions in some of the sites, and when they progressed, they often progressed at a single site, which allowed them to be salvaged, oftentimes with local therapy, and become long term survivors.

[Slide.]

This is an example of one of the partial responding patients. This patient at the time of Proleukin therapy had failed DTIC, interferon, and combinations of DTIC, interferon, and tamoxifen.

At the time of Proleukin treatment, the patient had lung involvement, multiple nodules, as you can see here, as well as liver involvement and lymph node and subcutaneous involvement.

[Slide.]

This shows the post-treatment after one course of Proleukin therapy, and you can see that there is marked response in the previous sites of disease, but we do have some persistent scan abnormalities.

In fact, the patient's x-rays haven't changed over the years, but upon completion of treatment, since they had persistent scan abnormalities, they were classified as a partial responder. The patient has received no further treatment.

[Slide.]

This is the same patient looking at the liver disease that she presented, again pretreatment, after one course, and the persistent scan abnormalities that the patient had at last follow-up in 1995.

[Slide.]

This is another patient who at time of Proleukin therapy had large volume subcutaneous disease, as you can see here. In addition, the patient had bony involvement, as well as lung involvement. This is after one course of

Proleukin treatment. The patient essentially was beginning to show marked regression of all tumor sites, continued on with an additional course of treatment. The patient was classified as a partial responder.

[Slide.]

An additional patient who really is of interest because they have this large, about 5 cm mass in their lung, 7 cm mass in the adrenal gland, after two courses of Proleukin treatment we can see the best response was a partial response with marked reduction of both the lung and the adrenal lesion.

A few months later the patient, although they continued to have regression in their adrenal metastasis, had progression in their lung metastasis, but since they had begun to resolve all other sites of disease, it was felt that this was surgically resectable.

The lung lesion was removed. The patient continued to regress in the adrenal gland and essentially remained disease free for a number of years.

[Slide.]

The five year clinical outcome is important in oncology because it is often at that time that we can consider patients cured of their disease, but it is rare to actually find reports of five year clinical outcome in

patients who have been reported in the literature.

So, we felt it was important to look at our five year data on the responding patients to Proleukin. What we can see is that 59 percent of the complete responding patients have maintained their responses for at least five years. Twenty-nine percent of all responding patients have response durations exceeding five years.

[Slide.]

With respect to survival, 76 percent of all complete responding patients are surviving five years, and 51 percent of all responding patients are surviving for greater than five years.

[Slide.]

Now, the studies were designed to look at response rate and response duration, but we felt it was also important to look at additional factors that may have predicted response to Proleukin therapy, and of the factors that were looked, there were only two that were determined to be associated with a response to Proleukin: the patient's performance status and whether or not they had received prior systemic therapy.

[Slide.]

What this slide demonstrates is that those patients, 191 patients with the ECOG performance status of

zero had twice the response rate as those patients with ECOG performance status of 1, and this association was significant.

However, it is important to note that responses were still obtained in patients with ECOG performance status of 1.

[Slide.]

Similarly, we demonstrated that there was an association between those patients who had not received any prior systemic therapy. Their response rate was twice that of those patients who had some form of prior systemic therapy. Again, this association was significant, however, we did see responses in patients who had received systemic therapy.

[Slide.]

We did look at the characteristics of the responding patients with respect to their patterns of metastatic involvement, and we did not see any association to response to Proleukin.

As you can see here, the patients with multiple sites of disease had a consistent response, as did those with the patients with a single organ site of involvement at time of Proleukin treatment.

[Slide.]

The patients with visceral involvement also had a consistent response compared to those patients who had no visceral involvement.

[Slide.]

So, in summary of the efficacy, we see that 16 percent of the patients treated responded to Proleukin therapy, and we saw responses in patients with visceral disease and multiple metastatic sites of disease.

Again, we could demonstrate that there were durable complete responses and that of those patients who achieved complete responses, 59 percent are cancer free at five years.

[Slide.]

I am now going to move on to the review of the safety. The toxicities associated with this regimen are well characterized. This is the same regimen that was used for treatment of the metastatic renal cell patients, and the safety in the current package insert summarizes over 500 patients, of which 102 patients included have metastatic melanoma, and these are patients that were included in this analysis.

[Slide.]

As a reminder with respect to the toxicities, the toxicities that are now well recognized with this regimen

were just being characterized at the time that these studies were conducted, and the major toxicities are capillary leak syndrome, which is manifested by hypotension and hypoperfusion in major organs, pulmonary, cardiac.

In addition, there is neurological toxicity associated with this dose-intensive regimen. These patients develop mental status changes, confusion, and somnolence while they are on treatment, and there is a sepsis-like syndrome associated with the regimen, however, these are now well characterized and, as the investigators gained experience with administering these dose-intensive regimens, there were treatment guidelines established because these became predictable.

[Slide.]

The treatment guidelines included screening of patients for adequate cardiac and pulmonary function, as well as renal function, and the introduction of supportive measures, the use of concomitant medications to alleviate the predictable toxicities, and also the use of prophylactic antibiotics to prevent a sepsis-like syndrome.

These treatment guidelines have been incorporated in the package insert since 1992.

[Slide.]

This is a summary of all adverse events reported

in the metastatic patients in at least 20 percent of the metastatic patients. Essentially, these are the same adverse events that were seen with this regimen when used in the metastatic renal cell patients, and these adverse events are listed in the current package insert.

[Slide.]

This slide summarizes all of the Grade 4 toxicities that were seen with this regimen, by number and by incident, and again these were essentially the same adverse events that were seen with treatment of the metastatic renal cell patients.

Although these Grade 4 adverse events do occur and are severe, the majority of them are reversible upon completion, one to two days of completion of the Proleukin regimen.

[Slide.]

This slide summarizes the early terminators on this study, an early terminator being defined as patients who did not continue on the protocol for reasons other than progressive disease or death.

We can see that the primary reason for patients terminating the study was acute toxicity and cardiac and respiratory being the primary reason, but you can also see that the long term effects, even if the patient terminates,

are rare.

There is one patient who had persistent cardiac dysfunction following a myocardial event on the protocol, and one patient had developed ischemic necrosis requiring amputation of metatarsals.

We also had 5 patients who, although they did not have life-threatening toxicities, could not tolerate the treatment and refused to continue. These patients, all their toxicities had resolved at the completion of treatment, and there are no long term effects on those patients.

[Slide.]

This is our drug-related death rate. The overall incidence of on-study drug-related deaths was 2 percent. All 6 deaths were related to the sepsis syndrome, and you will note that all the deaths occurred before 1990, before the use of prophylactic antibiotics was standard of care, and we know that none of these patients were on prophylactic antibiotics who died from this syndrome.

[Slide.]

So, in summary of the safety, we recognize that the toxicities associated with this regimen are common. They are severe, they can be severe, but they are also predictable and in most cases they are reversible upon

completion of treatment.

Importantly, most of these toxicities are not chronic or cumulative. There have been treatment guidelines incorporated into the package insert that have been put into practice since 1992.

[Slide.]

So now we can do a comparative observation between the metastatic melanoma database and the metastatic renal cell database. You can see that the patients enrolled for these two submissions are very similar, that the overall response rate is actually remarkably similar, and the breakdown of complete responding patients and partial responding patients is similar.

[Slide.]

But, more importantly, again, we have demonstrated that the duration of responses is durable, with the median duration of response in the metastatic melanoma patients, complete response patients, at least 40 months, and in the metastatic renal cell patients, at least 54 months.

[Slide.]

When we look at the drug-related on-study death rates, they are also comparable between the two groups, a 2 percent rate on the metastatic melanoma patients, and 4 percent on the metastatic renal cell.

[Slide.]

So, we feel that Proleukin treatment of patients with metastatic melanoma have a favorable risk-benefit, the toxicities of those severe are also predictable, they are manageable and reversible upon completing treatment, and that this treatment provides the opportunity for durable responses in patients.

[Slide.]

Thus, we feel that Proleukin is an important therapeutic option for patients with metastatic melanoma.

On behalf of the Chiron Corporation I would like to express our appreciation for the patients who participated in these studies and also to the physicians and the nurses who cared for these patients over the years.

Conclusion

[Slide.]

MS. O'HARA: Our conclusion was basically that we feel that Proleukin should be available for patients with metastatic melanoma.

[Slide.]

And we are recommending the same regimen that is in the current package insert, so that would include the same screening parameters, the same monitoring parameters, and aggressive management of the toxicities.

We would like to thank the committee for allowing us to share the data. Thank you, Dr. Swain.

DR. SWAIN: Thank you.

Committee Questions to Applicant

DR. SWAIN: We would like to open it up for questions. Dr. Raghavan, if you could start, or Dr. Ozols.

DR. OZOLS: One of the questions is how many patients with metastatic melanoma that you have screened would be eligible for this kind of a treatment protocol? It seems that from the database that they may be slightly younger, slightly more fit patients than what you would expect. The median age was 42, for example.

So, how many patients do you think are eligible for this type of treatment?

MS. O'HARA: Dr. Kunkel or Dr. Atkins?

DR. ATKINS: The majority of patients that we would screen for metastatic melanoma would be eligible. The patients who would be excluded would be those who would have CNS metastases or poor performance status, who probably aren't going to be treated on many of the other protocols.

The renal function, cardiac, and pulmonary hurdles that these patients have to clear in order to receive this therapy are really not a major problem for these patients who are usually relatively young and otherwise healthy, but

a few patients may have concomitant other diseases that would make them poor candidates for therapy.

DR. OZOLS: Have you re-treated any of the CRs who relapsed with IL-2?

DR. ATKINS: There is not a large database with that, but I think a few of the patients who were CRs here or even partial responders went on to receive other therapy that included IL-2, and some of those are durable responders, but the usual situation is that if you give more interleukin-2 after someone has progressed, you often see that that is resistant disease.

DR. OZOLS: One final question about the dose. Do you feel that this is -- obviously, this is a dose that is in the package insert for renal cell cancer -- what about other doses? I mean is this a dose that is accepted as being the optimal dose in the situation?

DR. ATKINS: There are a number of smaller studies in smaller series that looked at lower doses or doses of interleukin-2 combined with interferon, and the data is just not as good. We are still looking at combinations of this dose with other agents, such as various peptide vaccines or lower doses in combination with chemotherapy where the results are promising, but we just don't have long term follow-up in order to present that data now.

DR. RAGHAVAN: I have a few questions, as well, particularly for the group with just one side of metastatic disease. What proportion of those patients actually had biopsies to confirm metastatic involvement?

DR. ATKINS: All patients had biopsy-confirmed metastatic disease in order to go on study, and documented progression of disease before treatment.

DR. RAGHAVAN: In the context I think I understood the presentation that there were patients who were objective non-responders, but who went on to survive a lengthy period of time. Could you tell us a little more about what you think was going on there?

DR. ATKINS: Well, there were 30 patients who are alive at last analysis, 20 of those are in the responding group. The other 10 represent about 3 to 5 percent of the total population who may be the type of patients who would be alive with other treatment modalities.

A few of those patients were minor responders who went on to be able to have surgical resection or some durable minor response, but the majority of those patients received some other therapy, be it chemotherapy or aggressive surgery, or a few of them, other immunotherapy regimens that may have produced a durable response in another 10 patients. We are just saying that that is an

additive benefit to interleukin-2.

DR. RAGHAVAN: Can you tell us a little more about the sepsis syndrome, what are your thoughts about why it occurred, why this prophylaxis helped, et cetera?

DR. ATKINS: We have learned a lot about this. Actually, in our early studies when we were giving IL-2 with LAC, saw about a 20 percent incidence of bacteremia, which was primarily related to catheter site infections.

Most of these patients had central catheters in, and they had skin toxicity, colonization around the catheter site, and staph-related bacteremia, which was poorly tolerated in a group of patients that already have the side effects of interleukin-2.

We spent a lot of time trying to figure out what the cause of this was, trying to see if it was the surgeons, the nurses, the house staff who were responsible, and we couldn't attribute it to any group of people, so at our institution, and I think at other places around the country, we looked at neutrophil function in these patients and found that the neutrophils were paralyzed with response to standard chemotactic stimuli, such as F met leuphe, and this happened about five days into treatment around the same time when they were at risk for infection.

This is a situation that is similar to the

Chediak-Higashi syndrome, which is associated with staph-related infections, which can be benefitted by prophylactic antibiotics, so that led to a series of trials around the country with prophylactic antibiotic therapy, which has reduced the incidence of infection to about 3 or 4 percent, and the infections you see are more likely Staph epidermidis and Staph aureus, and not associated with the same serious consequences.

DR. RAGHAVAN: I was intrigued by the rather small series of Chiron patients against a background of rather larger series. It was n equals 5. Maybe it was explained, but I guess I missed it. Could someone comment on the purpose of that series?

DR. KUNKEL: The 5 patients that were enrolled on that Chiron-sponsored study were part of a larger study that treated many patients with different types of cancers. Those patients received essentially the same dose-intensive regimen. So, for completion, they were included in the analysis for primarily safety reasons, but also we looked at efficacy in those 5 patients.

DR. RAGHAVAN: It seemed to me that they actually received substantially a lower dose-intensive regimen than the others. Now, with 5 cases, I understand that we can't draw any major conclusions. Did you see any responses in

that group?

DR. KUNKEL: Actually, we did see a partial response in that group. It is actually one of the durable responses.

DR. RAGHAVAN: Which brings us back to Dr. Ozols' question, which is let's talk a little more about dose intensity, and moving away from the big data sets, could we look a little more at the other quanta of data that you have either in renal cell or melanoma. Dr. Atkins is well known for work in both areas, so take us through that.

DR. ATKINS: I think it is wrong to assume that just because those patients received a lower amount of interleukin-2 per dose, that they received a less intensive regimen. I think Dr. Kunkel showed that the amount of interleukin-2 that they received per course was very similar, and since not every patient receives all their doses during a course of therapy, they are sort of each treated to their individual MTD, and with both of those regimens, patients received as much interleukin-2 as they could tolerate.

In renal cancer, there is some data that lower doses of interleukin-2 have an activity that is fairly similar to what is seen with high dose bolus interleukin-2, but the data that is out there now seems to indicate that

the response durations are not the same with those lower doses, and it is more analogous to what you might see with interferon in renal cancer.

We don't even have that degree of confidence with lower doses of interleukin-2 in melanoma where the response rates are significantly lower if you give interleukin-2 by continuous infusion of subcutaneous administration in melanoma, and we don't even get to the point of talking about durability of response.

Where lower doses of interleukin-2 may find a role is when it is given with chemotherapy or potentially with other agents, such as interferon or vaccines or potentially interleukin-12, and we are also actively looking at ways of trying to dissociate the toxicity of interleukin-2 from its anti-tumor effect by giving a variety of agents that might block TNF or nitric oxide production.

DR. RAGHAVAN: You showed very elegantly the fact that many of the provocative Phase II trials in the melanoma literature have been disappointing when tested in randomized fashion. I am personally quite comfortable with the idea that you don't need to do a controlled trial of interleukin-2 in this context. However, I wonder if you could take us through your thoughts on what sort of studies you think might be needed at a later time to confirm the

biological activity of interleukin-2 and most particularly, what your strategy would be for looking at the question of dose.

Is it time to start asking a high dose/low dose question, do you need to compare interleukin versus interferon, and so on?

DR. ATKINS: At the moment, we don't have enough data with low dose interleukin-2 in melanoma to justify a high dose versus low dose study using interleukin-2 in metastatic melanoma.

We are looking at a national level at combination of chemotherapy plus interleukin-2 and interferon versus chemotherapy alone to see whether interleukin-2 adds to chemotherapy in that context using a regimen that is safe enough to be administered in a cooperative group setting.

We also hope to potentially look at ways of giving interleukin-2 together with various immunogenic peptides that might potentially be compared to interleukin-2 alone, as well as with various toxicity reduction agents that may allow us to look at dose relationship between toxicity and response.

I don't think the data with interferon is sufficient enough to justify its use in the majority of patients with metastatic disease although we still use it

quite frequently in patients who have had regional disease or metastatic disease resected, and we are looking at interferon in that setting, but the interferon regimen is a pretty dose-intensive regimen, as well, and lower doses of interferon that have been looked in the adjuvant setting have not been shown to be effective.

So, dose intensity seems to be a recurring theme for immunotherapy in melanoma.

DR. RAGHAVAN: I would accept that it is a recurring theme. I think I would also question whether it is necessarily a proven outcome, and it seems to me that you could certainly, with a regimen that isn't a panacea, it clearly has a role but isn't a panacea, I am puzzled that you are not interested in specifically exploring a lesser level of dose intensity to ascertain whether you actually do lose percentage survival points.

DR. ATKINS: We have done this a number of times, and the results just aren't there. The response rates are 5 or 10 percent, and durable responses aren't seen.

DR. RAGHAVAN: Numbers of cases that you have looked at in those series?

DR. ATKINS: Around the country, there are probably several thousand patients who have been treated with lower doses of interleukin-2 either alone or interferon

in a variety of different series.

DR. SWAIN: Dr. Simon.

DR. SIMON: There is one aspect of the data that I am having trouble understanding, and I wonder if you could clarify for me. It is sort of based on figure 4.2 of your NDA, which you show a Kaplan-Meier plot of progression free survival for the CR patients, the PR patients, and the responders combined.

For the CRs you show that the curve plateaus at about 60 percent, so since you had 6 percent of your patients were CRs, 60 percent of 6 percent is about 3.6 percent, so that would account for 3.6 percent of five year survival.

Now, for your PRs, that curve plateaus at it looks like about 10 percent, and you had 10 percent of your patients PRs, so 10 percent of 10 percent is 1 percent, so that would account for 1 percent of your five year survivals.

So, 3.6 percent plus 1 percent is 4.6 percent, but you are claiming you have a five year survival rate of 14 percent. So, who are the other 9 percent, are these patients who had progression, the PRs who had progression, but subsequently were treated with something else and became five year survivors or what?

MS. O'HARA: I would like to ask Dr. Yoshizawa to first comment on the curve, and then Dr. Kunkel.

DR. YOSHIZAWA: Carl Yoshizawa from Chiron Corporation. If I understand you correctly, you are referring to progression free survival in just the responding patients. That 14 percent that was referred to was overall survival, so it includes the time from start of treatment to response. Also, it includes all patients including non-responders.

DR. SIMON: That is what I am trying to get at. This curve would account for 4.6 percent of five year survivors. You claimed overall 14 percent, so there is a missing 9.5 percent. Are those accounted for by non-responders who were treated subsequently on some other treatment and went on to be five year survivors, or is it accounted for PRs who went on to be treated and became five year survivors?

DR. KUNKEL: At the time of the submission in 1996, we knew that there were 30 patients who were definitely alive and 39 patients who were lost to follow-up.

DR. SIMON: Thirty patients alive in five years out or just alive?

DR. KUNKEL: Alive. So, this is a Kaplan-Meier plot, the expected survival, five year survival. We also

knew that there were 39 patients lost to follow-up, so the numbers are calculated based on that.

With respect to the non-responding patients, there are 10 non-responding patients who are alive, and as Dr. Atkins indicated, we do have some follow-up on those patients.

MS. O'HARA: I would like to have our consultant, Dr. C. Fai Pang further elaborate on the plot.

DR. PANG: I want to clarify the definitions for progression free survival. For progression free survival, it is considered as an event if patients either progress or die. For survival, if they progress and didn't die, it is not an event. That accounts for the difference of the proportions.

So, for progression free survival, if patients progress, but did not die, so it is considered an event, and that is the difference between the progression free survival proportions and this survival proportion.

DR. SIMON: I understand that. I think what accounts for it is the 39 patients lost to follow-up. I think claiming a five year survival of 14 percent is not correct. I think the five year survival rate for the overall group is less than 14 percent, and because you have censored patients who are lost to follow-up, you probably

censored patients who died, and that gives us a biased Kaplan Meier curve.

DR. ATKINS: I think of the 30 patients who were alive at the last analysis, 10 of those were non-responders who are alive. Twelve of those are patients continuing in response, and 8 of those are patients who responded, then progressed, and got treated with some other therapy, sometimes just local therapy, who are still alive.

That accounts for the 30, and then you were talking about the difference in denominator.

DR. SWAIN: Dr. Santana.

DR. SANTANA: I want to get back to this issue of clarify for me about dose and early termination patients. Can you tell us what percent of patients had any temporary stopping of interleukin-2 during first cycle, and if so, what algorithms were used for dose modification? And a corollary to that question is, do you have any data relating the patients that responded regarding the dose received and their ultimate toxicity during the cycle 1?

DR. ATKINS: The median number of doses of interleukin-2 received, I believe is about 22, so the maximum could have been 28, but many of these patients therefore, probably almost all of them, had doses held, and the doses were held for the standard reasons, the same

reasons we hold doses for patients with renal cell cancer receiving interleukin-2.

They are patients who have low blood pressure that requires pressor support, patients who have neurologic toxicity, such as agitation or restlessness or sleepiness, patients with significant diarrhea, occasionally breathing problems, occasionally cardiac arrhythmias, and anytime where we thought that giving another dose might put a patient at risk, we held, allowed for some of the toxicity to resolve, and then proceeded once we felt it was safe.

DR. SANTANA: Proceeded at the same dose?

DR. ATKINS: At the same dose. There were no dose modifications at all in any of these regimens.

DR. SANTANA: And the next question was in these patients that responded, could you relate any toxicity or doses to their ultimate chance of having a response?

DR. ATKINS: We couldn't correlate severe toxicity with response. Many patients were treated to their individual MTD, but that could have been a different problem in each patient, but there was no correlation with degree of toxicity individual patient or as a whole in response.

DR. SWAIN: Dr. Vose.

DR. VOSE: Were there any specific sites of visceral involvement that appeared to have better outcome

compared to other visceral sites based on response rate?

DR. ATKINS: No. Liver and lung responded apparently equally to cutaneous and lymph node disease.

DR. VOSE: Was there any assessment by volumetric CTR or other analysis as far as low volume versus high volume disease as far as response rates or survival?

DR. ATKINS: The median tumor volume for all responders was slightly, just under 25 square centimeters, so this was significant tumor burden for the responders. I don't know the data is -- Lori, do you want to comment, is there data on how the responders compared to the non-responders in terms of tumor burden?

DR. KUNKEL: We do have data for the tumor burden on the responding patients. For the non-responding patients, target lesions were measured, so we don't have the complete tumor burden for those patients. They were only measured for target lesions, whereas, our responding patients, every lesion was measured.

We do have a slide that shows a relationship between progression free survival and tumor burden, but there is no relationship between that.

DR. SWAIN: Dr. Johnson.

DR. JOHNSON: I want to come back to an issue that Dr. Raghavan touched upon. Dr. Atkins, I thought an

excellent review of the subject matter, talked about a number of pilot studies that gave very excellent results from single or even multi-institutional trials, but then subjected to randomization in larger multicenter trials failed to live up to their expectation.

How is this different? How is what you are telling us now different than what these other data --

DR. ATKINS: I think the major difference is it is a different modality and the durability of the responses are different than what was seen with any of those other approaches, and this a 270-patient database treated at multiple sites using a specific regimen. So, I think that is the major difference.

DR. JOHNSON: If we accept that, why should we not perform a randomized trial? Why should we not take these carefully selected patients and subject this type therapy and then compare it to standard therapy?

DR. ATKINS: I think it would be very difficult to do a randomized trial when you have treatment option associated with durable benefit, albeit it in subset of patients, and to preclude patients from being able to receive that option.

DR. JOHNSON: If, in fact, there is a 10 percent survival, benefit of five years, we will give you that

benefit of doubt just for the sake of discussion, and 5 percent of patients highly selected that received DTIC, or frankly may not have received anything, were found to be alive at that interval of time, how could you argue the point you just made?

DR. ATKINS: I would say that this a bad disease and these patients need whatever treatment options are available.

DR. JOHNSON: It is not a question of denying them treatment option. I would say that probably the Dartmouth Group said that their therapy should be none denied to anyone, as well, when they first reported it.

DR. SWAIN: I guess one of the points that you are getting at is the selection bias --

DR. JOHNSON: Right.

DR. SWAIN: -- that is obvious in the studies, and you made several points in your discussion that disease free interval was important. Do you have any information on that in your patients, and the patients who were non-responders and survived?

DR. ATKINS: Do you want to comment on that?

DR. KUNKEL: We do have disease free interval for the responding patients. For the non-responding patients, that data was not collected prospectively. For the

responding patients, the median disease free interval was I believe 22 months.

DR. SWAIN: I agree with Dr. Johnson's point. I think you can liken it to bone marrow transplant for breast cancer, and everyone said you can't do randomized trials in that, you have to transplant everyone, but obviously, that is not true, so I think it really is a very good point because it is a very highly selected group of patients.

DR. KUNKEL: I would actually say that I don't think it is a highly selected group of patients. These were patients with poor prognostic factors, they had multiple sites of disease. They had multiple visceral sites, and they had failed the best available therapy at that time and at this time.

DR. SWAIN: I guess the correct way to say this, there is a selection bias just because it wasn't randomized.

DR. KUNKEL: That is correct.

DR. JOHNSON: You didn't present data, but you may well have the data, about gender. Did that make a difference in outcome? The very dramatic slides you showed were all women, and I am just wondering if women did better or was that a selection bias?

DR. KUNKEL: It may have been an inadvertent selection bias, but no, that was also analyzed in the

patient characteristic response to Proleukin, and there was no difference in gender with respect to either partial or complete response.

DR. JOHNSON: I guess for me personally, I still believe a randomized trial should be done. I will just state that for the record, but there was a group of patients that I was impressed about, and that was those 34 patients who had in fact had received chemotherapy and were treated with interleukin, and there was sort of a passing comment about those patients, but do you have more detailed data about that group of patient, and how did they fare?

DR. KUNKEL: Yes, we have a slide on patients who had received prior -- the responding patients who had received prior chemotherapy.

DR. JOHNSON: I am really interested in the 34 patients who had received prior chemotherapy, the totality of that data set. It is admittedly a small group, but I mean if you were to show me that the response rate was the same, the long term survival in that group was the same, et cetera, I would be personally more convinced.

[Slide.]

DR. KUNKEL: These are the responders with prior chemotherapy. Patient 014KC is actually one of our long term complete responding patients. The prior chemo was not

defined. These patients received platinum, melphalan, platinum DTIC, DTIC, and CVP, and melphalan and DTIC, which were fairly standard at that time or available treatments at that time.

As you can see, there are two of those patients, responding patients, alive.

DR. JOHNSON: And that first patient who only responded for 2 1/2 months, but who is still alive, what transpired next?

DR. KUNKEL: This was one of the patients that i talked about where they had complete remission at all sites of disease, and then progressed. I think this particular patient progressed in soft tissue site, and then was resected at that site, and has remained disease free since that time.

DR. SWAIN: Dr. Simon.

DR. SIMON: I just want to clarify one of the answers that was given. Someone said there were 39 patients lost to follow-up from the point of view of this survival curve. Was there some explanation for that? And then someone said there were 30 patients alive. I would like to sort of rationalize those numbers. I would like to get some information on you say if it is 30 patients, what is their time since last contact? How up to date is this data? I am

trying to decide whether there is an issue here of patients were listed as alive in your survival calculation, but actually, they were dead and you just couldn't find them.

DR. KUNKEL: Chiron has been committed to trying to find all these long term follow-ups, and we have recently updated the database. The date of last contact for all the alive patients was '96, so they have been recently followed up, and I am sure they are continuing to be followed up, but that data was at the time of presentation.

We have recently updated our database, and I don't know whether Dr. Litwin would prefer to speak to this since he has that data. That was not included in the submission because the update came after that.

DR. SIMON: What about the 39 patients lost to follow-up?

DR. KUNKEL: We actually have located death dates on all but 10 patients now.

DR. SIMON: Say that again.

DR. KUNKEL: We have located death dates on all but 10 of those 39 patients, so now we have 10 patients lost to follow-up of the entire database, or a 4 percent lost to follow-up.

DR. SIMON: But that is not reflected in this survival curve or your claim for five year survival.

DR. KUNKEL: No, because the data was presented at the time -- this reflects the submission data.

DR. SIMON: My impression is that your actual five year survival rate for the group of 270 patients is something of the order of 5 or 6 percent. Do you have some reason for thinking that is not the case?

DR. PANG: What we are going to be showing are the five year survival probabilities. It is overall survival as opposed to progression free survival, and this was in Dr. Kunkel's presentation.

[Slide.]

So, probably it would be more appropriate for you to use these rather than those for progression free survival.

DR. SIMON: I have seen the overall survival curve in your submission, but what I am saying is that is not a reliable figure, because of the lost to follow-up patients the way the Kaplan-Meier curve, it assumes there is not informative censoring. There is informative censoring, so I understand that. I am saying the data that goes into that curve I don't think is correct.

DR. PANG: I see Dr. Litwin is leaving.

[Laughter.]

DR. PANG: I think I will take the approach of

begging for forgiveness rather than asking for permission. We have updated survival curve with the information that Dr. Kunkel referred to with now only 10 patients lost to follow-up, and the survival curve doesn't change very much. The five year survival rate drops from 14 percent to 12 percent, but is not 5.

Perhaps there is a certain amount of informative censoring that took place, but not to the extent that you are suspecting.

MS. O'HARA: It is important to note that the follow-up information that we are talking about, those dates of death, these were all non-responders, so I wanted to clarify that point.

DR. SWAIN: I just had one question. Since 1990, have there been any deaths in your database using the compound?

MS. O'HARA: I am sorry. You are talking about the clinical database?

DR. SWAIN: Yes, and all the renal cell and all the database you have, because you said since 1990, you have instituted the prophylactic antibiotics, and I guess it has been widely used.

DR. ROSENBERG: There has been a great deal learned about dealing with the toxicity of interleukin-2,

and, in fact, in our first 200 patients, our treatment-related mortality was between 3 and 4 percent, but a randomized trial that taught us that prophylactic antibiotics could virtually eliminate completely septic complications, and the need to exclude patients that had ischemic heart disease led to a modification in the way patients were managed.

We, since 1990, have treated over 1,000 patients with a single treatment-related death, and in one consecutive series of 806 patients, all treated with a high dose regimen, did not have a single treatment-related death in 806 consecutive patients.

So, I do believe that with appropriate management, this high dose interleukin-2 regimen, in our case it is 720,000 International Units/kg can be very safely administered by groups with experience.

DR. SWAIN: Thank you. Dr. Siegel, I think you had a question.

DR. SIEGEL: I actually had a comment while we were talking about selectivity and perhaps, Dr. Rosenberg, you could stand by because you could comment on this more than I could.

One of the issues regarding selectivity was the fact, I believe for the intramural program, that there were

in many cases back in the late eighties, many more patients applying for the protocol than actually could be enrolled, and that there was some selection process based on assessment, largely ability to tolerate the protocol, that may have occurred, and there may also have been selectivity, and this is something I don't know about this protocol, but I would like to know.

There may also have been selectivity in that I believe that at certain times there was a delay of perhaps two or three months for when patients were initially proposed and considered to when they could be enrolled, so that in the presence of a rapidly progressive disease, which may have median survival of six to eight months, there was an additional factor that you had to not only meet eligibility criteria at the time of proposal, but then two or three months later, when ability to enter the protocol was there, you still had to meet the similar criteria.

Was that in fact the case?

DR. ROSENBERG: In general, there would only be a few week delay at most. These are obviously patients that have widespread metastatic disease, quite anxious about their situation, and I can't think of a situation where a two-month delay would have taken place because of an inability to put a patient in the hospital, for example. I

would think that is generally, probably about a two-week delay.

Now, when it comes to selection again, we take all comers. If we don't take a patient, it is only because we did not have an open slot. In our clinic, we can only take four new patients a week, but then it is just the first referral that comes in. It is really first come, first serve.

The patient selectivities are strictly on the basis of performance data, as well as the elimination, as written in the protocol, of patients that have evidence of ischemic heart disease, elevated creatinines or elevated liver function.

I would just take an opportunity to mention that in our own experience now, fairly extensively, with interleukin-2, but also with aggressive combination chemotherapy regimens, that the critical difference when one uses high dose interleukin-2 in patients with metastatic melanoma, is the durability of the complete responses, which I have not seen with any other treatment.

If you achieve a complete response, then the likelihood that you will ever recur is small. In our own experience, far less than half of the complete responders have ever recurred, and our longest responders now with

multiple sites of metastatic disease are out over 12 years.

So, I want to focus on the durability of complete responses, which I don't think you can achieve with any other available treatment.

DR. ATKINS: Within the Cytokine Working Group, there was very little selection bias except for the selection bias that is inherent in a patient being motivated enough to come to a referral center, but that is not much different than for the patients we would treat on other chemotherapy protocols.

DR. SWAIN: Are there any other questions from the committee? Mr. McDonough.

MR. McDONOUGH: Having had melanoma, I am Stage III. I am 4 years 9 months from surgery, 4 years 6 months from ending treatment on interferon. You are talking about the randomized testing. I went through that because at the time that I took interferon at the University of Pittsburgh, I was put in the computer mix, and it was high dose, low dose, no dose. I prayed for high dose. What would be the point of me going into a program and then coming out with a placebo?

I wanted to fight for my life. So, that is my vote on the randomization and testing. I don't know how you would achieve -- and I understand where he is trying to go

with lower doses do not have as much toxicity and the other problems involved.

I am asking these questions as a patient, and there may be some patients out here that would like to ask these questions. You are talking about an entry criteria. How stringent is that for the pulmonary and the cardiac? I mean you say stringent, but I don't know what that means, because I am 66, I am looking at people up there have a median age of 41. I am not going to blow up the balloon or I am not going to do the treadmill as well as they do.

DR. KUNKEL: The entry criteria from the standpoint of cardiac and pulmonary, which are probably the two most relevant to age, are be able to go on the treadmill and not have evidence of ischemic changes that you are about to potentially a myocardial infarction and therefore, would not tolerate low blood pressure, and to have pulmonary function that is what we call a FEV-1 of 2 liters, which is about half of what the normal FEB-1 might be for a patient.

MR. McDONOUGH: If I would flunk those tests, Doctor, could I waive it as a patient, the requirement? In other words, I want it anyhow. Could I say to you treat me, I will sign a waiver or whatever?

DR. ATKINS: If interleukin-2 was only being administered as part of the protocol, you wouldn't be able

to participate in the protocol. I guess if it were commercially available, you could then discuss it with your physician.

MR. McDONOUGH: Last question. I am sure you have had people in the study that have been treated by interferon as I have. From what I read up there, or what I saw, and what I have read here, I am least able to respond because I have had interferon? In other words, I wouldn't do as well as somebody that hasn't, or did I misinterpret that?

DR. KUNKEL: Actually, the study wasn't designed to look at patients who had prior interferon, although we subsequently did an analysis on that. There were only a few patients who had received interferon on the study. Although the response rate is higher in those patients who have received no therapy, we still saw responses, and we saw durable responses, in patients who had received interferon, but there were few patients on that study.

So, I hope that clarifies it.

MR. McDONOUGH: Last question. If approved, this would be available in Pittsburgh?

[Laughter.]

DR. ATKINS: It is currently being offered in Pittsburgh, and certainly if approved, would be available in Pittsburgh, and that is certainly one of the centers that

has a lot of experience.

MR. McDONOUGH: The protocol is going on still in Pittsburgh? I mean if I would relapse, I could get in at Pitt? This is all selfish. I realize it.

DR. ATKINS: I see the doctor who runs the treatment at Pittsburgh in the back nodding, but I would hope you would never be in that situation.

MR. McDONOUGH: After having been exposed to John Kirkland [phonetic], it is good enough for me.

DR. SWAIN: Are there any other questions from the committee?

If not, we will take a break and come back at 10 o'clock.

[Recess.]

DR. SWAIN: Next, we would like to proceed with the FDA presentation by Dr. Litwin.

FDA Presentation

[Slide.]

DR. LITWIN: Good morning. I am Dr. Stephen Litwin. I will present the CBER review for aldesleukin IL-2 as proposed for the treatment of metastatic melanoma.

[Slide.]

There were a number of CBER staff who participated very actively in this review. They are listed on this

slide. I understand I missed the opportunity to comment on Chiron's follow-up data. We received that data one day before this meeting, and although I really appreciate the confidence that the company has in us, it is enough time to work.

[Slide.]

Starting in 1985, and proceeding for about eight years, a series of studies were done which explored the treatment of solid tumors with IL-2. In January 1990, the Biologic Response Modifiers Advisory Committee, actually, the first meeting of it, discussed and considered the use of IL-2 for treatment of renal cell carcinoma.

They requested that the sponsor return with further information. A year and a half later, in 1992, they again reconsidered this. At this time, there was more data. They had moved from 174 to 255 patients. They were multicenter studies, and there was further follow-up data, and there was a favorable recommendation.

Approximately five months later, FDA licensed IL-2 for renal cell carcinoma.

[Slide.]

The current package labeling reads, "IL-2 induces durable complete responses or partial responses in a subset of renal cell carcinoma patients." The labeling further

urges that there be very careful patient selection and that cardiac and pulmonary function testing be done.

The rationale for approval by the FDA and the committee, and taken from transcripted records of the committee, there was less enthusiasm for the superiority or even comparability of IL-2 for renal cell carcinoma because of the single arm format of the study and the fact that there is a high reported incidence of spontaneous regression in renal cell carcinoma.

Rather, the favorable result was based on three factors pointed out at the time: the presence of durable remissions in the patients, the extensive regression of the tumor burden in these patients, and finally the fact that bulky disease would respond.

[Slide.]

The sponsor now proposes that there be a supplementary or additional labeling as follows: "Aldesleukin is indicated for the treatment of adults with metastatic melanoma."

The same dose and the same route of administration would be used. I won't go through this, you have heard it twice already.

[Slide.]

I will begin the first section of my presentation

which will deal with the experimental design and the study population and the results. I should point out that the data is very straightforward. CBER concurs fully with the sponsor in the endpoints and data presented, and the methods of analysis chosen. You have seen the data in the briefing packages, and you just heard it very clearly presented by Dr. Kunkel.

For those reasons, for all of those reasons I am going to try to concentrate, not on the description, but to a greater degree on those elements which we consider important for evaluation.

To start with the experimental design, 8 studies were integrated into a single-arm database of 270 patients. The data was obtained from a number of sources. There were 3 databases that represented information collected at the time the studies were done.

In addition, there was a retrospective audit by the company of all of the patients who responded, 43 responders, and there was also some audits of about half the patients for collection of safety data.

There were two doses used. The 720,000 IU/kg, I will refer to as the high dose. It was used in 147 patients. The 600,000 IU/kg was used in the extramural studies, the higher dose in the intramural studies.

There was no dose reduction. Doses were withheld for adverse events, either Grade 3 or Grade 4. There were 291 patients who were registered, 21 were considered ineligible, leaving 270 evaluable patients; 22, or about close to 10 percent, discontinued the study prematurely. I will discuss this further under safety.

[Slide.]

Moving on to the study population, the eligibility requirements were for histologically proven metastatic melanoma and patients who had failed standard therapy. They also had to have measurable lesions. Standard therapy was not defined.

[Slide.]

Initially, in the studies, the eligibility criteria were limited to those usual chemical and hematologic values that you usually find in baseline studies and illustrated on the first two lines, for Studies 0054, 0097, and 0053, which are all intramural studies. The number of patients is shown here.

In Study 0063, which is an extramural study, pulmonary function testing was added. Approximately five months later, in Study 0170, cardiac function testing was added included a stress treadmill and the presence of no evidence of heart disease.

Finally, about a year and a half after that -- this is in 1988, I believe -- the cardiac function testing was intensified with a thallium stress test.

Although the eligibility criteria were not completely specified -- and I think this issue has been discussed actually -- were not fully specified at the initiation of the study, there is I think ample reason to think that there was more intensive testing before eligibility was determined.

[Slide.]

In the study population, the stage of the disease at the time of entry to the study, using the current AJCC criteria, were not specified.

CBER did an analysis of 6 out of the 8 studies, 6 of the largest of the 8 integrated studies, and we found that 24 percent of the patients had sites of disease that were confined to cutaneous of subcutaneous or lymphatic sites. In the face of the lack of anatomic localization, these patients cannot be distinguished as either Stage IIIB or Stage IV, Stage IIIB being disease between the primary site and the metastatic site, and Stage IV being distant metastatic disease.

Once again, I think the discussion has already taken place. There was some issue as to whether any

selection, degree of selection bias may have been encountered in selecting these patients. We have no written information in what was submitted to us, whether patients were enrolled as they came through the door and met eligibility requirements or if some indirect selection bias was introduced, for example, by the ability of patients to be motivated enough to travel long distances to appear at the NCI for the initial screening.

[Slide.]

Turning to outcomes, there were no prospective endpoints designated. The sponsor has chosen to emphasize response rates and median duration of response as the critical endpoint for efficacy. CBER agrees that these are the appropriate pieces of data to be looked at in determining the results.

The response rate was 16 percent, that is, there were 43 responders out of 270 patients. Of these, 17, or 6 percent, were complete responders and 26, or 10 percent, were partial responders.

The duration of response for all of the responders was 8.9 months. The ranges are shown here. More importantly, for the partial responders, the median duration of response was 5.9 months.

[Slide.]

This rather busy but important slide takes up the issue of the consistency between the studies that were integrated into the database. I have listed five studies here. These are the studies that had most of the patients.

As a matter of record, let me indicate that the Chiron study, which was 5 patients and is not shown here, had 2 responses, 1 partial and 1 complete.

Turning now to the comparison that is shown in here, using selected features arising from the study, demographic features between these five studies were relatively similar. They are not shown.

The percentage of patients with visceral disease is shown on the first row of data. The percentage of patients with 2 or more sites of involvement is shown on the second row. Total IL-2 refers to the cumulative dose of IL-2 delivered. Once again, one can compare the sites, and they are reasonably comparable.

The last row shows two pieces of data in each cell, the complete response and the partial response. The number of overall responses, that is, objective responses, in these five studies ranged between 13 percent and 18 percent.

The conclusion from these studies, as far as they can be taken, is that the studies are reasonably comparable

and can be compared to determine efficacy and to collect safety data, despite the differences in doses. These three studies were intramural and given at 720,000 unit individual dose. Differences in time of performance ranging over 8 years, and the questions that arise about eligibility criteria.

[Slide.]

Turning to safety, virtually all the patients had severe adverse events. 95 percent of patients had Grade 3 adverse events, 35 percent of patients had Grade 4, life-threatening adverse events.

Doses withheld in most patients. We took a look at the number of patients who received at least 28 doses per course, which would be the maximum number that could be given. There were 18 of the 270, a little over 5 percent. All the remainder of the patients had doses withheld.

One can gain some idea also of the role of tolerance in this because if one compares the median dose, that is, the median number of doses for the first course, and those patients who received the high dose, that is, the 720,000, there was 16 as compared to those who received the lower dose, which was 22.

There were 22 early terminations, 16 were due to toxicity, 5 patients refused further participation. All had

listed adverse events. There was 1 patient who chose alternative therapy and left the study.

There was a 2 percent death rate, 8 patients died on study, 6 of whom were considered to be drug related. I will discuss that in a moment.

[Slide.]

The adverse events are listed here. This is the cumulative experience for Grades 1 through 4, for both the renal cell carcinoma data and the current metastatic melanoma data, and there are 25 patients involved.

Despite the fact that the adverse events seemed to involve virtually every body system, there were 4 major types of events, cardiovascular, pulmonary, renal, and, on the next slide, sepsis, which seemed to dominate.

About three quarters of the patients experienced hypotension, which in many cases required pressor agents, other cardiovascular problems included arrhythmias, both supraventricular and ventricular, and a more amorphous listing called cardiovascular disorders.

There seemed to be a coupling of toxic problems, often hypotension and, in some cases, followed by oliguria and anuria, and in some instances, pulmonary distress, suggesting that there was some series of events cascading.

The explanation of hypoperfusion of organs has

been forwarded.

[Slide.]

This is a further listing of adverse events of various body systems. Let me call your attention to the infections on the bottom. Although the percentage is relatively low, Grades 1 through 4, considering some of the other data that I have presented, infections were involved in 5 out of the 6 deaths, according to Dr. Kunkel, actually 6 out of the 6, a little hard to tell.

[Slide.]

These are the 6 IL-2 related deaths. These have been extracted from the clinical precis, a more full version of which is included in the briefing document. Once again, sepsis was involved in virtually every one of these deaths.

Also, there was a picture of multiorgan failure, which is somewhat hard to put together, associated with this. It should be noted that of the deaths, all 6 occurred at extramural sites. There were no deaths in the intramural sites.

[Slide.]

For purposes of comparison, I have listed here the IL-2 related deaths associated with earlier studies of renal cell carcinoma. They are, 2 of them, myocardial infarction, 1 cardiac tamponade, 2 episodes, 2 deaths due to sepsis, 1

due to bowel perforation and sepsis, 1 due to GI bleeding, 3 pulmonary complications, and 1 unknown.

[Slide.]

I would like now to move to the second section of my presentation, which focuses on a series of four review issues, which we think are central to evaluation of the data. They are the consistency of the eight studies, the definition of the patient population, the durability of the response, particularly in the partial responders, tumor regression, and some further analysis of the PR data, and finally, the prognostic variables.

I have already presented essentially the data with respect to consistency and the comparison of those five out of the eight studies for cardinal features.

I will move on to the definition of the patient population.

[Slide.]

In the first two rows of data, we can see that the age and the gender distribution is consistent, reasonably consistent with published demographic descriptions. The ECOG status and percentages of ECOG, PS 0, 71 percent, 1 and 27 percent ECOG 2, 2 percent.

Sixty-nine percent of the patients had visceral involvement, 71 percent had 2 or more sites of disease.

These data suggest that this is an advanced metastatic melanoma population. However, there are limitations in a very large part of the data. As mentioned earlier, the stage of the disease at the time of entrance into the study was not specified.

There is also limited information on prior therapy, on the type of prior therapy, for example, the breakdown of immunotherapy into the exact type of immunotherapy, the duration of prior therapy, the response to prior therapy, and the duration of that response.

This data is present partially for some of the responders, although not adequately, and is virtually absent along with tumor burden data for the non-responders.

[Slide.]

We turn now to the third issue, that is, durability and tumor regression.

[Slide.]

These are the complete responders. I have listed the complete responders who at the time the study was completed at the end of 1966, were still in remission. There were 10 of the 17.

Those who relapsed were 7 of the 17, and they are listed in the second group.

Among the complete responders, there were very

durable remissions. They ranged from 8 years plus 8 years to the last one, which is still 2 years. Among the complete responders who have relapsed, there were 3 which were relatively shorter. It depends on one's definition of a durable response. The data, I think, by and large, support the point of view that the complete responders showed durable responses.

[Slide.]

The picture is different for the partial responders. Of the 26 partial responders, only 3 were in remission at the time the study ended, and 23 had relapsed. Of those 3, one was quite long, at 7-plus years, another at 4 years. The third patient elected to have intensive chemotherapy and a bone marrow transplant, and was censored from the study.

If we look at the relapsed partial responders, you can see that there are many who have 1, 2, 3, 4-month limited durations of response. I think the data speak for themselves.

[Slide.]

This slide looks at the degree of regression of the tumor in the partial responding patients. The data has been organized into cohorts of sorts, so that on the first row of data, you can see those with 90 percent or more, and

that includes 7 of the 25 evaluable patients. There was 1 patient of the 26 partial responders in which there was no baseline data, and we could not evaluate.

80 to 89 percent response of 5 patients, et cetera.

At this point, the last two slides that I have shown, the durability of the responses and the partial responders, and the degree of regression of the tumor permits a comparison of the current data to those that were originally promulgated for the renal cell carcinoma licensure.

As I mentioned earlier in my discussion, the three major factors that were involved in the favorable response, the favorable result for the renal cell carcinoma included the durability of the response, particular in the responders, the degree of tumor regression, and finally, the ability of bulky tumors to respond.

In comparing them, let me start with the renal cell carcinoma. The durability of the renal cell carcinoma as mentioned by Dr. Kunkel now exceeds 20 months median duration of response. The durability, not shown on this slide, the durability of the response for the partial responders for metastatic melanoma is 5.8 months in comparison.

With respect to regression of the tumor burden, it was reported that for the renal cell carcinoma, over half, 15 out of 28 of the partial responders, had over 90 percent regression of the tumor burden.

In the metastatic melanoma patients, shown on this slide, the comparable figure is 7 out of the 25.

In both the renal cell carcinoma studies and in the current studies for metastatic melanoma, it seems evident that both bulky tumors can respond to the therapy. All in all, the responses seen here, particularly among the partial responders for the melanoma, seem much less dramatic than those noted for the renal cell carcinoma.

[Slide.]

A series of analyses are shown on the next two slides and summarized on the third of the partial responder group. Rather than list the average data for this group, what I have chosen to do is show just 7 of the 26 partial responders, and these are the partial responders who have the longest duration of response listed in order of longest duration on downwards.

The points that I will make, the next five or six points that I will make, however, although illustrated by this data, are equally valid for all 26 of the partial responders.

Let's begin with there is considerable tumor burden as noted from these cross-product measurements. Bulky disease can respond, but there is no necessary relationship between the tumor burden, at least that we could elicit, and the response.

Finally, the response duration and the tumor burden, when inspection is used, not quantitative, but inspection of the 26, shows that there is no obvious relationship between those who have long responses and less tumor burden or more tumor burden.

[Slide.]

This is the same 7 patients along this partial responders. The second column of data shows the number of days from the first dose of IL-2 to the declared objective response. The median for this was 133 days with a very long range starting at about 30 days and reaching well over to 800.

Almost all of the responses occurred after the first course of therapy, although repeat courses were given in some patients. In inspecting the data for the responses -- and I am talking about inspecting the serial data which is not shown here -- it was obvious that even after a partial response was declared, there continued to be regression of the tumor burden, often at a very late point

in time. This seems characteristic of the response.

The final column of data lists the number of days that patients received IL-2 after their partial response, their objective response was declared. You can see in this slide 5 out of the 7 that I show you. Among all 26 partial responders, actually 20 out of the 26 continued to receive IL-2 after they were in remission.

The impact of this with its attendant need for hospitalization and intensive care unit, and the attendant exposure to toxicity, I think is evident, and must be considered particularly for those partial responders of relatively short duration.

Other studies were also done. I will go through them relatively briefly. I do not have them illustrated. As I think mentioned earlier, both visceral and non-visceral sites of disease appeared to respond. On inspection of the data, there seemed to be no predilection for particular tissue sites that we could discern.

Of the complete responders, 2 out of the 17 had liver lesions which seemed to be somewhat underrepresented. Of the partial responders, the number was higher. It was 12 out of the 26 liver lesions would respond. There was only 1 patient with CNS disease. I believe those lesions also responded.

The degree of response in the lesions was also looked at. In general, patients who responded, all of the lesions would respond relatively comparably. There were a limited number of patients in whom some of the sites of disease would not respond. They are listed in the briefing package.

[Slide.]

To summarize the data that I have just presented, there are some partial responders who have a long duration of response. I have shown you the 7 patients who have over a year duration of response, but most partial responders do not.

There is a substantial tumor burden. Bulky tumors can respond to IL-2. The response duration does not appear to be related to tumor burden. Responses occurred at a median of 133 days and usually after the first course of IL-2, and 20 out of the 26 partial responders received IL-2 after their objective response.

[Slide.]

I will turn now to the last review issue, that is prognostic variables.

[Slide.]

I have listed first those factors that are not associated with the reponse. They include age and gender,

visceral involvement, yes or no, dichotomous, and the number of metastatic sites, at least in our hands.

[Slide.]

Of all of the analyses that were done of the variables, most of which were done by the sponsor, but some also by CBER, only two stood out as having a relationship to the tumor response, as pointed out by Dr. Kunkel. That is, the ECOG status and prior systemic therapy.

Patients with a performance status of ECOG zero had a 19 percent response rate as compared to 9 percent. Patients who lacked, who did not receive systemic therapy -- and we had grouped all of the systemic therapies -- which were 147 had not received any systemic therapy, had a higher response rate, 21 percent versus 10 percent.

Using an odds ratio, these figures are significant. The numbers did not cross 1.

[Slide.]

Once again, returning and reorganizing the data that I have just presented on ECOG status, in patients -- looking at it slightly differently -- looking at the first row, 71 percent of patients for ECOG performance status zero, 29 percent for 1 or 2. As I mentioned earlier, only a small percentage were 2, only five patients or 2 percent.

The response rate in the performance status zero

patients was 19 percent versus 9 percent. The death rate in the performance status zero patients was less than 1 percent versus 6 percent in the patients with ECOG PS 1 or 2.

If you look at this vertically then, a patient who is ECOG PS 1 or 2, had a 9 percent chance of responding, and a 6 percent chance of dying. We consider this data as strongly suggestive of a trend, which should prompt a review of the inclusion of these patients in studies.

On the other hand, these data should be interpreted cautiously. The number of patients with a PS 1 or 2, with respect to the response rate, is relatively low. In interpreting the number of deaths, it must be borne in mind that all of the deaths, 6 out of the 6, occurred in the extramural sites, and the extramural sites enrolled a higher percentage of patients with PS 1 or 2 ECOG status, so that the correlation that we may be looking at here could equally well be with the site of performance or the experience of the principal investigators, as well as with the ECOG status. That issue cannot be settled.

[Slide.]

I am going to more or less skip these. Dr. Atkins has really touched on these points much more deftly and in much more detail than we have.

We have pointed out the primary use of DTIC. The

response rate is often reported from 15 to 20 percent, a minority of the responses are complete responders.

Let me just make a general statement with respect to combination therapies that have been tried of all sorts. In general, although the data on the response rate is more promising, is higher, the number of complete responders remain low and as far as one can determine, the median durations or responses are also short.

In general, the current treatment for metastatic melanoma at this point is unsatisfactory.

[Slide.]

This simply is the second slide, which I will not get into.

[Slide.]

The last three slides summarize this presentation. First, the issues regarding the study population, that is, the definition of the study population, the definition of those patients who could benefit most and would have the least toxicity from the application of this agent.

It is very limited. There are a number of problems. I pointed out these again, the issue of patient selection has been discussed, the lack of staging by current staging methods. The dose given is difficult to determine.

First of all, because of the use of two different

doses in four and three of the studies respectively. Also, there was a great deal of variation between patients even in the same study with respect to dosing, because of the tolerance issue, and also the fact is that some patients received more than one course of therapy.

The impact then of dose effect is very difficult to judge. There is very limited data, I think I have touched on this point a number of times, particularly on prior therapy, and particularly in the remainder of the population characterized as non-responders.

[Slide.]

Issues regarding safety. The changes that have been made in the management of fluids and pressors, the introduction of prophylactic antibiotics, and the possible impact of changes in more stringent patient selection have been mentioned.

It is unclear as to what the impact of these alterations in the management, as logical as they appear. All of the 6 deaths occurred between 1987 and 1990. They were before the introduction of the antibiotics. There is no way to compare the death rate per year or in the studies with this limited data. There is no evidence that the changes in management have managed to deal with effectively the problems of severe toxicity.

ECOG 1 and 2 patients have higher toxicity and a lower response rate. That has been reviewed.

The risks of infections once again is unclear, and the etiology and the risk of cardiovascular adverse events is particularly uncertain at this point. Although there appears to be something there, they are certainly not fully known.

Finally, it should be mentioned that with a 16 percent overall response rate, there were 84 of 100 patients who unfortunately had to be subjected to the hospitalization and attendant toxicity that was involved without any discernible benefit.

[Slide.]

Finally, issues regarding efficacy. The majority of the partial responses are short, although there are some of duration. There is limited, though not insubstantial, tumor regression. There is a need for continuing therapy or at least as this protocol and studies were carried out after the objective response.

The major clinical value would be seen in the limited number of durable complete and partial responses, as I think emphasized by Dr. Kunkel, and in the availability of IL-2 as an alternative therapy.

Thank you for your attention.

DR. SWAIN: Thank you very much. We will open it up to the committee for questions.

Committee Questions to FDA

DR. SWAIN: Dr. Ozols.

DR. OZOLS: Is there any information on those patients who did not have any prior chemotherapy prior to getting IL-2, and then did not respond to IL-2, what their subsequent response rates were to standard chemotherapy?

DR. LITWIN: I don't have any information on that.

DR. KUNKEL: The follow-up data on the non-responding patients with respect to subsequent treatment was not obtained, however, I think their survival curve reflects what happened to those patients. They were eligible after one course of treatment with IL-2 to receive whatever available therapy was out there.

DR. RAGHAVAN: That was a nice review. Just to kind of set everything into clear context, you presented for us your overview of 250 patients, the vast majority of whom had a performance status of zero.

If you tried to summarize patient benefit from those 250 cases without specifying complete regression, partial regression, but just saying at the end of your analysis, what percentage of patients benefitted from treatment out of the denominator of 250, what figure would

you put on that?

DR. LITWIN: I would say of the 270 patients, 6 or 7 percent in the complete responder and the partial responder groups had durable enough remission, so that it would be no argument that anybody would have about that.

DR. RAGHAVAN: In the groupings that you looked at -- and you may have said it, I might have missed it -- there were clearly two dosage levels, the intra- and extramural dosing being about 15 percent different.

Were you able to assess differences in patterns of toxicity, requirements of hospitalization, that would give you any sense that even at that moderately high dose level, that there was a difference in the toxicity profile?

DR. LITWIN: We saw no differences in patterns of toxicity. The data on that point was limited, but we saw no differences. I will ask Dr. Kunkel if she has anything further to add to that, between the 720 and the 600.

DR. KUNKEL: What we did demonstrate is that those patients who received 720,000 tolerated fewer courses, so their toxicities would occur earlier in the cycle. It was 15 compared to 20, I believe, with the 600, but all patients, as Dr. Litwin has pointed out, really experienced Grade 3 effects, but they are reversible.

DR. RAGHAVAN: I actually do have one other

question while you are on your feet. Somewhere in the interaction with Dr. Litwin out of the room, and so on, there are some data floating around that I have just not captured that relate to late follow-up of patients who were lost.

You had 39, which is a significant proportion. You have tracked down 29 of them. Is it a reasonable assumption that those 29 patients died pretty close to the time that they were lost?

DR. KUNKEL: Actually, that is the truth. We were able to go back and locate death dates on all those patients. It didn't impact on our median overall survival.

DR. RAGHAVAN: But it would have brought your tail down some.

DR. KUNKEL: No, it didn't. Well, it would bring it down, I think 2 percent, 1 or 2 percent. You had 2 percent.

DR. SWAIN: Dr. Simon.

DR. SIMON: The fact that they did close to the time they were considered lost to follow-up would have the maximum effect, so it is not a comforting sort of issue, but I was going to ask a question to Dr. Litwin here.

This study basically shows durable remissions for -- well, there were 13 patients that are essentially still

going with long-term remissions, and you might make an argument that there were 20 patients who had some benefit.

So, it is basically somewhere between 5 and 7 percent of the patients had durable remissions. If you take basically DTIC or chemotherapy regimens for a similar kind of selection of patients, what percent would have long-term durable remissions? Is it zero, is it 2 percent, is it 5 percent?

DR. LITWIN: My understanding, which is only from the literature, and probably can be expanded by Dr. Atkins, is that it is a 15 percent to 20 percent response rate, and of those, about a quarter are complete responses, and those include most of the durable remissions, so we should be talking of a figure less than 10 percent, I think, and possibly much less than that.

DR. SIMON: Well, I mean for this series, it is only 5 to 7 percent, so if it is saying that it is less than 10 percent for DTIC, isn't saying anything really. I mean that gives us no basis for believing that this is any better than DTIC.

DR. SWAIN: Dr. Atkins, do you want to respond?

DR. ATKINS: I spent a lot of time looking at all of those articles that looked at DTIC and other combination chemotherapy, and paying particular attention to those that

were associated with greater than 5 percent, five-year survival, and there was always a tremendous selection bias for patient with surgically-resectable disease in that group of patients.

If you look at patients presenting with 70 percent, with more than one site of disease, comparable to this, the answer to your question is 1 or 2 percent in the ECOG and SWOG database.

DR. SIMON: With a proportion of performance data, zero patients similar to this series?

DR. ATKINS: Yes, exactly, and in the large reviews of DTIC alone that I presented, it is 1 to 2 percent, as well.

DR. SWAIN: Dr. Johnson.

DR. JOHNSON: That sounds like a reasonable figure, 1 or 2 percent. It also has to be tempered with the fact that none of those patients had a thallium stress test prior to being deemed a PS zero. So, even if you make the groups comparable on PS status, there is still an inherent selection here that goes towards selecting out, in this series of patients, a much better, more fit patient.

So, if you want to really make -- I mean if it is 1 or 2 percent, and let's assume that, and this is 5 percent, that additional 3 percent could be on the basis of

that selection factor, it seems to me. I hope it is not, but that is what we don't know.

DR. SWAIN: Ms. Beaman.

MS. BEAMAN: Maybe it was stated earlier, but I would like to hear again a comment on the general quality of life rating for those patients who discontinued the study.

DR. KUNKEL: A formal quality of life wasn't obtained on this study, but I think probably that that question is best addressed by the investigators who took care of the patients and, say, saw them in follow-up upon completion, so we will turn it over to Dr. Atkins first.

DR. ATKINS: Well, most patients experience significant side effects that required them to be in the hospital while they were receiving therapy. These side effects usually would begin to resolve as their treatment stopped. Patients would be well enough to be discharged within 1 to 2 days after finishing therapy.

By 3 or 4 days after finishing their first week of therapy, they would be well enough to get up and out of the house, and be eating again, and by the time they came back in for their second week of therapy, they would be pretty much returned to their baseline, get their second week of therapy, which would be, say, days 15 to 19, and then they would have a similar or maybe a day or two longer recovery

from that, so for about five weeks of time, they were either in the hospital or recovering from therapy.

Then, they would get evaluated a few months later and if they were getting benefit, they might get more therapy. If they weren't showing benefit, that was the end, they had their shot. The only durable effects of interleukin-2 therapy were vitiligo, which happened more frequently in the responders, some thyroid dysfunction, which was usually resolved spontaneously by 10 months, and in the subset of patients presented here, tumor response.

DR. KEEGAN: In addition to that, we might add that there were two patients with myocardial infarctions and one patient with an amputation.

DR. JOHNSON: What was the amputation?

DR. KEEGAN: It was for gangrene.

DR. SIEGEL: I think there were similarly in the renal cancer database, a small, a limited number of patients, but a real number of patients with ongoing toxicity effects.

DR. LITWIN: I should add that the data that we are talking about is for the 22 patients who discontinued the study. There was almost no data on the remainder of the patients, and I don't think that is a large enough group to be evaluated for the residual of this. It is hard to

believe that patients who suffer from oliguria, anuria, supraventricular tachycardias, et cetera, would not have some damage to an organ system.

DR. ATKINS: I was speaking from my personal experience of 10 years of treating these patients, and I have not seen any durable side effects except for what I have mentioned. I don't have personal experience with those two patients that were described in the 22, but it is really true. As a matter of fact, even we send patients out with elevated creatinines, maybe four or five, and we don't recheck them before we bring them back four weeks later for an I.V. contrast CT scan, because we know that it is going to be fine, and it always is.

DR. SWAIN: Dr. Raghavan.

DR. RAGHAVAN: While you are on your feet, Mike, can I just get you back to one question I asked, because it is kind of still bothering me a little bit, and it probably reflects my lack of knowledge of this area.

The issue of dosage. You answered my previous question by saying that there was a big database out there of low-dose IL-2, and I thought I heard you say that most of that was in combination studies done with interferon, other chemotherapy, and so on.

Could you just clarify for me is there a nice,

clean, well-conducted set of trials done by investigators of the Cytokine Working Group caliber, if I can call it that, intramural, extramural-intramural associated, whatever, that is just a very clean Phase II assessment of, say, half-dose IL-2, and what I am getting at is obviously the very important issue of cost, that if this goes through, while I know that the sponsor is going to be altruistic, my previous experience in life is that more costs more, and so I would like to be clear in my own mind that 600 or 700 is the level that is required.

The thing that is worrying me is the issue that I don't personally view chemotherapy as wonderful therapy for melanoma, and I wonder to myself if the studies were done in combination, is there the potential that low dose IL-2's effect has actually been vitiated somewhat by the association with other treatments. That is what I was trying to get at before.

DR. ATKINS: I was talking about combinations with interferon or with vaccines where I am not aware that there would be any detrimental effect. The Cytokine Working Group has looked at other schedules. We gave interleukin-2 by a continuous infusion which uses a lower amount of IL-2, probably about 20 or 30 percent of the amount of IL-2 although it does have toxicity of its own that requires

inpatient therapy. I think Dr. Dutcher was the first author of that paper, and there was a three-year, 5 percent response rate in that study.

DR. SWAIN: Are there any other comments? I guess we should go on to the questions.

Committee Discussion

DR. SWAIN: The first question. This license application describes the results of eight studies, enrolling a total of 270 patients, treated with a comparable dose and schedule of IL-2. Approximately 70 percent of the study population had visceral disease and more than one site of metastatic disease, 74 percent of the patients had ECOG PS 0 at baseline and all met stringent entry criteria regarding cardiac and pulmonary function.

The pooled data revealed an objective response rate of 16 percent and CR rate of 6 percent. The median duration of response for patients achieving a PR was 5.9 months; 10 of 17 complete responders remain in remission for over 2 years. The objective response rate for other single agents in this disease ranges from 5 to 25 percent with CR rates of 1 to 4.5 percent. Median response durations for CR patients treated with other single agent therapies has been up to 15 months.

The first issue is please discuss the type and

quality of the responses observed; and secondly, the population treated in this pooled data set.

Do you want to start with Dr. Ozols?

DR. OZOLS: We actually spent about an hour in doing this, this morning I guess.

DR. SWAIN: Summarize it.

DR. OZOLS: I think the issues regarding the patient selection aren't answerable. I mean there clearly was patient selection that went into place, but what we are hearing is that the majority of patients with metastatic melanoma would, in fact, be eligible for this type of treatment, and I think with the same caveats that we used for using this treatment in renal cell, I think we can select an appropriate population.

The quality of the response I think is the most compelling clinical aspect that we have heard. The duration of the response is very meaningful and very significant, and clearly associated with clinical benefit. I mean the response rate is obviously low, but the duration of the remissions is impressive.

DR. SWAIN: Does anyone else have anything to add? I think we all agree with that.

The second part is considering the rate, quality, and duration of response, can one conclude that IL-2

provides clinical benefit for patients with metastatic melanoma?

Dr. Raghavan.

DR. RAGHAVAN: I think that there is clear evidence that there is clinical benefit for selected patients and I think that, as Dr. Ozols said, the thing that is most impressive is the duration of responses. It has got to be held in the context that the alternatives that such patients have are very limited. So, I think the answer is yes.

DR. OZOLS: I agree.

DR. SWAIN: Is there any further discussion on that point?

We need to vote on that. Everyone who feels that there is clinical benefit from IL-2 for patients with metastatic melanoma, raise their hand.

[Show of hands.]

DR. SWAIN: It is unanimous.

The second question is about the toxicity. 95 percent of the patients experienced grade 3 toxicity and 35 percent grade 4. Treatment required hospitalization in an ICU setting during the IL-2 administration and in the post-infusion period. The treatment related mortality, 6/270 was not dissimilar to the treatment related mortality

of 11/259 in the renal cell studies.

Mortality was disproportionately higher in patients with a performance status of 1 to 2 versus 0. A logistic regression analysis indicated that the performance status of 0, lack of prior systemic therapy, and greater number of IL-2 courses administered correlated with a higher response rate. Current labeling for use in metastatic renal cell cancer restricts use to ICU facilities and to patients with normal cardiac and pulmonary function and notes that response rates were higher and mortality rates lower among patients with a performance status of 0.

Please discuss the toxicities of IL-2.

Would anyone else like to add any comments? I think we have discussed that pretty thoroughly. Any other additions?

DR. SANTANA: Could I ask a question?

DR. SWAIN: Sure.

DR. SANTANA: After the postmarketing of IL-2 for renal carcinoma, does the sponsor have any data whether the use of IL-2 in that setting, the intensive care units outside of clinical research units, has changed a different mortality figure or complication rate?

DR. SWAIN: Dr. Litwin.

DR. SIEGEL: It is very hard for us to have

denominators. Occasional events get reported, but in the postmarketing period it is hard to ascertain both the efficiency of reporting or the denominators in terms of usage. I am not sure we can get a good handle on that.

As has been indicated, over the period of studies, we did look at mortality rates in particular, and there were 2 or 3 or 4 deaths each year, '87, '88, '89, '90. As the fluid management and identification of patients clearly improved, the antibiotic prophylaxis started later, but there a very small portion of the database is in patients who are after '90 or so.

DR. KUNKEL: We need 40 and 43, 44.

[Slide.]

We were actually interested to see how the deaths related to the enrollment on these studies, so what this summarizes is the number of patients enrolled by years between 1985 and 1993. The red represents septic episodes, and as you can see, there was one episode in '87, two in '88, two in '89, one in '90, and then we did not have any episodes of death related to the sepsis-like syndrome in '91, '92, and '93.

[Slide.]

When we look at our postmarketing, which is slide 43, and as Dr. Siegel mentioned, of course, we don't have a

denominator for this. These are postmarketing reports of sepsis. This isn't necessarily reports of sepsis with the regimen that we are bringing to you today, but includes continuous I.V. and short I.V. infusion, as well as subcutaneous.

We do have 20 reports of sepsis and there were 7 deaths between that time period of '92 to '97 attributed to a sepsis-like syndrome.

What we also looked at was whether or not these patients had received prophylactic antibiotics, and none of the deaths associated with the sepsis-like syndrome had the patients been on prophylactic antibiotics.

So, that is what we know about our postmarket surveillance.

DR. SANTANA: We all recognize that is underreporting, too.

DR. OZOLS: I think you can also add to that what Dr. Rosenberg said, in the last 700, 800 patients, in their database they did not have any deaths, so I think the learning curve was clear, and the management of these patients has improved and the selection to some degree, but I think the death rate is markedly decreased into the acceptable.

DR. SWAIN: If there is no more discussion on that

point, in view of the responses and the toxicities, should IL-2 be indicated for use in metastatic melanoma?

Dr. Ozols.

DR. OZOLS: Yes.

DR. SWAIN: Dr. Raghavan?

DR. RAGHAVAN: Yes.

DR. SWAIN: Any other discussion? Okay.

So, we will take a vote and all that say yes, raise their hand to the question, should IL-2 be indicated for use in metastatic melanoma?

[Show of hands.]

DR. SWAIN: It is unanimous.

If approved, should the label further restrict the use of IL-2 to specific populations, such as a performance status of 0? Dr. Vose.

DR. VOSE: I would just like to make a comment that although there were not many of the patients that were lower performance did respond, there were a few patients that did benefit from that, so I don't think that one should restrict the labeling, but just as the precautions that are already there, that there is a higher response rate in patients with better performance status, and let the physician use their proper judgment.

DR. SWAIN: Mr. McDonough.

MR. McDONOUGH: If it came down to it for me, and I were in the 1 or 2 group, I would still want the right to take my shot.

DR. SWAIN: Does everyone agree or have any other comments? Dr. Krook.

DR. KROOK: I will play the devil's advocate with you only because I treat patients and there are times when you have to turn to a patient and say the side effects are worse than the treatment.

Now, you and I can go around and around, and I would probably bend to that, but I would give you a good argument that when I see the toxicity here, most of use who are treatment physicians don't like to precipitate death and all the problems, at least those of us who have been in practice for a while don't.

You may have the final say, but I certainly would talk to you a bit.

MR. McDONOUGH: Just a quick response to that. If I am PS1 or PS2, death is right around the corner anyhow, and if there is any chance whatsoever, it is like a on-side kickoff in the Super Bowl, I mean, you know, I am going to take that shot to save my life.

DR. SWAIN: Dr. Johnson.

DR. JOHNSON: I have had some trouble with this

application, but obviously agree with the majority to this point, and I appreciate Dr. Vose's comments because, as a clinician who takes care of patients is what I do, I like to have some leeway in decisionmaking.

This doesn't suggest that one doesn't have leeway by restricting it. It certainly gives a major indicator of precaution, and I think we have spent the better part of the morning talking about the high selectivity of this patient population, and I personally think we should include a restriction in this fashion, and as additional data are gained, and as that learning curve is improved, my suspicion is that we will gain additional information. I would be shocked, stunned if the sponsor weren't going to look at those data as those data came in over time. I would be equally shocked, stunned, if physicians didn't bend the rules a bit. I certainly agree with Dr. Krook's comments regarding patient decisionmaking and physicians participation in that activity.

It is always important to give the full story, and many physicians, most physicians in my estimate, don't want to participate in a patient's demise. Sometimes doing nothing is a wiser therapeutic choice than doing something, and I think that goes back to the first rule of medicine.

DR. VOSE: I understand and agree completely with

what you are saying, and my little bit of a concern about that is if we put it that way in the labeling -- and this is not going to be an inexpensive drug for those patients -- and I am concerned about the insurance companies going back and saying this patient should not have received that, and will not pay for it. So, I would prefer a very strong labeling as far as the increased response rate, but not to preclude those patients.

DR. SIEGEL: Let me simply comment that the way we have dealt with this to date in current labeling of this drug for renal cell carcinoma is not to contraindicate its use on the basis of performance status, but to note specifically that both under safety, that morality is higher with higher with a performance status of 1, as is the incidence of intubations, gangrene, coma, GI bleeding, and sepsis.

Those were selected ones, but selected obviously as side effects that happen to have significant impact and meaning, and similarly note that the response rate is lower, but at least to date have elected not to create either a contraindication or a specification in the indication in order to allow more physician judgment particularly given that ECOG performance status 1 can be very different from one patient to another in terms of assess ability to

tolerate toxicity.

That is where we are. We are asking if there is any reason to change that, but perhaps there is some consensus that that is where we ought to be.

DR. SWAIN: Dr. Raghavan.

DR. RAGHAVAN: I agree completely with Dr. Vose. I think that prescription here would just add to the complexity both for patients and for clinicians, and the reality is the ECOG performance status 1 is very much a function of the eye of the beholder, and I can't see the point of creating an interdiction on a prescription that is going to be hard to enforce and that will just force physicians, who have patients who want treatment, to commit perjury.

So, to my mind, given the fact that the database that related to performance status 2 was pretty slim, and, in fact, we have already said this is a very, very highly selective group of patients. It is the intramural and the extramural, and the reality is that most of the patients are extra extramural, and so we just can't make that knight's move of thinking.

The reality is we have identified the type of patient that is likely to have a chance of real benefit from this, and putting in artificial prescription I don't think

is going to help us, and I think giving the clinician and the patient, as Mr. McDonough said, some discretion, ultimately, it comes down to the patient being informed of what the risks are, looking at the lack of alternatives, and making an informed judgment, and you can do that on the labeling indication.

So, I think putting in artificial constraints would just complicate the issue and will create a problem for some patients where our very generous health funds will look at their performance status and create payment problems.

DR. SWAIN: Did you want us to vote on that?

DR. SIEGEL: I think we have a sense of most of the committee members already.

DR. SWAIN: Under the accelerated approval mechanism, drugs and biologics that have been studied for serious and life threatening diseases and "that provide meaningful benefit to patients over existing treatments" may be approved based on a surrogate endpoint that is reasonably likely to predict benefit provided post marketing studies confirm net clinical benefit. Under standard approval, post marketing commitments can be required of the sponsor for additional studies to optimize drug dosing for the patient population.

The question is: If there is an accelerated approval, what studies would be appropriate to confirm clinical benefit?

DR. SIEGEL: Let me clarify an issue or two about accelerated approval. I think this committee has had significant dealing with accelerated approval, and so this is probably a simple refresher.

For serious and life-threatening diseases, we will generally give a standard approval if there is evidence of clinical benefit that outweighs toxicities. Where a surrogate endpoint has been used, as we have done for a number of years, and as stated clearly in the Oncology Initiative of a year or two ago, we will accept surrogate endpoints with reasonable likelihood to predict benefit, and it was noted in that document, in that initiative, that responses in cancer based on experience, and this, of course, was discussed with this committee, were in many or most cases deemed to be reasonably useful as surrogates for clinical benefit.

That does not mean, however, that they cannot also be, as opposed to surrogates, as indicators of clinical benefit, and there has, as you know, been a history of giving full approvals on the basis of complete responses in particular, and durable complete responses, particularly in

refractory populations.

So, we potentially, coming in, could go either way with this. It is my feeling that on the basis of the remarks of this committee, and this is why I am putting this forward no, to hear if I am sensing wrong, but it would be my feeling based on the remarks of this committee that the nature, quality, and duration of these responses were, in fact, indicative of benefit of the complete and of the durable responses, not simply a reasonable surrogate that it was likely there would be benefit.

We would normally again, unless I hear otherwise, progress with a regular approval, and the implications regarding this question -- and that is why you are convened here -- are somewhat significant, because in the setting of accelerated approval, where there is proven benefit on a surrogate, the postmarketing commitments can require additional studies to prove that reasonable likelihood or predicting benefit indeed is true, that, in fact, there is a benefit conferred, whereas, in the setting of the more standard approval -- and there are stronger teeth in that and that withdrawal of the indication or of the drug can occur if the studies do not confirm benefit or if they are not performed with due diligence.

In the standard approval, we also speak to

postmarketing commitments about issues, such as dose optimization, optimization of the target population of patients and exploration of toxicity profiles, whatever those might be, those generally don't have the teeth of withdrawal behind them, nor of confirming efficacy, but just ensuring or promoting the ongoing appropriate development of the drug.

So, if I am sensing the committee right, it is more the latter type of trials that we would like you to be discussing as indicated by the second half of this question, what commitments for postmarketing studies should be sought.

DR. SWAIN: I guess I would just like to ask does anyone disagree with a standard approval?

Dr. Johnson.

DR. JOHNSON: I think for all the reasons that have just been stated, I don't a standard approval should be given at this point. I think an accelerated approval would be appropriate. I think without that, we won't get this information, and I think it is appropriate for us to make that clear to the sponsor, that this is something that we think ought to be done.

We are talking about a very, very small number of patients that will receive a very, very toxic and extremely expensive therapy, and I find it ironic that the committee

would worry about restricting PS because of access, but not worry about expending huge sums of health care dollars in a futile effort. I think that that is really the issue in my mind.

I want to ask one question now of the FDA. Has approval been withdrawn from any company that received accelerated approval, because we may be making a mistake here.

DR. SIEGEL: I don't recall drugs that have been withdrawn from the marketplace. Indications have been withdrawn. That is what we are discussing here in the sense that this has another approved indication. Indications have been withdrawn I know in the case of antiretroviral nucleosides. There were some indications, some combination uses that didn't pan out in further studies that led to change of labeling, but that has not been a common practice. The typical practice with accelerated approval is that postmarketing studies have been reasonably well completed and usually or in virtually all cases, have obviously, if they yield new information, not entirely predictable, but have confirmed the utility for the drug.

DR. OZOLS: I am not in favor of accelerated approval. I am in favor of a standard approval. I think that while we all agree that you need more additional

studies with this drug, I think if the accelerated approval mechanism, the type of trial that you would be almost forced to do in that case would be what you mentioned before, perhaps a randomized trial against DTIC, which I think would be not a good trial, and I would have a hard time participating in that. I think it would be a waste of patients and resource, but that is almost what you would be asking for.

Again, I think with standard approval, the type of things that we do want to find out are really dose and optimization, and ways in which it may be combined with other agents, and so forth, so I think it is just a different focus of the research that you would be asking for.

DR. RAGHAVAN: I think that Dr. Johnson's comments are very appropriate in the context of health care environment with reducing level of finance. On the other hand, I think we made a good decision in terms of not restricting the labeling for the reasons I said before, but I think that a critically important issue where I must say the sponsor has not convinced me at all is the issue of dosage. We need to look at issues related to dose of the product in identically administered schedules.

I listened carefully to the difference of doses in

infusion and with other drugs, with the assumption that other drugs wouldn't lower the response rate and survival. I am unconvinced by that. I think the reason we do two-tail survival curves is always accepting that patients can have worse results with innovative approaches, and so I think that it is in the patient's interest in terms of safety, in terms of identifying potentially the chance of a remission at less cost of quality of life terms, and certainly in terms of saving the community a substantial amount of financial dollars, but that is absolutely a key question that needs to be addressed by a good group in very carefully defined circumstances.

DR. SWAIN: Are there any other comments?

Did you want us to vote on standard versus accelerated approval?

DR. SIEGEL: Sure. That would be interesting.

DR. SWAIN: Okay.

[Laughter.]

DR. SIEGEL: That would be informative.

DR. SWAIN: All right.

First, I will ask all those in favor of accelerated approval of IL-2 for metastatic melanoma, raise their hand.

[Show of hands.]

DR. SWAIN: Three yes.

All those not in favor?

[Show of hands.]

DR. SWAIN: And abstentions?

[One abstention.]

DR. SWAIN: Then, I will ask the other question.

All those in favor of standard approval of IL-2 for metastatic melanoma, raise their hand.

[Show of hands.]

DR. SWAIN: Hopefully, this will work out the right way. You never know.

All those not in favor of standard approval?

[Show of hands.]

DR. SWAIN: Abstentions?

[One abstention.]

DR. SWAIN: Great. So it did work.

If there is no more discussion, then, we can conclude this morning's meeting. Thank you all very much.

[Whereupon, at 11:30 a.m., the proceedings were recessed, to be resumed at 12:30 p.m.]

AFTERNOON PROCEEDINGS

[12:40 p.m.]

DR. DUTCHER: We will start with Dr. DeLap.

DR. DeLAP: It has recently come to our attention that this is going to be the last Advisory Committee meeting for our wonderful executive secretary, Jannette O'Neill-Gonzalez. I just want to say thanks very much for a wonderful job, and we are deeply grateful and we will miss you. Thank you.

LT O'NEILL-GONZALEZ: Thank you.

[Applause.]

DR. DUTCHER: On behalf of the committee I want to thank you very much, and thank you for all your E-mails and faxes.

LT O'NEILL-GONZALEZ: You are welcome.

DR. DUTCHER: We are going to now move on with the afternoon session, which is the application for Neomark (broxuridine) for the use as a cell proliferation marker.

We will begin with the Applicant's presentation.

Dr. Govier.

NDA 20-806 Neomark (broxuridine for injection)

NeoPharm, Inc.

Applicant's Presentation

Introduction and Overview

DR. GOVIER: Good afternoon, Dr. Dutcher, members of the committee and staff, Dr. DeLap, members of the division, ladies and gentlemen.

[Slide.]

I am Bill Govier, President and CEO of NeoPharm, and we are pleased to be here today to discuss Neomark as a cell proliferation marker to determine the tumor labeling index in breast carcinoma. The generic name for Neomark is broxuridine.

[Slide.]

I will make the formal presentation today, but I am not here alone. With me to help answer questions are Dr. Tony Dritschilo, Dr. William Goodson, Dr. Seema Khan, Dr. Tim Kinsella, Dr. Ted Lawrence, Dr. Jaye Thompson, and Dr. Fred Waldman. Each of these individuals has experience with some aspect of this compound.

My agenda will be to first provide some general background and overview information relating to Neomark, then go into a more in-depth consideration of the clinical data, and then conclude.

[Slide.]

NeoPharm obtained the right to submit a New Drug Application for broxuridine under the terms of a creda with

the NCI. Under that creda we had access to the data generated by investigators working under the NCI IND. Actually, broxuridine has been administered to humans for various purposes since 1964, almost 34 years ago, but no one has had the opportunity to seek marketing approval until now.

[Slide.]

Neomark is a tool to rapidly obtain prognostic information about a breast carcinoma. It is a prognostic indicator. We believe that the information it provides is important for both the physician and the patient, but there are several things that Neomark is not.

[Slide.]

It is not a therapeutic agent in this indication. It does not treat the tumor. It is not a diagnostic agent. It doesn't tell you that the patient has cancer. It does not direct the physician to use any specific therapy.

I would like to comment on that last point a bit. People often try to ask too much of a prognostic indicator. A prognostic indicator does not indicate a specific therapy.

If you think about it, among the hundreds of laboratory tests which are done today, you can probably count on one hand the number that absolutely dictate which specific therapy to use.

Most tests provide information which the physician uses along with a lot of other information to decide what to do. For example, a blood pressure reading of 150/95, a cholesterol level of 290, a PSA level of 300 do not dictate which therapy to use. An injected contrast medium does not specify what therapy one must use. Yet, each of these tests is a prognostic indicator.

The labeling index was not intended to identify a specific therapy as being the most appropriate. It probably should not do that. The therapy of breast cancer is in a constant state of flux, and we hope is always improving with new therapies becoming available. If a test is directed toward one specific therapy, that test could become out of date very quickly.

As it is, the labeling index has stood the test of time as a prognostic indicator over the past 30 years. The value of the index is that it is an excellent indicator of how aggressive a particular tumor is. This information may never be identified using any other considerations.

This knowledge may lead the physician either toward or away from alternative therapies. Any edge which the physician can have to help decide which patients should be considered for perhaps non-standard protocols should be helpful.

As you will see, the labeling index, or LI for short, is an independent indicator of the prognosis of a particular patient. It provides information over and above that obtained from considering any other characteristic of the patient.

As I continue, I hope that you will recognize the utility of that information. We wish that there was no need for a prognostic indicator. Ideally, we would like to be able to identify the tumor and immediately know that we can cure it. A prognostic indicator would not be needed in that case. In the real world, however, that is not the case.

[Slide.]

Broxuridine is a thymidine analog and the structure is shown here. The methyl group of the thymine is replaced by bromine, here, and the cell basically cannot tell the difference.

[Slide.]

Broxuridine is incorporated into the DNA of actively dividing cell, that is, cells in S-phase, as a substitute for thymidine. It can be identified in those cells using standard immunohistochemical techniques.

This information yields a calculation of the tumor labeling index, which is the percentage of actively dividing cells in the particular tumor.

[Slide.]

The general concept of the utility of the LI is fairly simple and has been known for many years. Malignant tumors have actively dividing cells. The more dividing cells the tumor has, the more malignant or aggressive it is likely to be, and by definition, the higher the LI will be.

It is generally accepted that highly aggressive malignant tumors are more likely to kill the patient, thus, the general principle is that the higher the LI, the more aggressive the tumor, and the more likely it is to kill the patient. This principle holds regardless of any other characteristics of the patient.

[Slide.]

Determination of the LI is straightforward and reproducible. When a patient comes to the doctor, the lesion is first identified as a malignant tumor by fine needle aspiration biopsy or other suitable technique. It is generally not possible to determine the exact type or stage of the tumor at this time.

A small dose of Neomark is administered intravenously in a 30-minute infusion to the patient just before the surgeon removes the tumor. We use an infusion because the half-life of the compound is very short, about 10 minutes.

This labels the entire tumor, as well as any metastases which may be present. A small piece of the excised tumor, which is now labeled with Neomark, is sent for immunohistochemical analysis.

The labeled cells are easy to distinguish and are counted with a microscope. Approximately 2,000 total cells are counted in several microscopic fields and the percentage of Neomark-marked labeled cells is called the labeling index.

[Slide.]

This is a photomicrograph of a ductal invasive carcinoma to show how easy it is to recognize the labeled cells. I hope that projects properly for you. There is really no uncertainty in identifying these cells because they are stained brown.

I would like to emphasize that although we will present data today showing that the LI results have independent predictive value, and that they correlate with survival and recurrence, the idea that this correlation exists is not at all new. What we are presenting is a new method to obtain labeling index information by using Neomark.

[Slide.]

The classic way to determine the LI has been by

using tritiated thymidine. Such work with this compound began at least as early as 1967, and the database is quite large.

Over 10,000 breast cancer cases using this technique can be found in the literature, and there is a strong correlation between high LI and decreased survival. That is, patients having a high LI are less likely to survive for any extended period. This concept I think has been well accepted on the basis of a very large database generated by many investigators over a 25-year period, and I will show you some of that data as we go along.

[Slide.]

Tritiated thymidine has been the classic way to obtain this information, however, it has significant disadvantages which make this technique generally not clinically useful.

It is a radioactive material, and this carries with it a well-known set of handling problems. This also means that realistically, it can only be used on in vitro specimens. We believe that this is a significant disadvantage, and I will show you why.

Another problem with tritiated thymidine is that you have to use radioautography to get the results. Depending on the amount of radioactivity in the specimen,

development of a radioautograph typically takes weeks or months, and the results are not available when the patient needs them. Neomark provides a much simpler technique.

[Slide.]

It was shown to substitute for thymidine in DNA in 1957. To identify it in the cells at that time, however, required fairly laborious chemical techniques. The availability of a specific antibody, first made in the Livermore laboratory in 1982, made identification easier and encouraged both preclinical and clinical work. Since that time, more than 5,000 patients, having many different kinds of tumors, have had their LI determined using Neomark with great success and safety.

We are presenting data from about 200 prospective breast carcinoma patients in this NDA using Neomark with up to an 11-year follow-up period. The published literature documents that the Neomark results correlate very well with the tritiated thymidine results. We have found that Neomark provides prognostic information at least as good as that previously obtained in thousands of patients with tritiated thymidine, and the technique is easier and faster.

I will also present analyses which indicate that the Neomark LI provides information over and above that obtained by considering any of the other standard widely

used prognostic indicators, that is to say, the LI is an independent indicator and the results have prognostic value regardless of any other characteristic which the patient may have.

[Slide.]

The advantages of using Neomark over tritiated thymidine we believe are quite clear. It is not a radioactive material, it permits in vivo determination of the LI, and the results are available in 1 or 2 days.

[Slide.]

We believe that the ability to use this test in vivo is an important advantage. Unless extreme care is taken from the time the tissue is removed, an isolated piece of tissue in vitro does not behave the same as the intact organism. This kind of care is not typical of an ordinary production lab.

The in vivo technique labels the entire tumor rather than just the surface cell layers. It provides as homogeneous a distribution of the label as possible in the tumor because the label is carried in the circulation.

It can be used with very small tumors. This is important because if the tumor is very small, the pathologist may initially need the entire specimen. The presence of Neomark does not interfere with their ability to

read the sections and the immunohistochemistry can be done after they are finished.

Because the entire tumor is labeled, the LI information can be obtained from the worst looking, most aggressive region of the tumor. We think that this is very important since the overall behavior of the tumor probably reflects the behavior of its most malignant part.

Breast cancer is a heterogeneous entity and the tumor contains more than one cell line, each with its own degree of aggressiveness. We want to provide the most relevant information possible about the tumor, and to do that, we identify the most active portion. This is something which cannot be done with some other techniques which measure S-phase.

The in vivo technique also eliminates any problem with non-viable or poorly metabolizing cells in the in vitro preparations, and well as any problems due to lack of cell penetration in tissue slices.

The tissue must be metabolizing normally for the test to have meaning. This is almost never the case with in vitro preparations.

Finally, because the tissue has been labeled in vivo, one can go back at a later date, recut sections, and do additional analyses with it if it becomes desirable to do

so.

Having said all of this, I would like now to jump ahead and show you a simplified version of our final results to set the stage for a more complete discussion of the data.

[Slide.]

This is a figure showing standard Kaplan-Meier survival curves for our entire database. We have used a cutpoint of 8 for breast carcinoma, and I will show you why we selected this number later. All of the Kaplan-Meier curves will be presented in this manner.

The top curve shows the survival pattern out to about 11 years. Those patients who had a low LI, defined as 8 or less, most of them have survived. The bottom curve shows the pattern for those patients having a high LI but greater than 8. Many of them have not survived. The difference is striking.

As you might expect, these two curves were highly statistically different, the p is 0.0001, but perhaps the more important is they are clearly clinically different. Patients having a lower LI survived longer.

I would like to emphasize that this is true for all patients regardless of any other characteristics that they may have, that is, whether or not they are pre- or post-menopausal, ER or PR positive or negative, node

positive or negative, or the stage of the tumor. I will show you later the data to support this statement.

[Slide.]

The next figure shows similar Kaplan-Meier curves for recurrence-free survival. Again, the difference is quite striking between those patients with an LI of 8 or less, and those greater than 8, and the difference is highly significant. Again, p is 0.0001.

[Slide.]

We have also calculated the risk ratios for these patients depending on their labeling index. This table shows data obtained when we consider the LI by itself as a dichotomous variable with a cutpoint of 8.

Looking at the entire database, if the LI is greater than 8, the patients have a 16-fold greater risk of dying than if it is 8 or less. Similarly, they have a 4-fold greater risk of recurrence. We could equally well consider the LI as a continuous variable instead of dichotomous, and I will also discuss that later.

[Slide.]

Finally, because some people like to think in terms of survival rates, such as 5 year survival rate, I pulled them from the data, and will show them for various time points. This slide shows 3, 5, and 7 year survival

rates for patients with an LI of 8 or less and for those having an LI greater than 8. The differences I think are striking.

[Slide.]

We believe that the LI is a very valuable piece of information for both the patient and the physician. First of all, and most important, it is a safe test. In the more than 5,000 patients in the NCI database who received a small dose of broxuridine to measure the labeling index, there are only 3 reported mild adverse events that we are aware of.

These were 1 episode of mild hypotension, 1 of a mild headache, and 1 episode of vomiting. There is also now 1 report of a rash, which occurred almost a month after Neomark administration and after other treatments, but that almost certainly was not related to Neomark.

[Slide.]

The LI information is useful to both the physician and the patient. It describes how aggressive the tumor is. Our analysis shows that it is an independent prognostic variable and provides information over and above that from other commonly used indicators, such as node status, tumor size, ER and PR status, menopausal status or tumor stage.

[Slide.]

I will also show you that the LI can separate the

patient's traditional prognostic factors into good or poor prognosis groups.

A high LI identifies patients who typically do poorly with standard therapy and who may be candidates for alternative therapy.

A low LI identifies patients who would be expected to do well with standard therapy even though they may have poor traditional prognostic factors.

The LI describes the characteristic of the tumor, but it does not attempt to tell the physician what therapy to use.

I would like to provide two different examples of the clinical utility of the labeling index at this point.

In the first case, consider a patient who presents with a small tumor and no positive nodes are found at surgery. Generally, one might think that this tumor can be controlled quite easily. However, if the LI is high, our data indicate that there is a cause for great concern and that this patient is a high risk.

[Slide.]

Here are some specific patients from our database who fit this description. These patients all had small tumors, no or at most one positive node at surgery, and a high labeling index. These patients did not survive for

extended periods.

The point to be made here is that if the LI is high, there is a significant risk that the tumor has metastasized whether or not you find the positive node at surgery. The high LI has highlighted this possibility.

Another way to say this is that these patients are at great risk of being staged incorrectly, and this could lead to the use of an inappropriate therapy. If the LI had been considered, a different therapy might have been selected. We think that the LI may offer a potentially useful way to stage a tumor.

[Slide.]

As a second example, consider patients who present with a good-sized tumor and positive nodes are found at surgery. One would ordinarily expect these patients to have a poor prognosis. Using the LI, however, we are able to identify a group of these patients with low LIs who do better than expected.

The patients shows on this slide are still alive even though they had large numbers of positive nodes and big tumors.

Utility to the physician is obviously important, but we believe that the patient considerations are equally important. The physician-patient interactions are much

different today than they were 10 or 20 years ago. Patients these days have a definite desire to know everything they can about their disease.

The moment of a diagnosis of breast cancer is probably the most important point in that patient's life. They do not want to be talked to in terms of generalities, such as overall 5-year survival rates. They want information that is as specific as possible for their own individual situation.

[Slide.]

The Neomark LI provides information which is specific to their particular tumor. Obviously, it does not permit the physician to tell a patient exactly how long they will live, no test does that, but it does provide information about the potential risk. It should help the patient to participate in a more informed manner in the therapeutic decisions which will affect their lives.

I would now like to show you the details of the patient database which we have obtained from Dr. Bill Goodson of UCSF and Dr. Seema Khan of SUNY in Syracuse. I will discuss these two similar studies together and show you both separate and combined data.

The study required that the patients receive an infusion of Neomark over a 30-minute period just prior to

their surgery. The recommended dose is 200 mg/meter squared or approximately 350 to 400 mg total dose.

[Slide.]

A portion of the tumor is examined in the immunohistochemistry lab to determine the LI. These were prospective studies. The investigators offered participation to each of their patients who met the entry criteria.

No therapy decisions were made on the basis of the LI.

The two studies are comparable and roughly match the general breast cancer population. In our analyses, there was no evidence that any possible patient selection bias could alter the correlations which we found between the LI and patient survival or tumor recurrence.

Following the immediate post-treatment phase, patients were generally followed, first, at 3-month intervals, then at 6, and then at 12-month intervals as their schedules would permit. Follow-up was done by means of office visits, visits to other physicians as recorded in the tumor registries, or by telephone.

[Slide.]

The patient demographics for each study separately and the two combined are show in the following three slides

showing both patient numbers and percent. There are a lot of numbers on these slides, and I don't plan to dwell on these numbers. They are provided in your set of slides, but at this moment, suffice to say that our analyses indicate that it is acceptable to pool the two sets of data. I will also show you that each study independently separates the patients nicely at the cutpoint.

[Slide.]

The second slide shows more of the demographics and ending at the bottom with some therapy considerations. After their surgery, the patients received whatever additional therapy was considered appropriate for their tumor at the time, and the patients were followed.

[Slide.]

This slide shows a different way to look at the therapy and shows you that the patients received a standard therapeutic regimen for the characteristics recognized at the time. For example, a high proportion of the pre-menopausal, node-positive patients received chemotherapy.

[Slide.]

We have identified a labeling index of 8 as the cutpoint for this prognostic indicator in breast carcinoma. This value was initially selected because it is the median

value of the group. However, we have examined many other cutpoints and overall our judgment is that 8 is the best.

As an example, this slide shows a quartile analysis. The first and second quartiles are clustered together and are not statistically different, and the third and fourth quartiles are also clustered together and are not different.

[Slide.]

We have also examined several other cutpoints, most every other number, in fact, we have looked at, using a Cox analysis considering both the likelihood statistic and the risk ratios, we concluded that 8 was still the best cutpoint.

Purely on the likelihood results, one could argue that 9 is better because the one with the lowest likelihood score should be the best, but because the likelihood statistic was very similar at 8 or 9, very tiny differences between point 1 and point 9, we selected the one that produced the higher risk ratio, which was 8.

We also recognized that there is some degree of uncertainty right around the cutpoint. This uncertainty is not unique to this test, however. Rather, it is the rule even for things as common as blood pressure.

For example, an arbitrary cutpoint of 140/90 for

hypertension requires the physician to make a judgment as to whether to treat patients whose readings are relatively close to that value on either side.

Now I would like to show you the results of several analyses of the database.

[Slide.]

In study 1, the Kaplan-Meier survival curves, using 8 as the cutpoint, looked like this. Again, the top line shows patients with an LI of 8 or less, and the bottom line shows those with an LI greater than 8. There is a highly statistically significant difference in survival between the two groups. the log-rank p value is 0.0001. Clearly, those with a low LI survived longer.

[Slide.]

In study 2, here are the same curves, again showing a statistically significant split between the two groups using the cutpoint of 8. In this case, p is about 0.03.

[Slide.]

When you combine the patients from the two studies, we get the set of curves which I first showed you in my introductory part. The two groups of patients, as defined by the cutpoint of 8, are highly statistically different as to survival, and we believe that this

information is clinically significant, as well.

Note that these curves represent the entire patient data set. We will discuss various subsets in a moment. The test also provides information as to its probability of a recurrence.

[Slide.]

This figure shows you the results as it relates to recurrence-free survival in study 1. Patients with an LI of 8 or less have a much greater chance of surviving recurrence free. Note that there are many more events in the patients having the higher LI. The difference is highly significant again at 0.0002.

[Slide.]

A similar figure showing the results from study 2. Again, the results are significantly different. Here, the p is 0.02.

[Slide.]

And the combined results look like this, with a highly significant difference between the two groups of a p of 0.0001.

I would like to show you the results of the analyses of several subsets of patients using the combined data set.

[Slide.]

Here are the Kaplan-Meier for pre-menopausal patients. We are able to separate these patients into two groups based on the LI, with the group having the lower index having a significantly better survival probability. P is 0.006.

[Slide.]

Here are the curves for post-menopausal patients. Again, we can discriminate very clearly between two groups. Next, I will show you the curves for stages 1, 2, and 3.

[Slide.]

In stage 1, there is a separation, but the data are a little immature and the difference is borderline significant at 0.07.

[Slide.]

In stage 2, the difference is highly significant with a p of 0.004.

[Slide.]

In stage 3, the curves are obviously different, but because there were no deaths at all in the group with the low LI, we lose statistical power and the p value is borderline, but approaching significance at 0.13.

[Slide.]

Similarly, here are the curves for patients with negative nodes. The two groups show borderline significance

with a p of 0.1. If this data matured a bit longer, this p will almost certainly fall. Now, to demonstrate that this is likely, I would like to refer to the literature for a moment concerning this subset of node-negative patients.

Rosella Silvestrini, at the National Tumor Institute in Milan, just this year published on a group of 3,800 node-negative breast cancer patients who had a labeling index determined with tritiated thymidine, and they were followed over a 20-year period.

[Slide.]

Their evaluation using a Cox model showed that the LI was the only feature -- considering at this point the LI, ER status, and the tumor size -- the LI was the only feature which maintained a significant correlation with survival, distant metastases, and local recurrence over the entire time period, and their p values are indicated in blue. ER status, for example, was not independently predictive, and tumor size maintained predictability for recurrence and metastases, but not for overall survival.

The point here is that this very large group of patients, 3,800 of them, shows the prognostic value of the LI in the node-negative subset of patients, and Neomark is directly comparable to the tritiated thymidine results.

[Slide.]

If we consider the node-positive patients, there is a highly significant difference between the two groups. Those having a low LI showed much better survival than the group with an LI greater than 8. Node-positive patients are generally considered to have a poor prognosis, of course, but the LI can identify a group of patients who would be expected to do well, and the p was 0.0001.

[Slide.]

Here is the figure for the recurrence-free survival of patients with positive nodes. Again, we are able to separate two distinct groups. The overall data are quite compelling to us that the Neomark labeling index can provide prognostic information relating to any subgroup of patients, as well as the overall group.

Next, we asked does it add information that is not otherwise available. To answer this question, we have run Cox proportional hazard models with the data. I will show you the analyses using the data combined from both studies.

[Slide.]

This slide shows the potential prognostic factors which we ran in a univariate Cox model for survival. They were then tested for inclusion in the multivariate model.

[Slide.]

We can consider the LI as either a continuous

variable or as a dichotomous one with our cutpoint at 8. The final model, when we consider the LI as a continuous variable is shown here. This model suggests that of all those examined, the most important factors in predicting survival are labeling index and node status.

After adjusting for node status, the LI adds information for predicting survival. The risk ratio for LI imply that as it increases one unit, the risk of death increases 1.08-fold. Log likelihood tests have confirmed that the model fit is improved by adding the labeling index.

[Slide.]

Now, if you consider LI in a dichotomous fashion, the final model contains the same factors, LI and node status, as significant predictors of survival. In this case, the risk ratio becomes 12.4 for LI after taking into consideration the node status, suggesting that the patients have a 12.4-fold greater risk of dying if the LI is greater than 8.

[Slide.]

This bar graph shows the risk ratios for a number of prognostic factors. In this case, we have considered each one separately, using the labeling index split at 8. In this model, the risk ratio associated with the LI is 16. This suggests that if you know only the labeling index, the

patients with an LI greater than 8 have a 16-fold increase in their risk of death compared to patients with labeling index of 8 or less.

In comparison to the other factors, the LI is obviously important.

[Slide.]

We have looked at the Cox proportional hazard analysis for the endpoint of recurrence-free survival, as well, and the final Cox model looks a little different than the overall survival model. This model suggests that the most important factors for predicting recurrence-free survival are the labeling index, menopausal status, and cancer stage.

After adjusting for the other prognostic factors, the LI adds information for predicting recurrence-free survival. The LI risk ratio implies that as it increases one unit, the risk of recurrence or death increases 1.04-fold looking at the LI as a continuous variable.

[Slide.]

If you consider LI in a dichotomous fashion, with a cutpoint of 8, the final model contains the same factors, and in this case, the LI risk ratio becomes 2.2 after adjusting for the other factors.

[Slide.]

This bar graph shows the risk ratios for relapse-free survival of several prognostic factors, again each considered separately. This suggests that patients with a labeling index greater than 8 have a 4-fold increase in their risk of recurrence or death. Again, LI is important.

[Slide.]

The utility of the labeling index has been demonstrated previously using tritiated thymidine as the label. Over 10,000 cases have validated the prognostic ability of this test. The many problems associated with the radioactive label have meant that this test for all practical purposes did not have clinical utility.

Neomark, which is also a thymidine analog, provides results which are at least as good as those with tritiated thymidine, but with much greater ease of use, and the results are available in a timely fashion.

Neomark has been found to be very safe to use. In over 5,000 labeling index cases, there have been only three reports of mild adverse events. Neomark clearly can determine the labeling index. We believe that the labeling index information is helpful to both patients and physicians.

For the patient, it provides information which is

specific to their individual tumor. It may help them understand their degree of risk and help them participate in therapeutic decisions.

For the physician, the labeling index predicts the likelihood of survival and the recurrence-free survival in all groups of patients. We are clearly able to separate groups using a cutpoint of 8.

The Cox analyses indicate that the labeling index provides information over and above that obtained from considering any other generally used prognostic indicator.

Beyond its general utility, we believe that the labeling index is particularly valuable in certain instances. Patients with small tumors, negative nodes, and a high LI appear to be at much higher risk than would be thought without knowing the LI information. This information may lead the physician to use different therapy.

Conversely, patients with positive nodes and a low LI appear to be at less risk than would otherwise be thought. This, too, may lead the physician to consider different therapies.

The LI may also help the physician decide what to do in cases where the situation is borderline after considering all other relevant factors. It also appears that the labeling index may be a very useful way to stage a

tumor.

When all is said and done, if we had a clearly defined therapy which would cure all cases of breast cancer, we would not need a prognostic indicator. All patients having a high labeling index would then survive as long as anyone else.

We would hope that this may be possible in the future. With the current therapy, it is not the case and we do need prognostic indicators. We do not claim that Neomark is the best prognostic indicator that will ever be available. There is always room for improvement, and more ways to use the existing information will be identified over the coming years.

We do believe that the data show that the Neomark labeling index is a safe and effective prognostic indicator at the present time. As with any test, it should be considered in conjunction with all the other available information when considering how to treat a patient.

There is currently no prognostic indicator approved for use in breast cancer. We believe that this test should be made available and we ask for your recommendation to approve.

DR. DUTCHER: Thank you.

We will open it up to the committee for questions

to the sponsor. Dr. Simon.

Committee Questions to Applicant

DR. SIMON: I want to make maybe some questions and some observations that may give you an opportunity to sort of clarify things.

One thing I am concerned about is your use of 8 as a cutpoint and then computing statistical significance values based on having selected the cutpoint that you felt was the best discriminant in the data. You know, there is a growing statistical literature that says that the resulting p values are invalid when you do that.

There is a variety of ways of trying to adjust your p values for that, so all of your p values you give and all of these log-rank tests from Kaplan-Meier curves or even the Cox regression analysis that are based on a binary representation of the labeling index seem to me to be sort of invalid.

That would be one thing. I guess the other thing is you bring up Silvestrini's data, and she used a labeling index of 3 percent as a cutpoint for tritiated thymidine defined labeling index, and her results indicated that -- I think it is important that you do bring up her results, because I think really to try to claim anything about what is important prognostically or what adds to anything else

based on the size of data set that you present is difficult. Usually, the statistical reliability of an analysis depends on the number of events you have. You have only something like 23 deaths and I think 38 failures, and I think to try to tease out of that what is important for what subsets of patients, when you have a very mixed group of patients, almost all of whom have received systemic therapy, to try to provide advice for clinicians in terms of therapeutic relevant advice for really a minuscule small data set like that, I think is problematic.

In Silvestrini's data, in which there were over 2,000 patients that were followed, who had had tritiated thymidine and labeling indices done, there basically were no subsets of patients that she could identify, and those 2,000 patients who were followed, none of whom had systemic therapy after surgery, and she just divided them into patients who had labeling index less than 3 percent versus over 3 percent based on -- and that was not sort of a data-derived sort of thing, and there were even for the low labeling index patients, she had something like a 35 percent recurrence rate at 10 years without systemic therapy.

So, there were no subsets of patients who could be identified based on the tritiated thymidine labeling index that would provide any sort of basis for withholding

systemic therapy, for example, for those node-negative patients.

So, I am concerned basically that you are drawing strong conclusions from really a minuscule small data set and that your conclusions are biased by the use of a cutpoint, a data-derived cutpoint.

I wanted to give them an opportunity if they had any sort of information that would indicate they had done any other analyses that mitigated any of those concerns.

DR. THOMPSON: I am Jaye Thompson. Concerning the cutpoint, I remember a pre-NDA meeting when we arrived at the FDA and we discussed how best to describe this phenomenon that we are seeing of predictability, and they suggested that we find the optimal cutpoint, which really does entail data dredging.

We investigated thoroughly numerous cutpoints and I can assure you that almost every single possible cutpoint does show this phenomenon. This is not new. Tritiated thymidine has been showing this, as well.

I believe there are other that speak to this, but since tritiated thymidine is not done in vivo, it does not uptake as much, and that is why the medians are so different, the 3 percent versus closer to the 8 percent. I believe there are other people who could comment on the

other two questions that you have.

DR. GOVIER: I wouldn't draw any importance between the cutpoint of 8 that we used and the one which, for example, Silvestrini and other investigators have published with the tritiated thymidine. Their numbers are always lower, but I think that is a function of the test, and each test has to be looked at separately in each specific tumor, and we found that 8 was the one that was the median value initially, and as our analyses showed, came out statistically to be the very best.

Dr. Goodson.

DR. GOODSON: My name is Bill Goodson. I am formerly from UC in San Francisco. I am not really a consultant to NeoPharm. I am here basically in compliance with the terms of the creda and the fact that the NCI, in the process of the creda, asked me to cooperate with NeoPharm, and I have that in a letter if you would like to see it at some point. NeoPharm has paid the supplemental costs of part of the analysis, but that it is the only thing, and all of these patients were accessioned before NeoPharm had anything to do with the drug, at least as far as I knew, so that I have really no conflict of interest in this.

There are two comments I would like to make.

First of all, our original analysis was designed to divide this simply as above and below the median, so that from the standpoint of what we did in our own work, the median in our way of looking at it is just -- I think it is actually 7.9 rather than 8, and that was a predesigned and not derived from the data.

The second comment I wanted to make is in comparing the value that Silvestrini gets and the value that other persons have obtained with labeling indices, I recently was at the San Antonio breast cancer meetings and had a conversation with John Meyer, who as you probably know is the other person who has done a very large number of patients with tritiated thymidine in vitro labeling, and John's comment to me in talking about this was that he said specifically that Silvestrini uses a 30-minute incubation whereas he uses approximately a two-hour incubation, and that he at least in his own comments to me said basically that he thinks that that is probably the reason that she gets a much lower number.

I think before focusing on a lower mean or median, that would be something that would need to be tracked down, I mean if that is important to you, but I think clearly that you are looking at a number that we, in our minds, selected initially, and if you look back at our original publication

back in 1991, long before any of these other analyses were done, we were talking about the mean and median at that point, and that is what we had chosen.

Thank you.

DR. SIMON: Do you have any information on inter-laboratory reproducibility of your assay?

DR. GOVIER: I can answer that in -- well, I can answer that in two ways. We do have information based on the laboratory which was primarily responsible for the data that is shown here, and that is the one at UCSF, and that is in the --

DR. SIMON: No, that is just two technicians reading the same prepared slide. I am talking about inter-laboratory reproducibility, if you cut the slide into two sections, and one is read and processed at one lab, and the other is --

DR. GOVIER: I understand. We do not have that information as yet. Our proposal is that if this were to be approved, we would have a Central Laboratory do the work for this test, and the plan then is to train that Central Laboratory, so that they will be up to the standards of the one at UCSF.

DR. SIMON: What about intratumor variability, two sections on the same tumor? You make the point that the

value of in vivo labeling is because of tumor heterogeneity and therefore you get to look at different parts of the tumor, but in your application, you never actually do that, so how much --

DR. GOVIER: Let me ask Dr. Waldman to respond to that.

DR. WALDMAN: I am Fred Waldman, also with Bill Goodson at UCSF. I want to say that I am leading the program project on prognostic markers in breast cancer that Liam Smith led for a very long time, and one of the original goals was to bring things from the lab to the bedside, and that is sort of we are real proud to come out here and try to support this, because that is sort of one of the things that translational research-wise we want to do.

To answer your question, Rich, the goal of the scoring is to come up with the most aggressive labeling index of that tumor, which is generally not in the central necrotic area of the tumor, but will be along the growing edge usually. That may vary, you know, on some tumors, since they are all labeled, we can check in different blocks, we can look at one block, comparing it to a lymph node, and so forth.

When we do that, they all correlate very well, but generally speaking, one block is sent to us for scoring as

representative by the pathologist, and so for most of these cases, that is the block that we looked at. In a handful of cases, we have looked at every block and done scoring to look at the correlation, and the tumor heterogeneity exists in different regions, but if you use this method of scoring rather than a purely random picking out fields, but rather looking for the higher labeling regions and going to multiple, high-powered fields to get a labeling index of 2,000 cells, there is a pretty good correlation among those different scores.

DR. DUTCHER: Dr. Swain.

DR. SWAIN: I had a couple of questions about the implementation of the study and the primary objective. Apparently, I guess 21 percent of patients were excluded from the study done in San Francisco. Can you just comment on that and why they were actually given the drug?

DR. GOVIER: Gail was trying to put her finger on it, but there is a table. Briefly, these people received the drug, but there were reasons why they, actually, were then, in fact, protocol violators, but many of them were because they were stage zero tumors, and one couldn't tell really directly ahead of time. Others were excluded because there was no labeling index value obtained, most often that was because of essentially a logistics snafu, if you will,

in the system, and the tissue samples actually never got to the laboratory.

DR. SWAIN: Couldn't it have been obtained afterwards?

DR. GOODSON: Could I comment on that? I do take -- this one, I sort of feel like, you know, I take this as something that I am responsible for. A fair number of those patients, the original design of this was as part of a program project, and we were interested, not just in primary tumors which is what this application is based on, but we also had a series of patients with recurrent tumors, we had a series of patients with in situ only tumors, and we had a series of patients, actually, one patient with a sarcoma and a couple of patients for whom there was no residual tumor in the breast, but for whom we could be a labeling index on an axillary lymph node.

There was a fairly small number of patients in whom there was a biopsy done, and the biopsy had grossly positive margins, and we then gave the BVR and went back and did the analysis afterwards, and despite what the pathologist had told us it looked like an incisional biopsy initially, there was no identifiable tumor that we could count in the residual specimen at the time of doing further surgery.

So, I mean most of these -- I mean we can go through them one by one if you want -- but I mean that is basically what happened.

DR. SWAIN: So, you intentionally had the recurrent patients in there, that was your study design initially.

DR. GOODSON: Yes. Our intent was to -- this is part of a study tied to markers, and all of this stuff would sort of disappear in the laboratory and be looked at in multiple different ways, and we were interested in recurrent, as well as primary tumors in what was initially set up, and that is really -- I mean I think if I had a recurrent tumor, I would still probably do it on the basis of what was supposed to go into the program projects as opposed to what is going into this creda and drug application, et cetera.

DR. SWAIN: I have another question for you on the follow-up. On the slide, it was shown that the patients were followed up every 3, 6, and 12 months, and I think in the application, it said every 6 months, and then a lot of patients actually didn't get follow-up for recurrence regularly, plus I think survival hadn't been calculated for about 19 percent of the patients for the last year and a half. Can you just comment on follow-up?

DR. GOODSON: I think the closing date of this actually was sometime in late '96 was the date that the data was followed up. Patients were followed up every 3 months for the first year, every 6 months for the first five years, and then once every 12 months after that.

I would say roughly about 20 percent of the patients have left the Bay area, at least of my own patients, and have been in contact with these patients by telephone on an annual basis, and at least as of December 1996, all of the patients that I was involved in, 100 percent of them had either had a follow-up or a death certificate or some other confirmation within the year 1996, and I haven't done that for '97 yet.

DR. SWAIN: And they are all your patients, there is no other investigator?

DR. GOODSON: There are other investigators, but those other investigators, I can say the same thing for them, we have tracked them down.

DR. SWAIN: I noticed that a lot of the patients received as treatment radiation therapy, I think about 58 percent, whereas, some 90-some percent had mastectomy. Were they in a poor prognostic group?

DR. GOODSON: I am not quite sure. I have this feeling that there may be an error in the 90 percent

mastectomy. That number I have not seen, and my feeling is that the radiation -- just I know in my own general practice -- is probably more like about a 40-60 split.

Their radiation therapy was given postoperatively as an adjunct only in patients with more than 4 positive nodes or when grossly positive margins is in keeping with what was considered standard of practice at our institution or else as an adjunct after a partial mastectomy.

DR. SWAIN: Okay. Then, I had a question about the follow-up. Was there a different -- median follow-up I think for the whole study was about 5 years or a little over 5 years -- was the median follow-up longer or shorter or different for the patients who were less than 8 versus greater than 8, because this was an interesting accrual in that patients were accrued over 9 years, so it was a very long accrual period.

In other words, did the patients who did well have shorter follow-up?

DR. GOODSON: I don't know. The median follow-up for the whole group is somewhere at about four and a half years, but I do not -- you know, you may know whether or not there is a different median for the high labeling index, low labeling index patients. I really don't know that.

DR. THOMPSON: I can't say -- I can tell you the

numbers, we don't have a back-up slide on that -- but I would say that we saw no pattern at all. We detected no pattern in the linked follow-up.

DR. SWAIN: Another question. Did you do time trend analyses because of the length of time that this study accrued to see if changing in treatment patterns or whatever would change the outcome?

DR. THOMPSON: No, we did not investigate that. The investigators were blinded concerning the labeling index as far as just citing therapy and treatment. So, we believe that they were receiving what was optimal or standard therapy at the institution at the time, and that, of course, probably did change in 10 years.

DR. SWAIN: Just one last question. Can someone comment on the mutagenicity of this compound at all in this dose?

DR. GOVIER: Yes, we do have a back-up slide with some of those points on it, which we will come to in a second.

[Slide.]

I guess the top line point there is that the LI dose in humans is about 5 mg/kilo. The studies which have looked at mutagenicity with broxuridine have found that it does produce teratology in mice, and the comments are that

it is a fairly specific teratology in mice, but the threshold dose for doing that is 60 mg/kilo.

There is some teratology in hamsters, again at a much higher dose, and the studies have also commented that even though they have seen teratology, they have not seen carcinogenicity. So, all of these effects are noted at much higher doses than we are using. There is a comment, as well, in the literature, which says that the effects that they see look a bit like high-dose vitamin A effects, and they noted that there were no point mutations produced, and for long term effects, one would expect to have point mutations.

DR. DUTCHER: Dr. Santana.

DR. SANTANA: How did you determine that the test was a satisfactory test for an individual patient, knowing that immunohistochemistry is notoriously, sometimes difficult to perform? Did you, within the same individual patient, look at another normal tissue like breast tissue or, for that matter, a bone marrow that notoriously has a lot of cell cycling? How did you determine for each individual patient that the test was satisfactory in terms of a control?

DR. WALDMAN: In terms of the assay itself, we of course run negative and positive controls. There is no such thing as a -- a negative control without the primary

antibody of course -- a positive control we run colon, normal colon tissue, which was done in vitro at BRDU, because there is a very specific pattern of labeling at the base of the crypts where the proliferation is going on.

We also run on every assay day a positive control of a breast tumor, which has a known labeling index, to look for interassay variability. Within a subject, if there appears to be zero labeling in a tumor, or very low labeling, there is really no great control for that.

We can look at normal ducts and normal lobules if they exist in the same section, and we see a very low labeling index, approximately 1 percent. Interestingly, it varies with the degree of dysplasia, but it is still very, very low, but we can see that there are labeled cells within the normal regions of the breast.

DR. DUTCHER: Ms. Beaman.

MS. BEAMAN: I wanted to know, do you have any Neomark data showing data labeling index by ethnic group, and secondly, did I understand you to say that in situ doesn't index?

DR. GOVIER: No, it is not that the in situ cases cannot be done, but they were not part of this protocol, and so we have excluded them from the analysis. We do have cases of carcinoma in situ, however, with labeling index

performed. We have an information, and it doesn't deviate from anything that we have shown you.

The question of ethnicity, I don't believe we have enough examples of different ethnic groups to really make a statement on that. Most of them were, in fact, caucasians, but I don't think we can draw conclusions from the small number of other groups that we had. I wish we could, sorry.

DR. RAGHAVAN: I have two questions. Looking at the data that you showed about safety, you cited 5,000 cases with 3 adverse events, and my understanding is this involves injections of a relatively benign substance.

In my experience, I can't think of a trial that was placebo controlled that had such a low rate of complications. That leads me to ask the question, how did you determine the complication rate? This does better than any other placebo I have ever seen.

DR. GOVIER: I don't know if it is better than a placebo or not, but clearly, these are the results which were reported to the NCI by the investigators who did the actual studies, and they were compiled on an annual basis by the NCI and put out in annual report form, and we obtained the information that way, so these other cases are not breast cancer patients, but they run the gamut of almost any tumor that you can imagine, and all I can say is that those

are the results that were reported to the holder of the IND.

DR. RAGHAVAN: Second question, I guess to Fred. The techniques for looking at ploidy haven't really been mentioned at all today in the results. It is another way of looking at a similar area of the cell cycle, and I am puzzled that you haven't even mentioned the "p" word, so there are clearly some concerns about the way of analyzing ploidy.

Put it in a context please.

DR. WALDMAN: The way it was explained to me is that this is not a requirement of this proposal to compare it to other methods per se, but, of course, academically, we are very interested in comparing it, and flow cytometry analysis for S-phase is a standard that has been used for the last 10, 15 years, and it is what I started doing with Joe Gray out at Livermore prior to the antibody being available.

You mentioned ploidy, of course, which is just whether the DNA content is diploid like normal cells or abnormal aneuploid, and that has not established itself as being independent predictive ability, whereas, S-phase by flow has in a number of large studies.

We started doing that and switched over to the BRDU because our success rate was much greater for doing

BRDU. We can do an assay on 95 percent of the cases with BRDU, whereas, by flow cytometry, the success is dependent on having a tumor that is big enough that we can make thick sections to get nuclei out of that, and even then, if there is broad interference with the flow, we are only able to be successful 70, 80 percent of the time.

In any case, we picked out of these cases that Bill had, 135, we picked out the 95 cases where we were able to do S-phase and BRDU.

[Slide.]

This just recently, I apologize very much for the form of these two curves on the ordinate is still recurrence-free survival, and the abscissa showing time in years, and for these 95 patients -- and again, as Dr. Simon said, it is a small number -- there was a significantly -- well, let me just -- these are the two curves for BRDU and S-phase by flow.

When Dan Moore, our statistician, did a randomization test, Monte Carlo, to pick out whether there was a significant difference between these two curves, he tells me that there is.

So, on this early data, unpublished, the BRDU appears to be more frequently successful as an assay and perhaps better predictive in this overall set of patients.

DR. RAGHAVAN: Could you explain, with the second graph there, how did you approach defining a cutpoint?

DR. WALDMAN: Also, the median. The reason that we used the median is that we don't -- you know, we don't want to optimize the cutpoint and then feed back and so with S-phase, the median among different groups, in fact, in the literature, is different, so they are mostly basing it on their on median rather than any internationally valid median for flow S-phase.

DR. DUTCHER: Dr. Simon.

DR. SIMON: One thing I noticed -- maybe I noticed it wrong -- but what you presented in terms of the final Cox model for relapse-free survival didn't seem to agree with Table 8 in the NDA. Table 8 actually showed that labeling index in that final Cox model for relapse-free survival was not statistically significant, and in the slide you showed, unless I am mistaken, you indicated that it was.

DR. THOMPSON: There may be some number changes. The original NDA has been amended twice, and each time we have gone back and gotten the latest survival information and recurrence information, and the numbers we presented today are from the most recent amendment.

DR. DUTCHER: Dr. Johnson.

DR. DAVID JOHNSON: In addition to the S-phase,

were any other proliferation markers that were performed by immunohistochemistry applied to this group of patients?

DR. WALDMAN: We are in the process of doing the Q67, but let me say that the correlations are interesting. When we just look at correlation between BRDU and S-phase, the r-squared is about I think 0.16. Between BRDU and Q67, in a different set, not inclusive of all of these patients, and that is why we are still adding on the later patients, the correlation was on the same order, but it is not a perfect correlation. Q67 is measuring growth fraction in lots of cells that are not necessarily in S-phase, so we are really measuring different things with the different markers.

DR. DAVID JOHNSON: Before you step down, I didn't see in any of the prognostic factors that were investigated that tumor grade was included. We have histopathology. Maybe that was intended to also include that, but I am just asking, was that looked at independently?

DR. THOMPSON: We did have some of that data available, but it was not available on very many of the patients, and so it would end up excluding too many of the patients from the analysis. Let's say maybe 60 or 70 patients we are missing that information.

DR. DAVID JOHNSON: Really.

DR. DUTCHER: Did you not have Central Pathology review for this?

DR. DAVID JOHNSON: The question was, was there Central Pathology review, and does the pathologist not assign a grade to the tumor?

DR. GOODSON: Our cases are all independently reviewed by a single pathologist who goes back and double-checks things, and this is actually -- I can't comment -- I know that is what we do, and I am not sure where this discrepancy comes from -- I can't comment on that, but I can answer that we do do the review.

DR. DAVID JOHNSON: It may not be worth asking, but what about -- you indicated here that another prognostic factor investigated was tumor stage. I interpreted that to mean the T size, the size of the tumor. Was that done based on actual size, though, as well? In other words, the actual size of the tumor, not the T size, not T1 versus T2, but 1 cm versus 3 cm versus, and so on.

DR. GOODSON: I think you actually did this as a continuous variable on tumor size, but staging, it was the UICC classification, which is T1, you know, and M-zero, M-zero, and then T1 or T2 with M-1, and so you were just staging -- it is kind of harder to do it as a continuous variable. You are either at stage 1 or stage 2 or stage 3,

but the tumor size was done as a continuous variable, as were the number of nodes involved.

DR. DAVID JOHNSON: I am particularly interested in those that were node-negative and whether or not the LI was able to discriminate in the patients based on LI versus tumor size.

DR. GOODSON: I think I understand your question. I can't give you that off the top of my head, I really don't know.

DR. DUTCHER: Other questions? Okay.

Thank. I guess we can take a quick break. We will back at 2 o'clock.

[Recess.]

DR. DUTCHER: FDA presentation.

FDA Presentation

[Slide.]

DR. KAREN JOHNSON: Dr. Dutcher, members of the Advisory Committee, FDA colleagues and guests, I will be presenting the medical or clinical summary for NDA 20-806, an application that pertains to the use of bromodeoxyuridine for the determination of labeling index.

[Slide.]

This summary was made possible by the effort of an extended team of reviewers who are acknowledged on this

slide.

[Slide.]

It has been proposed that bromodeoxyuridine, when administered intravenously, is a cell proliferation marker that can be used to estimate the labeling index of malignant breast tumors. The proposed dose is 200 mg/meter squared administered over 30 minutes in the hour before surgery.

[Slide.]

In looking at the options for determining clinical benefit, one of those would be an examination of the correlation, its strength and quality between survival and the bromodeoxyuridine labeling index.

A second option for looking at clinical benefit would involve the clinical relevance of separating patients into prognostic groups based on the bromodeoxyuridine labeling index.

[Slide.]

As far as some background here is concerned, the intravenous use in investigational studies began in 1979 under IND 21-97. Prior to that time, bromodeoxyuridine had been given intra-arterially and this was not feasible for labeling index uses.

This application is supported by two clinical trials. The first clinical trial was begun in August of

1986 at the University of California, San Francisco. The accrual period for that trial ended in March of 1995, and the follow-up continues.

The second study was begun in May of 1991, and it was conducted at the State University of New York at Syracuse, and the accrual period for that study ended in April of 1995.

The initial results from these studies were available for an NDA submission in December of 1996. At that time, the follow-up data cut-off point was the end of October 1996. Since then, updated information has been provided, and in August of 1997, we received information that extended the cut-off data up to the end of July 1997, so that was an additional eight months of follow-up data.

For the cut-off period at the end of October 1996, there were 54 patients who had not had follow-up information included in the data set for the year prior to the cut-off. With the amendment in August of 1997, that number was reduced from 54 to 30.

[Slide.]

So, in describing the study submitted for review, I am going to proceed first with the larger of the two studies. You will see that both studies involved a single arm and survival was an endpoint that was available in terms

of the data set information.

The San Francisco trial was much larger than the one at Syracuse, involving 163 patients, and of course this was the largest difference between the two studies, the amount of data. Other differences included the size of the dose, the larger dose at San Francisco being 200 mg/meter squared versus 100 mg/meter squared at Syracuse.

Also, there were slight differences in the way the drug was administered. The drug was given over 30 minutes an hour before surgery at San Francisco, and at Syracuse it was given over the 30 minute prior to surgery following a prior dose of bromodeoxyuridine.

[Slide.]

So, in looking at the larger study, the objectives indicated that female breast cancer patients would be followed and that a labeling index would be obtained for each of these with bromodeoxyuridine.

The objectives clearly state that the proportion of patients recurring and the time to recurrence would be obtained.

[Slide.]

However, the protocol did not specify the primary endpoints for analysis, and there was not a methodical basis for assessing the recurrent disease status in these patients

since they were followed in a variety of different practices, and those practices could be different.

Also, clinical documentation or recurrence was not provided with the application. So, we conclude that a consistently determined recurrence endpoint cannot be verified, and that leaves survival as the primary endpoint for our review.

[Slide.]

In looking at the San Francisco study, the patients involved were females with a good performance status and normal organ function, and there was a requirement for cytologically or histologically confirmed diagnosis of resectable stage 1, 2, or 3 breast cancer.

By including all three stages in the patient population, this generated a very heterogeneous patient population. A number of labeling index studies in the literature actually focused on a smaller segment of the patient population, for instance, in node-negative patients, so we were looking at quite a broad category of patients, a very heterogeneous group of patients.

[Slide.]

There has been some debate about the size of the patient population for the San Francisco study, so I wanted to review this information with you. There were 207

patients who were identified in the institutional database as potentially being people who participated in T86-0217. All of these patients were targeted for having a determination of labeling index.

However, 5 of the 207 did not receive intravenous bromodeoxyuridine, but portions of their tumor were sent to the lab for an in vitro determination of labeling index. So, these patients technically were not in the study population. Another 3 patients received intravenous bromodeoxyuridine, but this was received as part of another protocol, and finally, a single patient was assigned two study accession numbers, and so there were two entries for that single patient.

So, the conclusion here is that there were actually 198 patients who had an intravenous infusion of bromodeoxyuridine for labeling index determination as part of the study done in San Francisco.

[Slide.]

Not all of those 198 patients, however, could be used for analysis. The sponsor has excluded 35 patients out of the 198. Thirteen of those patients had no labeling index determined. For 8 of those 13 patients, the samples were not sent to the lab for labeling index determination. However, since the application mentions that a strength of

this approach is that the pathologic sample that is formalin-fixed could be used, it is surprising that at least some of these 8 samples were not available for labeling index determination.

There were 11 patients who were in the data set of 198 who had recurrent disease. There were 7 patients with carcinoma in situ. There were 3 samples that came from lymph nodes, and one of the tumors was a sarcoma. So, these constitute the 35 patients who were excluded from analysis by the sponsor.

Of those 35 patients, there were 22 who did not meet eligibility criteria.

[Slide.]

Now, in looking at the sponsor's data set, the FDA concluded that there were 3 additional patients who might have been excluded from the analysis, 2 because there was no invasive cancer residual in the surgical specimen that was removed at the time of bromodeoxyuridine administration, and another patient was classified as being a stage 4 patient.

Also, we noted that there were three protocol violations in that 3 male breast cancer patients had been included, however, we did not remove them from the FDA data set for analysis because we had no basis to say that these 3 cases were performed differently from the others.

So, for the additional comments here, you see that the FDA has identified 3 additional patients for exclusion from analysis, and there were, in fact, 4 patients who did not meet protocol eligibility criteria.

[Slide.]

So, what we see from this is that 38 of 198, or 19 percent, of patients were excluded from the analysis either by the sponsor or the FDA, and there were 26 of these 38 patients who did not meet eligibility criteria.

These deficiencies in the data may have affected the results.

[Slide.]

Another question about the results comes up in terms of patient follow-up. There was no follow-up data for at least one year prior to data cut-off -- that is July 31, 1997 -- for 30 of the 163 patients who remained in the sponsor's data set.

We examined these 30 patients to see if there were any basis for determining that they were unlike the patients for whom follow-up was available, and we were not able to determine that there was a difference. We looked at various characteristics and outcomes. An example of this would be the labeling index, and the labeling index was similar for the group of patients who did not have follow-up compared to

the ones who did.

[Slide.]

So, now to describe the patients in the San Francisco study. The median age was about 51 years. The large portion of the patients, about 50 percent or half of them were stage 2 patients. The great majority received adjuvant systemic therapy, and this was another source of heterogeneity in the data set.

About 40 percent of the patients received cytotoxic chemotherapy. Another third received hormonal therapy only. About 20 percent received a combination of chemo-hormonal therapy.

The median duration of follow-up was nearly five years, but for some of the patients, there was substantially less follow-up than that, and the median value of bromodeoxyuridine labeling index, as you have heard, was 7.9. For purposes of the analysis, this was rounded to 8.

[Slide.]

Next, I am going to show you this histogram of the patient population, and you can see that the distribution of labeling indices for the various patients was asymmetric around that median value of 8.

So, the columns on this histogram are organized according to two-unit differences, and the patients in this

first column have a labeling index that is less than 2, in the second column, the labeling index is 2, but less than 4, and so forth, and you can that half of the patients are falling into the group with a labeling index of less than 8, and approximately half are falling into the group with a labeling index greater than 8.

[Slide.]

And then you have seen the Kaplan-Meier curves corresponding to those groups, and so this upper line here shows you the survival for the patients with labeling index less than 8, and there were only 2 events in this group, and then the survival curve here for the patients with labeling index greater than 8, and there were 20 events here, and I have excluded the one event. That was in the stage 4 patient.

[Slide.]

So, you have heard that using this breakpoint that was determined by using the median generated a relative risk of approximately 14-fold, so that patients who had the indices above 8 were 14-fold more likely to have a death event than those who had labeling index less or equal to 8.

[Slide.]

We have raised the question of whether there might be another way to determine a cut-off point for making

prognostic assessments, and the method that I have chosen to show to you today is based on receiver operating characteristic analysis.

I, of course, am indebted to my biostatistical friends for providing this information, but this method depends strongly on sensitivity and specificity, so I want to talk a little bit about those two concepts in the context of this study.

For the sensitivity determination, the event that we are predicting is death, and a positive test is defined as a labeling index greater than the cut-off, and I am going to be concrete and use a cut-off value of 8 just for the purposes of this discussion.

So, the sensitivity ends up being the deaths that occur in patients who have labeling index above 8 divided by the total number of patients dead.

The specificity, on the other hand, is based on patients who have the cut-off less than positive test, the people who have a negative test and an outcome that is not predicted by the test. So, here, for specificity, we are looking at the number of patients who are not dead, with a labeling index less than the cut-off of 8, divided by the total number of patients who are not dead.

These two parameters are important because they

allow us to compare the chances we have of estimating the true positives based on the test versus the false positives based on the test, and so the indicator we have of false positives is 1 minus the specificity.

So, we could use as an informative breakpoint one in which the odds of correctly predicting an event exceed the odds of an incorrect prediction. To illustrate that now I am going to go to the receiver operating curve.

[Slide.]

Here, you see that we have a sensitivity over here, which is a representation of the true positives. We have 1 minus the specificity, which is a representation of the false positives, and we have a criterion line or break-even point where our ability to designate the true positives is greater than the risk of identifying false positives.

So, we have a curve based on a series of cutpoints that shows us where the test is operating in an advantageous method.

Now, where this value of cutpoint crosses the line here corresponds to your labeling index of 11.7, and where the line leaves the curve up here corresponds to a cutpoint of 3.6. So, you can see there is a whole range of cutpoints which would meet the criterion for allowing us the chance of

making more correct predictions than incorrect predictions.

However, the point of this exercise was to find the cutpoint that was the optimal one, the one that maximized the ability to do that, and that corresponds to a value of 9.1.

[Slide.]

So, moving on, what I want to do now is to compare the results that we get when the breakpoint used is 9.1 rather than the 8.0 that the sponsor has told us about, and what you see here is that the relative risk turns out to be 7.7 rather than 13.9, and the point here is that the relative risk is highly sensitive to the selection of the breakpoint, and you might expect that on the basis of the confidence intervals that the sponsor showed us earlier.

So, it is inherent to the use of these relative risks that we acknowledge the uncertainty that surrounds them.

[Slide.]

Next, I want to go on to a scatterplot of survival for the patients in the study who had a fully defined prognosis, that is, the patients who are dead, and I want to point out that this slide is different from the one in the handout. The one in the handout was mistakenly introduced, and it should be crossed out. So, this is the correct

slide.

What you see, as we have mentioned before, is that there were two patients who had a labeling index less than 8, however, they died in spite of the good prognosis that you would think they might have based on labeling index. So, this is an indication of misclassification on the basis of using labeling index in the case of the individual patient.

Looking at the other side, patients who have extremely high values of labeling index without having a recurrence, who are as many years as eight years of follow-up and a value of 34.

I think that you will have to agree that using the labeling index in the case of the individual patient really requires that other information be taken into consideration, that you cannot use labeling index on its own and when you are considering prognosis.

[Slide.]

So, now I have a few general remarks about the clinical use of prognostic factors, and with breast cancer patients, the use of prognostic factors have focused largely in relationship to decisions about adjuvant therapy.

Now, historically, we have been moving to a situation where adjuvant therapy has been used ever more

widely. So, you have to look for examples of subgroups of patients where prognostic factors would find clinical usefulness, and, of course, the sponsor gave an example of that, and I have provided a few examples.

These are not meant to be limiting, just examples only, but it is conceivable theoretically that you could identify a subgroup of good prognosis patients, who could be spared adjuvant therapy.

This is the complementary group to the one that the sponsor mentioned, the group of patients with small tumors, who had a high labeling index, would get adjuvant therapy, the ones with very small tumors and a low index might be targeted to forego, say, cytotoxic therapy.

Another situation where you might use the prognostic factor would be in poor prognosis subgroups, so that you could base a decision to use more aggressive therapy on a prognostic factor like labeling index.

Although we like to consider the possibilities inherent in having prognostic factors available for use, it appears that just having these factors is not always beneficial. There is at least one paper in the literature that has demonstrated that making an increased number of prognostic factors available for clinician analysis does not lead to a more precise estimate of prognosis, but actually

introduces more uncertainty into the prognostication process.

So, in considering prognostic factors, I think we have to ask the question about how much information can we provide to assist clinicians in making good use of the prognostic factor.

[Slide.]

The pitfalls in evaluating new prognostic factors have been very elegantly summarized by Dr. Barry Clark, and as one of the pitfalls for evaluating prognostic factors, he mentions the univariate analysis.

Since no single prognostic factor is sufficiently correlated with outcome to serve as a definitive measure of prognosis, the univariate analysis can be misleading. Individual factors may be alternative representations of the same biological phenomenon, so an integrated prognostic model offers the advantage of adjusting for the correlations.

Also, pitfalls in evaluating prognostic factors include the use of small studies to do that, and the use of studies where treatments are heterogeneous.

[Slide.]

Another consideration in trying to reach an optimized prognostic model involves the uncertainty about

labeling index at the borderlines between prognostic groups, especially dichotomized prognostic groups, so it is conceivable that in using a labeling index of prognostic factor, that you might want to focus on specific segments of the labeling index distribution, very low or very high perhaps under certain conditions.

[Slide.]

So, I have spent the bulk of the time reviewing the larger study at San Francisco, but what I want to do now is to go back to the smaller study in Syracuse, and you will see that again this is a much smaller study, and there were differences in the patient characteristics for the patients in this study.

[Slide.]

The median age was slightly higher, at age 52. There was a greater number of patients in the stage 2 category here. Rather than 49 percent you saw before, there are almost 20 percent more patients in the stage 2 category.

The median value of labeling index here was 6.35, which is somewhat surprising since the people with stage 2 might have been expected to raise the median labeling index in this group.

The median duration of follow-up here was very brief, only 2.3 years, and the receipt of systemic therapy

was also different in this patient population, only 57 percent compared to the 78 percent in the larger study.

[Slide.]

Now, although we have been talking about 28 evaluable patients, there were also 5 patients in this study who were not considered to be evaluable for the following reasons. There was 1 patient who had a benign tumor. There were 2 patients who had cancer other than breast cancer, colon and ovarian. There was 1 patient with an unreadable labeling index, and in one specimen, there was no residual tumor to be read.

So, among the 28 evaluable patients, there were 6 events, 3 deaths and 3 recurrences, and the univariate Cox model was attempted, but the model did not converge. That may have been related to the small size of the data set.

[Slide.]

So, our conclusions about this study are that the data statistically is uninformative because the sample size is so small and the event are few in number, that the size of the study from Syracuse does not allow a determination of whether the two study results are compatible for merging, and that the data from the San Francisco study must stand on its own. There is no substantive advantage to combining the data from the two data sets in terms of evaluating the

results.

[Slide.]

Finally, a few words about safety. For the two studies supporting this application, there were 231 patients who received a single dose of bromodeoxyuridine, and no adverse events were associated with the study drug in the conduct of these studies.

The sponsor has indicated that there were 5,000 other patients who received from 50 to 500 mg/meter squared of bromodeoxyuridine in cell kinetics or labeling index studies, and in this group of people, only 3 mild adverse events were observed.

Other considerations for toxicity involved the fact that this method requires an intravenous administration, and so there could be some potential for problems with that, either a mistake in dose or wrong medicine given.

Another issue that is related to safety involves the fact that bromodeoxyuridine is a mutagen, and the effect in humans who are pregnant is unknown, and so the studies have been conducted with the caveat that patients could not be entered if they were pregnant.

Another consideration, because the bromodeoxyuridine is a mutagen, is to try to assess the

extent of the risk that is taken by this single dose, and it is really hard to do that on the basis of data, however, many of these patients go on to receive cytotoxic therapy and it would seem, by comparison, that the single dose of this mutagen may not be overwhelming compared to the mutagenicity of the therapy for breast cancer itself.

So, in conclusion, for the safety aspects of the application, there seems to be no overwhelming safety problem with the bromodeoxyuridine, and this brings me to the overall conclusions of the review.

[Slide.]

They are that a group of stage 1, 2, and 3 breast cancer patients were evaluated in the San Francisco study and it was apparent that there was a correlation between survival and bromodeoxyuridine labeling index.

The study procedures for assessing relapse-free survival were not sufficiently defined to warrant the use of bromodeoxyuridine labeling index for prognostication of relapse. An integrated multivariate prognostic model with an optimized breakpoint has not been defined, and the potential usefulness of this test in treatment planning has not been established.

So, that brings me to the end of my remarks, and if there are any questions, I would be glad to entertain

them.

DR. SWAIN: Thank you. Are there questions for FDA?

Committee Questions to FDA

DR. SWAIN: Dr. Johnson, I just wanted to ask you again about the 30 patients that did not have follow-up as of I think October '97.

DR. KAREN JOHNSON: Actually, that follow-up cut-off point was updated in the August amendment to July 31, 1997, so there were 30 patients then, but for the October cut-off, there were 54.

DR. SWAIN: Of '96?

DR. KAREN JOHNSON: Yes.

DR. SWAIN: So, there are 30 patients, and what kind of lack of follow-up is there? Is it patients who hadn't been followed for five years or just hadn't had a visit for a year?

DR. KAREN JOHNSON: The majority of those 30 patients would have fallen into the category of follow-up between one and two years before the cut-off, but there were a handful of patients who had had significantly long periods without follow-up.

DR. TEMPLE: I realize there are a lot of ways to analyze these data, but can you say something about the -- I

don't know -- robustness of the choice of cutpoint? For example, if everything was graded 8, and fell off rapidly at 7.5, 9.5, you did show some other values, how flat is the discrimination curve?

DR. KAREN JOHNSON: What I would do would be to go back to the receiver operating characteristic curve, and what you will remember from that curve is that there were multiple cutpoints aside from 9.1 that were nearly as good as 9.1.

Now, as part of that analysis, we did not make a listing of those values, but the relative risk for 8 is known, so we have no burning interest in specifying that 9.1 has to be the cutpoint, however, we would like to see a rationale for the selection of the cutpoint and comparing the methods, so receiver operating curve versus use of the median versus any other potential method that someone might apply.

DR. TEMPLE: The receiver operating curve is an important analytic tool, but it doesn't have the same tangible feel that looking at what the ratio or the survival ratio is, or something like that.

DR. SIMON: Maybe I could comment. I think there is a couple of dimensions of cutpoints, and I think there is a lot of confusion about cutpoints.

When I spoke previously, what I was concerned -- there are two issues. One is how do you assess whether there is a significant effect in some group of patients, and if you try to make that assessment by using a cutpoint, there is a potential depending on how you got that cutpoint to bias that assessment of significance. That is different from the issue of once you have established that there maybe is a significant effect there, then, the question is how does risk vary with your assay or with your prognostic factor.

I personally don't think -- you know, we are going to get into this I guess with the discussion -- but when you get into clinical relevance and clinical decisionmaking, it is not necessarily going to be based on a cutpoint, and so doing lots of things to determine the optimal cutpoint may not be really relevant.

I think what is more relevant is that if there is clinical relevance, then, the question is how does the risk vary with the level in this case of labeling index, and that is I think where you need a representation --

DR. TEMPLE: I think that is what I was asking, that is, if it was in some say a graded risk with labeling index.

DR. SIMON: But it is not an issue of cutpoint, it

is a matter of as you are labeling for a given set, for a woman with negative nodes, ER positive, tumor less than 2 cm, who is not receiving systemic therapy, how does her risk vary as a function of labeling index.

DR. TEMPLE: You actually want to see it for relevant clinical subsets.

DR. KAREN JOHNSON: I would like to go back to your question about robustness, too, and just consider the relative risk and remind you of what the confidence interval was around the relative risk that the sponsor showed, which was really quite broad.

DR. DeLAP: I think the other way I like to look at cutpoints -- and you can correct me if I am wrong -- but it really is a matter of what you are looking for. In other words, if you want to identify, say, a very favorable prognostic group of patients, you may pick a cutpoint that is very strict, say, very low, and only maybe 10 or 20 percent of the patients will fall below that point, but you will have identified a group that has a very low risk.

Alternatively, if you want to identify a very high risk, you may pick a cutpoint that is high, and the receiver operating curve I think just shows you the area over which the cutpoint will have some validity in making distinctions, but within that area, you can pick the cutpoint you want to

make the kind of discrimination.

DR. SIMON: I agree with that, but I think we have to recognize, if there is clinical relevance to a marker, you have to recognize that practitioners and patients may have their own cutpoints, and, you know, that they need to know what the trade-offs are and what the risk is as a function of the assay, but there is no unique cutpoint that is going to be best for everyone.

DR. TEMPLE: Actually, you just said exactly what I want to know. The receiver operating characteristics are not as tangible, at least not to someone who doesn't use that all the time, as it would be to see what the choice of a cutpoint does to the ratio, to the outcome ratio in each of a series of defined subsets, in other words, just a very simple curve that goes from a cutpoint of 1, which shows a ratio of -- it shows nondiscrimination, say -- or all the way to a ratio of 11, which shows the best discrimination, or whether you see it plateau after 8 or after 6 or things like that.

It is a different kind of operating characteristic that I think is what you are saying a person would want to see. Then, you would know whether to say okay, over 8, under 8, it tells you all you need to know, or you might say up to 2 is one group, 2 to 4 is another group, 4 to 6 is

another group. I mean there is a lot of ways to define how to use a diagnostic test, which is all this is in some ways.

Committee Discussion

DR. SWAIN: Discussion? Do you want to go straight to the questions or do you want to make any comment?

DR. SIMON: I guess there are three issues for me. One is the study that was presented or the studies that were presented, to me have limitations in the sense of they are not big enough, there are not enough patients to really answer the kind of questions that you want to ask, and that is the basic problem with those studies.

In other words, to me you would really want to say, okay let's take women who have node-negative disease, who have not received systemic therapy, and let's see whether this assay permits me to identify a set of women whose prognosis is so good that I may want to consider withholding systemic therapy, or let's take a set of women who have node-positive disease, maybe they are ER-positive, and let's see whether -- you know, how their risk, who have received, say, chemotherapy -- and let's see how their risk depends both on number of nodes involved and this, and whether this assay adds to number of nodes.

This data set is just way too limited to do that.

Now, on the other hand, there is a lot more data available about tritiated thymidine labeling, and that does permit one to answer some of the questions, and so although I am critical of what we can do with the data set presented, at the same time I don't feel that we should limit our potential consideration to trying to make BUdR and labeling available to women, just with regard to this data presented, because we know that there is the experience with tritiated thymidine, and we can try to draw on information that says that they are basically measuring the same thing.

So, I temper sort of my concern about the particular small trials presented with that, and then I think it comes down to is there really clinical relevance to the use of labeling index today, and that gets into, well, do we really need, you know, what do we mean, do we really need clinical relevance to be established, and what do we mean by clinical relevance, do we mean something less than just plugging it into an algorithm, which was sort of the straw man, you know, that was presented, but maybe we need something more than just saying it is measuring cell proliferation, that there has to be some kind of level of clinical relevance between those two extremes that we would like to see.

And I sort of see two potential areas of clinical

relevance. One is for the node-negative woman, does this identify a set of such women who you could potentially consider withholding systemic therapy from, either tamoxifen or chemotherapy, and I look then at Silvestrini's results and I say no, you know, it is a big body of data, 2,000 women who didn't receive systemic therapy, and I don't think she was able to identify that subset of women.

So, then, there is the other side of the coin, and does it permit you to identify women who you want to give more intensive, say, systemic treatment, than you otherwise might if you didn't have available labeling index.

The problem there is, you know, are the more intensive therapies of established, you know, bone marrow transplant or things sort of almost bone marrow transplant, are they of established value, and I don't think quite yet they have established the value.

So, then you get to the dilemma, well, okay, so there is not really a clear clinical benefit, and on the other hand, there may be -- that story may change within the next couple of years. So, to me, that is my sort of concern on the issue.

DR. SWAIN: I would just echo a lot of what Dr. Simon said at the beginning, about the small size of the study, the small number of events, and I have great concern

about the follow-up of the patients, the short follow-up in some patients, and really the reliability of the endpoint in the study.

So, I have a lot of problem with the data and extrapolating it to clinical use. As far as the issue of clinical benefit, at least in this study, I don't see how we can determine any clinical benefit here because it was such a heterogeneously treated group over so many years, over 9 years period of time.

So, I would feel like at least in this study, we really can't find clinical utility. As far as more generally finding it, I am sure that Dr. Simon and everyone knows that CLGB is looking at exactly the study he just described, but using S-phase and small tumors with node-negative disease, and then patients who have larger node-negative tumors who are getting chemotherapy, looking at treatment with patients who have the high S-phase, so that study is really being done in a large number of patients in a group study.

DR. DeLAP: I am thinking about this as it is being discussed, and it kind of reminds me of how for some patients now, the importance of axillary dissection or node sampling, or whatever is done, is clearly becoming, in at least some patients, is more of a prognostic issue rather

than a therapeutic issue.

You may have a patient who you are quite confident based on your initial surgical procedure that you are going to offer this patient adjuvant treatment, and the axilla is clinically negative, and yet you may want to do that for prognostic value.

I think the questions, as we get down to the questions, what we are asking in the questions is, does this give you prognostic information that is worth having, in other words, I don't think the questions are asking whether you can make any treatment decisions on this, because I think we have already concluded that you don't have the information here that tells you how to treat the patient, and so it comes back to does this offer the prognostic information and is that something that is worth having.

Those are the focus of our questions.

DR. SWAIN: I would just repeat what I said basically. I feel like it is unreliable with the survival follow-up not in 30 patients, and heterogeneous treatment, so it is hard for me to really answer yes to that question.

DR. OZOLS: I think that while the issue of prognostic information, what you articulated is clear, if this test is available, it will be used, I think unfortunately by some, as a discriminator of a treatment, so

I think that is the potential down side, that the patients will be treated on the basis of this without appropriately being factored into all the other factors, you know, prognostic situation, so there is a potential for misuse.

DR. TEMPLE: I suppose that is possible, but right now I imagine people would be treated on the basis of their stage, things like that.

What this tells you, if I looked at it, is that this index is more informative than stage, and so I was curious to hear the number of patients described as small. Now, small depends on how many endpoints there are, not how big the study is.

If you look at the stage 2 figure -- I am just looking at their figure number 41 with about 100 patients -- it is very easy to see the difference in survival, with about 30 percent or something like that.

DR. SIMON: But that is not the issue, Bob.

DR. TEMPLE: Does it guide your treatment? No, we have assumed it doesn't guide your treatment.

DR. SIMON: No, no, even that, because when you look at survival, it is confounding all of the other variables, and you are looking at it in a univariate sort of way, and stage is sort of a straw man, too, because people don't treat patients based on stage. Stage doesn't even

take into consideration the number of nodes.

DR. TEMPLE: I guess the point is that in whatever the number of nodes in the people characterized as stage 2, over the course of their follow-up, which isn't all that long, I suppose, there is almost no deaths in one group, and there is a lot of deaths in the other. I mean it is a very wide discrimination using this test alone, which means there can't be any subsets that go the other way. There aren't enough deaths for them to do that.

DR. SIMON: This is a very heterogeneous set of patients, heterogeneous set of treatments, there is lots of prognostic factors floating around in this. I feel still a bias selection of a way of dichotomizing the things to present those curves. So, I don't agree.

DR. TEMPLE: Help us understand. There was only one death in people who were at the low end of the index.

DR. SIMON: So, if you had used a cutpoint that was lower, you would have had a lot of those patients who did well would be above it.

DR. TEMPLE: Well, maybe. I think that is the question I was sort of asking before. I don't think that is so. It's very flat.

DR. SIMON: You would have because a cutpoint of 8, there are a lot of patients who had labeling indices

around 8, so if you had moved the cutpoint down, those women, all of whom survived, some of them would now be in the high labeling index group.

DR. TEMPLE: Well, if that were true, that would be a problem, but we need to look and see whether that is the case. I don't think it is.

DR. SANTANA: I thought the FDA presentation did show the distribution of patients between two units very clearly. There was a graph that was shown by the FDA where the patients were distributed in two-unit increments, so you are correct. If you cut it at a lower point, a lot of those patients that are survivors are going to be in the high risk group.

DR. TEMPLE: That is for the whole population that we don't know about, stage 2.

DR. RAGHAVAN: Coming back to Bob's question, Bob Temple's question, one of the problems is it looks at first glance, and maybe second glance, it is very simple to say there is only one or maybe two deaths when you take an labeling index of 8, but the problem I think that Dr. Karen Johnson alluded to is the fact that the confidence intervals for risk were really quite broad. It was 3 to 68. So, that means at an increased risk.

The difficulty then comes in, in terms of

identifying the level of confidence of that observation, and so if you then take a relatively small data set that is heterogeneously managed to produce that data, and you then say how confident are we that this can be applied usefully, clinically, to reproducible data sets, that is I think the problem that several of the committee members are having.

So, in looking at the data as presented, it looks really potentially quite interesting, but with small numbers, the fact that the second data set essentially become uninformative on top of the first data set, other concerns that relate to such issues as the definitions of toxicity, because that comes into it, how are the toxicity data generated, were they generated in a fashion that would miss lots of minor toxicities.

You then start to get a higher level of lack of confidence in what has been presented.

MS. CARROLL: I guess as the patient representative, I will just stake my claim as having breast cancer recurrence twice now. I don't see where it offers a lot of value to the woman except from a standpoint of if I am under the cut-off, I might not have to worry as much as if I am over the cut-off.

They stated that it doesn't have any clinical or therapeutic values, so therefore, it has little of

importance except as to where I stand as far as a recurrence or perhaps not, but there is nothing to determine specifically whether you will or will not regarding wherever you fall on that scale as to what is ultimately going to happen with your life.

DR. TEMPLE: Can we be clear on that? I mean one of the questions is whether having -- let's say for the moment that it was a useful prognostic indicator, and that it told you something that other information didn't, let's say that that was true for the moment, not that that is true, but let's say that -- are you saying that is not worth having?

MS. CARROLL: It is worth having, but it is not going to necessarily conclude what is going to happen, your therapy, your treatment is what is going to make the difference.

DR. TEMPLE: Well, actually, there is some uncertainty about that statement, you know.

MS. CARROLL: Right. That's the whole point.

DR. TEMPLE: Once you finish your surgery and finish your adjuvant therapy, but this probably tells you more than what your treatment tells you.

MS. CARROLL: This tells you more than what your treatment would tell you?

DR. TEMPLE: Yes. Once you have done the things you know work, which you are probably going to do anyway. It is not so clear how much -- from data we have -- how much difference the treatment makes.

But I want to ask a different question. Let's say this does not tell you how to treat, let's assume that. Everybody seems to believe that, and the company doesn't claim differently. Is something that merely tells you what your prognosis is of value? Even if that is all it did, if it did that, and if it did it in a way that you couldn't get from staging and other stuff, would that be of value? That is sort of the fundamental question.

MS. CARROLL: But I don't see where it is telling you something different that you can't get from anywhere else.

DR. TEMPLE: I am asking you assume that it did for the moment, would that be of value?

MS. CARROLL: It doesn't. Why assume?

DR. RAGHAVAN: Why don't we try answering it a different way. Let's just take the assumptions in your question. So, if this were an independent prognostic variable that gave you information within other subsets, so let's say, for argument's sake, that you accepted that estrogen receptor-positive patients had a different

prognosis, let's call it a better prognosis from estrogen receptor-negative patients, and you wanted to dissect estrogen receptor-positive patients, and this test did it, then, I think the answer is yes, that would be helpful if it gave you reliable prognostic information that you didn't get from elsewhere.

Now, you would then to start to get into the trap of comparative data acquisition. In other words, having a technology such as looking at percent S-phase done an in optimal way, which is a non-invasive test versus a test that requires an injection of a substance, so you can't answer your question unless you broaden the frames of reference that, as I understand it, the FDA already defined.

So, they told the sponsor that the sponsor didn't have to worry about comparative testing for whatever reason. So, I think if you want your question answered, then, sure, if you get an independent prognostic variable that is a new one, and particularly if it is more powerful than the other non-prognostic variables, then, it identifies a subset of patients where you might try harder, and the one that I would see down the track would be the group where you might want to broaden the indications for transplant, if transplant is proven to be of benefit, or for some new drug.

But, unfortunately, what we are stuck with, while

it's an acute hypothetical question, the problem that we are stuck with is N equals less than 300 with five or seven different regimens of treatment, some of which are unspecified, and no real way that we can sift out whether they are non-random variables that would confound this observation with blackboard confidence intervals.

So, what you are asking is a perfectly reasonable question in a perfect world. If you then take it further and say, okay, so what data does the sponsor need to demonstrate that they have got a winning prognostic factor, I think that going back to the Silvestrini approach, uniform patient set, treated uniformly, identifying where there is really new prognostic information with adequate follow-up.

DR. DUTCHER: I think that is why the inter-group study uses S-phase by flow, but everyone is on a treatment protocol that is identical, and they are all getting follow-up that is identical, and there are going to be endpoints that are going to be looked at to determine if that has prognostic significance.

DR. TEMPLE: They are asking a different question, aren't they? They are trying to find out whether intervention -- they are trying to test an intervention.

DR. SWAIN: Well, actually, they are doing both. There is one group that is not treated, the small tumors.

DR. TEMPLE: Okay. But they are trying to see whether treatment benefits. They are comparing treatment with no treatment, so it is particularly important to them to have groups that are comparable.

DR. SWAIN: Well, actually, in one group they are just trying to see if they can find prognostic factors in those small tumors that will determine who should go on to get treatment.

DR. TEMPLE: Okay. And since they are not treated at all, they have uniform treatment you are saying.

DR. SWAIN: Right.

DR. JUSTICE: I would like just to respond to one of the Dr. Raghavan's comments.

I think this was presented to us as this is the data we have, what can we do with it, and we did discuss doing prospective trials to really try to find a better role for this agent, but I think it is just not feasible for the sponsor.

DR. TEMPLE: The matters like sensitivity of the cutpoint that Rich raised seemed very important, but for the moment let me ignore that one, too. The differences between people with the factor and without the factor here seem very large compared to anything related to treatment as we know it.

So, I guess I would ask how important is the fact that the treatments were not different. I mean all of these people had whatever surgery was appropriate, I guess, but we are not talking about different kind of chemotherapy. None of them have effects that are this large in anything we have ever seen, so important is that aspect of the question? Again, deferring the question of whether there could be spuriosity induced into the whole thing because of the choice of cutpoints, leaving that question aside.

DR. DAVID JOHNSON: I don't think the cutpoint is the issue that any of us or at least I am wrestling with. I mean we have data that I think are of proven utility in assessing patients, and I don't know that -- I mean they have even told us that 60 patients, they don't know certain, I think very basic, bits of known prognostic data that could be all lumped into one group here, which could account for that marked difference that you have ascribed hypothetically to the labeling index, and I think this the point that Rich keeps coming back to.

DR. TEMPLE: This is histopathologic grade.

DR. DAVID JOHNSON: Well, certainly grade would have a major impact on outcome. I mean there are groups of women who have low grade tumors that one would not treat, that might survive for 20 years before they recurred. In a

small group of patients with high grade tumors, small lesions that might recur quite quickly.

I mean those are some of the issues that we are wrestling with, that those data are not included in the analysis as far as I can tell.

DR. SIMON: I mean I think labeling index is clearly a prognostic factor. Now, you know, because I basically believe the tritiated thymidine experience, you know, it is just that -- so I mean that is what I keep going back to.

I guess I would have liked to have seen some comparison between in vivo BUdR labeling and in vitro tritiated thymidine labeling just to show that basically you are getting a linear relationship, but I am sort of assuming that you are and therefore I am believing the tritiated thymidine experience, and that is that it is a prognostic factor for most subsets of patients.

My difficulty with it, though, is finding a treatment-related clinical relevance, and I don't see that there is. In terms of whether it is of value to a woman to know that, I mean I am not trying to comment on that.

MS. BEAMAN: Do I hear someone saying that a small, less than a centimeter malignancy, no lymph node involvement, would constitute a different method of

treatment if the labeling index is above or below the 8? Is that what I am hearing?

DR. SWAIN: No, I was describing a study where they are looking for prognostic factors in the small tumors, and those patients aren't being treated based on any factors at all, they are just on the observation arm.

MS. BEAMAN: On observation.

DR. SWAIN: Yes, they are all on observation to see if they can find prognostic factors to predict which of the small tumors will recur.

DR. TEMPLE: No one is alleging that this tells you or should tell you to change your treatment. People have speculated that it might anyway, and they are a little nervous about that because it is not clear that that is merited without the actual studies, but the contention here is not that it tells you how to adjust your treatment. It just gives you a better idea of how you are going to do, that is all. That is all this can support now at best.

DR. DUTCHER: Shall we discuss Question No. 1?

DR. SWAIN: That is what we have been talking about, yes.

You can all read the initial paragraph.

The broxuridine labeling index breakpoint of 8 was based on the median value for 163 patients with primary

breast cancer evaluated at a single institution. There is no information in the NDA linking broxuridine labeling index with choice of therapy, nor is such information likely to be forthcoming. Does the broxuridine labeling index provide clinically meaningful information for physicians and breast cancer patients?

I would answer no based on my previous comments.

DR. SIMON: I am not sure what -- I can't answer really whether it would provide clinically meaningful information to a woman with breast cancer, but in terms of I think of clinically meaningful in the sense of it helping with treatment decisions, and I don't think we have any evidence at this point that it does that.

DR. TEMPLE: Rich, you have got to be specific. If that is what you want to say, please say it.

DR. SIMON: I have said it many times.

DR. TEMPLE: No, but say specifically I don't think prognostic information alone independent of treatment information is clinically meaningful. Is that what you mean?

DR. SWAIN: This specific prognostic factor based on the data we are seeing.

DR. TEMPLE: Okay. It is important to separate --

DR. SWAIN: That is different than just saying any

prognostic factor.

DR. TEMPLE: That is fine. Let's be sure that we understand the whole answer. One, that this doesn't provide any information. Two, prognostic information alone is not useful. Even if the question doesn't say that. I just want to be sure we understand what you are saying.

DR. SIMON: I am not sure of the distinction you are making.

DR. DUTCHER: Prognostic information is useful. I don't think anybody will say it is not. I think it needs to be tied to clinical outcome. The problem here is the heterogeneity of the data that has been presented versus an ongoing study looking at similar questions with a different technique versus this technique in in vitro studies which didn't help us with clinical useful information.

DR. TEMPLE: I am sorry to be a pest on this. Suppose for the moment that nobody had discovered yet whether the prognostic information that this gives you could be altered by treatment, in other words, they haven't gone the next step and said okay, people with a high labeling index, if you treat them this way, they do better than they would otherwise. Let's say that information is not there, but you believe that it was useful prognostic individual independent of treatment. Those are two separate questions.

DR. DUTCHER: You have got two problems. First of all, if you have a prognostic marker that says if you are above this, you are going to die, and if below this, you are not, that is like trying to decide if you are going to look for the Huntington's gene.

DR. TEMPLE: That is exactly right.

DR. DUTCHER: So, then we have to look at these ladies and say do you want to know that.

The second thing is, though, does this data set make us confident that this particular test provides that definitive information, and I think that is what we are concerned about.

DR. TEMPLE: That is exactly right, but if you answer the first question no, we don't have to go any further.

DR. DeLAP: Maybe it would be helpful to rephrase a little bit and say the first question is does the panel believe that prognostic information per se is useful, and I think that we have heard that you think the prognostic information per se is good to have.

DR. DUTCHER: Tied to a desire to be interventionalists.

DR. OZOLS: Well, prognostic information is very useful when you are designing new treatments and you try to

develop and identify groups of patients when you want to try new things, and it is very important, but prognostic information, when you apply it to an individual patient has become a very different issue. Like Jan said, I mean are you really looking at something that is in the context of what are you going to do with that information, what is that woman going to do with that information.

If it is such a high discriminator that it is black or white, all or none, that is one thing, but there is no such prognostic factor. Some patients at stage 4 do good, some patients at stage 1 do bad.

So, on an individual basis, the prognostic information for a patient is much less important than it is really for -- because that individual patient could be anywhere in that spectrum. But for identifying important prognostic factors, you try to develop new therapies, it is very important.

So, this is important to have a good prognostic factor again more for identifying groups of patients and populations, but not on an individual patient. It become very difficult to use that information.

DR. TEMPLE: So, your view would be that if it didn't have treatment implications that you knew, it would not be useful to have a satisfactory discriminator that told

you what the prognosis was?

DR. RAGHAVAN: I don't think that is what we are saying. I would like to come back to your previous question and say that I am sorry you are being a pest, as well, because you are putting an onus on us that is not a reasonable one.

You are putting us in the role of the naysayers based on poor quality data, and I think that this panel would probably agree that having a really reliable prognostic index does have a utility, and I think that our patient advocates would feel that if they could be given information that let them know -- if they there was in the universe of knowledge information that would let them know what the future held, some of them would want it, and some wouldn't. The one who might want it would say, well, I kind of would like to plan, and the ones who wouldn't would say, well, I am hoping things get better.

But that is not the issue here.

DR. TEMPLE: Yes, it is.

DR. RAGHAVAN: No, no.

DR. TEMPLE: It is one issue here.

DR. RAGHAVAN: Let me finish. The issue here is not predicated on that question. The issue here is predicated on our level of confidence that this is a

prognostic marker that will add to the universe of our current knowledge, and Bob Justice said it. He said the sponsor came to us and said this is the information we have, how can we apply it, can we use it, can we do something with it, and it is not our responsibility in this panel to create data that just aren't there.

So, you are asking us to make statements about a prognostic variable that might be a very powerful one, but also could be a bust. In the range of confidence intervals with the number of confounding variables, one of which is number of data points, you are asking us to redefine acceptability in a way that we shouldn't have to do.

We have said to you repeatedly now that this might be an important prognostic test. We have also said to you that we are unable based on the data presented to separate it from other non-prognostic factors or to rank it versus other prognostic tests, and the reality is we are stuck with the same database that you are, and Dr. Johnson, reviewing the data, has come up with very important statistical questions that relate to the adequacy of the information that we are presented.

So, for us to let it through, it's fine if we happen to luck out, but if we then introduce into clinical practice a prognostic factor that is actually wrong, that

allows a patient with small node-negative breast cancer to have to go through aggressive treatment for no reason because it was wrong, that is a bad thing if that patient dies getting chemotherapy.

And the flip side, if I have a patient who has a 5 cm breast tumor and 4 out of 6 nodes positive, but their LI in this test says no risk, and I say to them, well, all my conventional knowledge tells me you are really at high risk, but I have this one test, based on an inadequate set of data (that the FDA let through), so we are going to watch and wait and hope for the best, that would be a very bad mistake. We can't make that decision today.

DR. TEMPLE: Okay. I am going to understand from that, that most of you at least think an effective, good prognostic indicator, without properties of other kinds, would be a useful thing.

DR. RAGHAVAN: That is what I think.

DR. TEMPLE: You just may not think this is that.

DR. DeLAP: Even if you couldn't use that hypothetical prognostic indicator as a basis for making a treatment decision.

DR. RAGHAVAN: As a clinician, I think being able to offer a patient air-tight information based on our current knowledge about their prognosis, being able to offer

it to them, not insisting that they take it, I think it is potentially useful. There are all sorts of clinical and personal decisions, more the personal ones, that would be useful for a patient to know.

It comes back to what we were discussing earlier, when George Canellos said, Melanoma is a disease that gives cancer a bad name," I mean I think patients with melanoma suffer with that knowledge, because they know they have a potentially very risky thing. Whether that then helps them depends on a range of anecdotal experiences, but that is not germane to the discussion here.

DR. DeLAP: Well, I think we have the precision we are looking for here. We just wanted to make sure whether the way you are headed with the first question was predicated on some belief that just having prognostic information per se was not of value or if you are looking specifically at this case and saying in this case, you don't have --

DR. TEMPLE: It is very important to distinguish those two because people, in pursuing this further, need to know what their hurdle is, if it has to be linked to, you know, prognostic tests, or you have to do a follow-up trial showing that you can intervene and change it, that's one thing. If you need a better test, that is another thing, so

we do need to know that.

DR. DUTCHER: Do you want us to vote on this? Yes. Okay.

DR. TEMPLE: We need to know formally what you think about the last question, but we are going to assume based on the previous discussion that if it provided good prognostic information that was reliable, that would be of value, and now you are voting on whether you think the test, as studied so far, does.

DR. DUTCHER: Well, I think that is what we have to do is change the sentence to say does the broxuridine LI data as presented --

DR. TEMPLE: That's okay. We will understand that.

DR. DUTCHER: -- provide clinically meaningful information for physicians and breast cancer patients.

We have heard Dr. Swain and Dr. Simon's opinion, so let us vote.

All those who would say it does, yes, it does, please raise your hand.

[No response.]

DR. DUTCHER: All those who would say no?

[Show of hands.]

DR. DUTCHER: Nine out of nine would vote no.

The next question is: Is there sufficient evidence to conclude that a single, pre-surgical infusion of broxuridine at a dose of 200 mg/meter squared for in vivo tumor labeling is safe based on the study that was presented?

DR. SWAIN: We certainly haven't seen anything to indicate that it is not safe, but I would really go back to Dr. Raghavan's comment on this, and that the toxicities were extremely low and we don't really even know how they were evaluated in each case, so I am not sure we even have enough information to conclude that it is safe, plus I think in the informed consent it did not include a pregnancy warning although patients were apparently told that, so I think that is also problem. Isn't that correct, Karen, it did not?

DR. KAREN JOHNSON: No, that was included in the review that we submitted.

DR. SWAIN: But it wasn't included in the initial informed consent.

DR. KAREN JOHNSON: The San Francisco study did not have the warning in the consent form about the pregnancy.

DR. SANTANA: I think, in reality, when you are giving an agent prior to an immediate intervention right after that, you have to recognize that the period of

observation to determine safety is limited. Somebody made a comment that if a lot of these patients who received this agent go on to chemotherapy, that confounds the whole issue of teratogenicity and mutagenicity, et cetera.

If all the patients were also having surgery within an hour, you have got to be careful that you don't confound the surgical safety issues and complications of surgery, hypotension, or something that could happen.

So, I think you have to be very careful to define the period of observation of safety here, and that is fair, because you do have a window there of one hour or half an hour where you can determine some of these safety issues.

But I also agree with you that I haven't heard anything really dramatic that says it is not safe.

DR. DUTCHER: Well, the question says is there sufficient evidence to conclude.

All those who think the answer is yes, please raise your hand.

[Show of hands.]

DR. DUTCHER: Five yes.

All those who vote no?

[One negative vote.]

DR. DUTCHER: One.

Abstentions?

[Show of hands.]

DR. DUTCHER: Three.

Do you recommend that broxuridine be approved as an infusion at surgery for labeling index determination to assign primary breast cancer patients to a higher versus a lower risk group? If not, what additional studies should be performed?

Do we have some discussion?

DR. SIMON: Well, I would say that since we answered No. 1 as no, that we would have to answer this as no.

DR. TEMPLE: It is the second part that is important.

DR. DUTCHER: What additional studies --

DR. TEMPLE: -- could resolve this deficiency.

DR. SWAIN: Well, I think if they treated a group of patients, a homogeneous group of patients who are getting a homogeneous treatment, and I couldn't really tell you the number of that, that that would be an adequate study, and certainly with better follow-up and a lot of tight endpoints, and the safety issues that we discussed earlier, I think those would be important, too.

DR. TEMPLE: Do you mean one particular prognostic tumor grade set, something like that?

DR. SWAIN: No, either node-negative, node-positive.

DR. TEMPLE: So, homogeneous here would be node-positive?

DR. SWAIN: Right, and getting the same treatment.

DR. SIMON: How would you feel about if they showed compelling evidence that the BUdR labeling index was essentially measuring the same thing as tritiated thymidine index, and then appealed to the tritiated thymidine prognostic factor studies, that large database, for evidence that it is a valid prognostic factor?

DR. RAGHAVAN: I don't think that would help us. I think that, you know, a sponsor only has so much that it can do, and to send them off down a pathway that could end up with another committee that then said, yeah, but that's not really a surrogate of anything, I don't think that would be helpful.

I mean it is kind of frustrating because reading the submission and listening to Fred and others, who I respect, my guess -- and that is the problem -- my guess is that they are probably on to something. Our role is not to guess, and so it is kind of frustrating to look at data and say, you know, there probably is something there, and I am sure that is why Temple was being irritating, because he

probably feels the same, and he is doing his job.

I think so what you want to see is, in a perfect world, someone who helps the sponsor to do the right study, and Bob Ozols has just muttered to me that what you want to do here is plug it into a good, well-powered adjuvant trial, and what you can then get out of that is a good, solid, multivariate analysis that is unbiased, where you can actually be looking at the utility of this test versus ploidy, this test versus estrogen receptors, et cetera, where you have uniform management and in the era of current management, it gives us a reality.

The problem is that if you try to get surrogate steps, if you go back and say well, let's compare this versus thymidine, you know, that will take a year or two, and then you will probably get a committee that will say, yeah, but everything has changed, so they get nothing for their buck.

DR. TEMPLE: Could you be specific on what trial? Sandy, you said what a narrow group would be, like node-positive people who were treated more or less similarly. But that is still a one-arm study, that is like this one, but looking for prognosis.

Is there a controlled trial that you think would be informative here?

DR. SWAIN: I think it would be really hard to do a controlled trial in that you would base the treatment on your result.

DR. TEMPLE: It beats me.

DR. SWAIN: I don't think you could do that.

DR. TEMPLE: This is what people are talking about, and I am trying to figure it out.

DR. SWAIN: I think pretty much everyone now, even with node-negative diseases, getting chemotherapy, so I think that -- I mean you could do it in one of the larger NSABP studies or one of the other cooperative group studies if you were that inclined on this specific factor.

DR. TEMPLE: Before you leave that, it could be part of an ongoing trial.

DR. SWAIN: Right.

DR. TEMPLE: And you could even do it in both arms, so you can see if the treatment makes any difference, and still see if the prognostic indicator was --

DR. SWAIN: Right. Just one other thing I wanted to add. I think some questions came up about that at the beginning, was about inter-laboratory variation. I think those kind of studies need to be done, too, to look at it to see if there is any variability or what the consistency is in the result.

DR. DAVID JOHNSON: Since Bob has been asking us a lot of questions, let me ask Bob a question.

DR. TEMPLE: You have got to be Bob-specific.

DR. DAVID JOHNSON: Well, I will be Bob-a-leftic here and go to that side.

In one of the CALGB trials, adjuvant trials, HER2-NEU was assessed, and found to be suggestive at least of the worst prognosis in women that had been treated in a fairly comparable way, and more importantly perhaps, in that same trial, the data suggests that HER2-negative women were not benefitted by any alteration in dosage of adjuvant therapy, whereas, HER2-positive patients perhaps benefitted by a higher dose.

Now, this is retrospective analysis of data and needs to be prospectively analyzed, but in a sense, HER2-NEU is sort of the same thing here. It is a prognostic factor. That is the question you were asking earlier, is it important to know that.

At the moment, the answer to that question is for me no, it is not very important, because I don't know what to do about it yet. I have some ideas about what to do about that. I kind of have an idea that is what you were driving at with the labeling index here.

I think it would be very difficult to append this

technique to an ongoing adjuvant trial in the cooperative groups. It might be possible to do. It might be possible to do in an NSABP trial perhaps, but I think it would be very difficult to do that, and short of doing that, I am not sure what specific trial one could do to get at the kind of information all of us I think would like to have.

This is the dilemma that we wrestled with the other day with our colleagues who were worried about neoplastic meningitis.

DR. TEMPLE: Well, I guess I can think of one. If you were doing some intervention trial with two treatments, say, and there was a period in which you had to remove the tumor, so you therefore had an opportunity to do this labeling, you could administer it to both arms, and from each arm you would get prognostic information, one with one treatment, one with the other.

You could also, if you incredibly lucky, see if in some way there was an interaction between the prognostic variable and the treatment.

DR. DAVID JOHNSON: Well, what I was going to say is NSABP does have a trial ongoing at the moment for pre-op chemotherapy. That was the point I was going to make, and I don't know how far along that trial is in its accrual goal, but if there ever were a trial in which this type of

approach might he reasonable to append to a trial, that would be the one, because those are women who are getting chemotherapy, who are all undergoing the surgical procedure, and are going to get injected with something anyway at some point, and one might be able to make a strong argument to do that. I don't remember the number of that trial. Sandy, what is that?

DR. SWAIN: B27.

DR. DAVID JOHNSON: Do you know how far --

DR. SWAIN: I think there are about 700 or so patients on the trial. I think that is a nice idea, I don't think it is going to happen because I think the accrual -- it is hard to go back, I mean that is the bottom line.

We just had a meeting and talking about a lot of different things, so I don't think that is going to happen in that study, and the current trials will not be preoperative. I think the preoperative might make it more difficult anyway because then you would have to do more biopsies. So, I don't think that is ideal. I think probably the postoperative adjuvant therapy would be best.

DR. TEMPLE: But I guess you have to get this all done before the opinion.

DR. SWAIN: The problem I have with it, too, frankly, is that it is I.V., and it is hectic, it is busy,

people are delayed for their surgery for hours, and I just think it is just one more thing for the surgeon and the patient to undergo, that really right now I don't find any reason for, and there are other possibilities, proliferation markers, to me, that would seem more interesting, that you can just look it under paraffin or whatever. I know that is not what you asked.

DR. TEMPLE: Well, that has to do with practical ability to get this off the ground and whether anyone will use it, which you are right, we don't usually worry about that. That is somebody else's problem.

DR. DUTCHER: I think the logistics are an issue because it is hard to append things onto ongoing studies, and yet those kinds of studies would be the best way to find out. I mean we have all been down the road, if something is a wonderful marker, and then we find out that nobody can figure it out, and you are still better off having a set of patients that you can say were as uniform as you could make them.

So, I think that would be in the best of all possible worlds, the best way to be able to look at it.

DR. DAVID JOHNSON: What about taking that same set of patients, a smaller set of patients -- and I won't define the set of patients -- but I am just trying to think

now like an investigator and to do what Dr. Simon was suggesting, and actually do thymidine labeling and this labeling, and see if there, in fact, is a correlation, and then perhaps, with a certain set of patients -- I don't know what that number would be -- but one that would give you a certain degree of confidence if there was correlation, and then perhaps you could extrapolate it.

It wouldn't be quite the same thing, Derek, it seems to me. I agree with your earlier comment about putting another committee in the same dilemma, but it seems like if you could make that correlation with a set of patients, one might feel more comfortable extrapolating it.

DR. TEMPLE: That only really works if you are confident that you can say something about the thing you are linking it to, which we haven't gone through. We don't know that.

DR. SIMON: Silvestrini has data, in other words, I alluded to the 2,000 -- actually, there were 3,000, but 2,000 were follow-up women who had no systemic therapy. She has another group of women who received tamoxifen as the only systemic therapy, and another group who received CMF plus or minus adriamycin, so these are very large data sets in which the minimum I think is probably 500 patients, in which the patients are very uniformly treated.

So, to me, those body of data are relatively compelling that the tritiated thymidine labeling index is a prognostic factor that adds to the other prognostic factors, although it doesn't permit you to make treatment decisions.

DR. TEMPLE: So, if one then could try to link some outcome on the tritiated thymidine index to an outcome on this index, and show that they correlate well, you think it is at least possible with the large database available for tritiated thymidine, that there would be something really to go from one to the other.

DR. SIMON: Yes. I think the potential value of this drug is that it may give more reproducible results than the KI67 and the other things that are going to be done just on tissue blocks.

DR. TEMPLE: One other question. On the matter of homogeneity, it seems to me in some ways what you want to know is how this performs in a variety of settings, not just one, so what I am taking from this is that you would like to be able to look at this in defined groups of people, but there might be more than one of them, there might be several of them.

DR. SIMON: Right.

DR. DUTCHER: I think that is it. Thank you all very much. Have a happy holiday. See you in the new year.

ajh

[Whereupon, at 3:40 p.m., the meeting was adjourned.]