disease-free survival curves.

The agency once again performed estimates of 3-year disease-free survival. The results, 81.6 percent for the AC treated patients and 81.2 percent for the AC plus Taxol treated patients, provide no evidence that the Taxol treated patients who had ER positive and/or PR positive tumors evinced any benefit from the addition of four cycles of Taxol in adjuvant therapy for their node-positive disease.

The findings in the ER positive and/or PR positive subset of patients prompted the FDA to perform an additional analysis on those patients who had hormone receptor positive tumors and received tamoxifen. Even though this represents a more specific subgroup than the previously identified group, it consisted of a sizable number of patients at close to 2,000. The analysis of this subgroup is even less suggestive of a trend toward Taxol effect with a hazard ratio of close to 1.

The most closely related analysis performed by the sponsor is disease-free survival in all tamoxifen treated patients. As can be seen in the sponsor's graph, there is no appreciable difference in the disease-free survival curves for Taxol treated patients compared to the control group.

In summary, the agency is in agreement with the

sponsor on the overall positive effect of Taxol. However, these overall positive results are based on the findings in the ER/PR negative group of patients. The evidence for a Taxol effect in the receptor positive or tamoxifen treated patients appears to be insufficient.

In this trial, the efficacy endpoints were disease-free survival and overall survival. Objective disease relapse was used to evaluate disease-free survival and was defined as the appearance of local recurrence or distant metastases at any site or death due to any cause. The most common reason for failure was the occurrence of distant metastases, with the second most common reason for failure being local disease recurrence.

Taxol demonstrated efficacy in decreasing the odds of both distant recurrence and local recurrence. This chart shows that the effect of sequential Taxol in decreasing the odds of recurrence was similar for both distant and local rates of recurrence.

Before I go on, I would like to present a quick overview of the other definition for objective disease relapse in this protocol which was death due to any cause. At a median follow-up of 30.1 months, a total of 342 deaths had been reported. 192 deaths had occurred in the AC treated group, which is comparable to 12 percent of the population, and 150, or 10 percent, of those treated with

AC plus Taxol had died. The corresponding percentages of survivors are shown on the right-hand side of the figure.

As we saw in the analysis of disease-free survival, according to the three identified subgroups, when we interpret the results in overall survival with respect to the same three subgroups, a similar pattern emerges. The positive results for the entire study population are driven by the very noteworthy beneficial effect of Taxol in the ER negative/PR negative population.

The first graph, this graph, and all subsequent graphs were taken from the sponsor's submission. This first graph compares overall survival in receptor negative patients treated with AC versus AC plus Taxol. Those treated with sequential Taxol derived a substantial survival advantage. Sponsor and agency hazard ratios were consistent. The sponsor reported a hazard ratio of 0.72 with a corresponding p value of 0.11.

In those patients with ER positive and/or PR positive tumors, there was no appreciable difference in overall survival when the AC treated group was compared to the AC plus Taxol treated patients. The sponsor calculated a hazard ratio of 0.83 with a corresponding p value of 0.31.

The lack of evidence for effect with sequential adjuvant Taxol after 4 cycles of AC is even more pronounced

when comparing AC treated versus AC plus Taxol treated patients who had hormone receptor positive tumors and received tamoxifen. The sponsor's hazard ratio of 0.92 and p value of 0.63 reflect all patients treated with tamoxifen.

Since the reported toxicities for AC were comparable and occurred with equal frequency during the AC part of treatment in all patients, I will not repeat them here. Instead I will focus on the toxicity associated with 4 additional cycles of Taxol, which is not without risk.

The early population, as the sponsor indicated earlier, consisted of the first 325 patients that were accrued to the trial. The protocol specified complete reporting of all adverse events that were grade 2 or higher for this cohort of patients. Therefore, the figures in blue represent the most accurate toxicity profile for Taxol in this trial. The incidence of adverse events were reported as the worst grade per patient. This does not tell us if the same worst grade toxicity recurred in subsequent cycles of therapy. Women of all age groups experienced more non-hematologic toxicities with the addition of Taxol. The risk profile is expected based on the known toxicities associated with the use of Taxol with the most notable toxicities including hypersensitivity reactions, neurosensory events, arthralgias/myalgias,

diarrhea, and neuromotor toxicity. In summary, the impact of 4 additional months of therapy should not be discounted. The women suffered some morbidity and some decrease in quality of life.

82 patients, or 6 percent, of those randomized to treatment with AC plus Taxol discontinued therapy during Taxol due to drug-related toxicity. In comparison, 15 patients withdrew from therapy in the AC arm, and 17 patients randomized to the AC plus Taxol regimen withdrew during the AC portion of their treatment.

2 patients died acutely from Taxol toxicity. 1 patient had a brain infarct subsequent to sepsis, and 1 patient experienced a hypersensitivity reaction. The patient who died during AC treatment died of respiratory disease which was attributed by the investigator to disease progression and not related to drug toxicity.

Some issues to consider. For the entire study population, the overall results of the trial are very positive. The use of Taxol reduced the recurrence rate or risk of recurrence by 22 percent with a hazard ratio of 0.78 and reduced the risk of death by 26 percent with a hazard ratio of 0.74.

Although the FDA usually views subset analyses with trepidation and great caution, the agency feels that the results in this trial with respect to the identified

subgroups are compelling. The subgroups are large, with a notable number of events occurring in each. The subgroups represent medically plausible populations. In fact, the protocol specified different treatment for patients in each subgroup to receive or not receive tamoxifen.

And finally, the overall results of this trial seem to be driven by the findings in the receptor negative population treated with Taxol.

Furthermore, 4 additional cycles of chemotherapy are not without risk. As we saw, 82 patients discontinued Taxol therapy because of drug-related toxicity and 2 patients died acutely of drug-related toxicity during Taxol therapy. Based on these data from an interim analysis, it seems to me that the lack of evidence of a Taxol effect in patients with receptor positive tumors treated with tamoxifen would not justify the added toxicity of 4 additional cycles of Taxol chemotherapy.

In summary, based on the current interim data, the net beneficial outcome in disease-free survival and overall survival reported for all AC plus Taxol treated patients appears to be derived from those patients with tumors that were hormone receptor negative for both estrogen and progesterone. This group comprised about one-third of the entire study population.

I believe there is sufficient evidence to

approve Taxol as adjuvant therapy subsequent to the combination of doxorubicin and cyclophosphamide in patients with node-positive breast cancer who have tumors that are negative for both estrogen and progesterone receptors. This recommendation is based on the striking improvement demonstrated for disease-free survival and overall survival in this subgroup.

Two-thirds of the study population had tumors which were hormone receptor positive. Per protocol these patients which received tamoxifen at the first interim analysis of this trial, there seems to be no evidence of benefit from 4 additional courses of chemotherapy with Taxol after AC in patients who will receive tamoxifen. The effect of Taxol cannot be discerned in this group of patients.

Therefore, based on the currently available interim data, I do not believe there is sufficient evidence to recommend approval for Taxol as adjuvant therapy sequential to the combination of doxorubicin and cyclophosphamide in patients with node-positive receptor positive breast cancer. This recommendation is based on the near unity in the hazard ratio and no trend toward statistical significance, along with 3-year disease-free survival estimates showing no difference. I must say that the result of future interim analyses and/or the final

analysis may alter this current recommendation.

Thank you for your attention.

DR. NERENSTONE: Thank you. We'll now open up to questions from the committee. Dr. Williams?

DR. WILLIAMS: I want to make a statement for the team.

I think we've had a very good discussion with breast cancer experts and with the company and the team's presentation.

We made a recommendation here but I really think that at this point in time we're really more asking what's the right thing to do. I really think that this is a very tough call. I just wanted to sort of communicate the FDA's current position on this.

DR. NERENSTONE: Thank you.

Dr. Kelsen?

DR. KELSEN: It seems to me that the major issue that you've raised, since there's general agreement on a recommendation for non-estrogen receptor and progesterone receptor patients is now bad is it to take 4 cycles of Taxol for ERP/PRP positive patients when we don't yet have full evidence of benefit, but you're basing it on a subset analysis.

As I look at, I guess it's slide 32 from the sponsor's presentation, looking not at grade 2 toxicity but

grade 3 or 4 toxicity because, although no one wants any toxicity, the key issues are serious toxicities, those numbers are very small for the AC followed by T arm for serious grade 3 toxicity unless I'm misreading this, either hematologic or non-hematologic toxicities.

DR. O'LEARY: I believe it was in the range of about 15 percent --

DR. KELSEN: Leukopenia, 9 percent; granulocytopenia, 21 percent; less than 1 percent or 1 percent for everything else, including cardiovascular, nausea, vomiting, whatever. Slides 34 and 32.

So, we're basing our recommendation to not give therapy to ER or PR positive patients on a subset analysis with trends that are slightly below the unity point. And that's not a very comfortable feeling to withhold therapy that may change the cure rate. So, you have to be pretty comfortable I think that it's the right thing to do because it will be several years before we know for sure that this is not effective therapy in making this decision.

DR. O'LEARY: The next interim analysis will occur? Can the sponsor tell us?

DR. BERRY: The 900 will be probably 12, 18 months from now. I'm not sure.

DR. CANETTA: If I can just make a point. I wonder whether it is appropriate to call these interim

analyses because the definition of interim analysis applied to the stopping rules for the protocol. This study that's been reported has not been stopped. It has been completed. So, I don't think that there is a compelling reason to go back to 900 events or 1,350 events as the protocol wrote that would have been done in the event that the protocol had to stop. The protocol has been completed.

DR. WILLIAMS: I think the protocol was designed to perform analyses based on number of events, and I call that an interim analysis. I think we'd be interested in the data as they were designed to be collected and we would make decisions based on those at each particular time. I'm not quite sure I understand your distinction. Certainly we can't stop the trial, but we're certainly going to look at the data when there's twice as much as there is now.

DR. NERENSTONE: Dr. Temple and then Dr. Margolin.

DR. TEMPLE: I don't think interim here was meant to imply that there's anything wrong with it. I think Dr. O'Leary was just expressing the hope that perhaps with more data, there might be a benefit seen in that subpopulation. I imagine everybody sort of hopes for that. It wasn't a statement that the data aren't persuasive for some information now.

DR. NERENSTONE: Dr. Margolin.

DR. MARGOLIN: I think, although we don't have this data and we won't really from this study or maybe from the next or next after that early trialists group, we have to consider that the addition of Taxol is going to have an impact on all groups similar to the addition of chemotherapy to hormonal therapy in patients with ER positive disease.

Since there are often different levels of limitation or caution that can be placed on drug approvals, one option that we've seen the FDA do sometimes -- and I would wonder if that's being considered -- is not to limit the actual sentence that's written for the indication in the approval, but to have very prominently in the package insert the data from this trial cautioning that the proof of benefit of Taxol in the ER positive patients who receive tamoxifen has not yet been demonstrated beyond all doubt.

DR. NERENSTONE: Dr. Justice?

DR. JUSTICE: The answer to that question is yes. We can put in the clinical study section, if the committee recommends, a full disclosure of the issues. We have that in indications as well, but definitely in the clinical/pharm.

DR. NERENSTONE: Dr. Kelsen.

DR. KELSEN: This is a procedural question.

We're hopefully going to see large scale trials in a number of solid tumors over the next few years, many of which may not have a subgroup analysis planned of this type. What will the position of the agency be, let's say, if we do a colon cancer trial and we're lucky enough to get 5,000 patients in it? And there are a number of subgroups in colon cancer. We're not going to do subgroup analyses in all of them. How shall we approach that as these adjuvant trials come through?

DR. NERENSTONE: Dr. Temple?

DR. TEMPLE: Well, we're usually on the other side of this argument.

(Laughter.)

DR. TEMPLE: We're historically skeptical about subgroup analyses, especially when they try to salvage an otherwise negative study.

I think the theme here is that this sort of grabs you by the hair more than most of them do. We are, in general, resistant to making much out of the many possible subset analyses that show up in trials. So, we have the same attitude that the company is expressing. It's just that when you see two-thirds of a study with a hazard ratio of approximately 1, you sort of have to say, well, what should I do with this? So, I would consider this quite exceptional. We don't usually celebrate the

small differences that are inevitable in any trial. So, it's not a difference in attitude. We're very skeptical. But as Jim said, this sort of grabs you.

DR. NERENSTONE: Are there other questions from the committee?

(No response.)

DR. NERENSTONE: Okay, thank you very much.

At this point, I've been asked to reopen the public hearing and Dr. Marissa Weiss would like to address the committee.

DR. WEISS: Good morning. My name is Marissa Weiss. I'm a physician oncologist specializing in breast cancer, and I'm here today representing my nonprofit educational organization, Living Beyond Breast cancer, which is Philadelphia based but a national organization. Our mission is to help all women affected by breast cancer live as long as possible with the best quality of life.

I am here on my own. I was invited by myself. Bristol-Myers is one of many companies that buys a few seats at our table for our annual gala, which is next week, and all of you are invited. There will be 800 people there.

(Laughter.)

DR. WEISS: I'd just like to start by putting this into perspective. We're all here in the room for the

same reason, which is 40 percent of 180,000 newly diagnosed women with breast cancer with have their lymph nodes involved, and as Dr. Henderson said, of the 3,000 people on this study, over half were expected to have a recurrence. So, this is a large group of women, 72,000 women diagnosed each year, with nodes involved, and over half are still predicted to recur over the long term. So, we desperately need effective treatments for these women.

I am struck by the incremental benefit that Taxol offers to women who have already completed their Adriamycin and Cytoxan chemotherapy. It's very impressive, and the shape of the curves, two parallel curves, over time -- those two points of analysis -- they're identical. But also the curves start to plateau out. So, I feel comfortable with the reliability of that data.

Also, we've had a longer experience with Taxol than just this study. This is not the first study. We have a lot of information about toxicity, not necessarily after AC chemotherapy.

These data do cover the highest risk period in this particular population of women with nodes positive, the first 3 years being the highest risk period. These data are just short of 3 years.

Just to say for all of us in the room who have already given our patients the benefit of Adriamycin and

Cytoxan chemotherapy, what this study does show is at least dose intensification of Adriamycin doesn't buy you anything more. So, we've got this group of women who have gotten the benefit of the best standard chemotherapy and giving more of it doesn't do a damned thing. So, the point is what more can we do for these women that's substantially different, and it seems that Taxol does do that without significant incremental side effects.

Clearly additional chemotherapy being involved for 4 more months, quality of life issues are definitely there. But we all know that for those women on this study -- and most of them are young women in the prime of their lives. They're going to choose it. They can trust that they with their doctor can have a discussion that says, based on this potential incremental benefit in your situation, do you want to accept these additional incremental sides effects. I have to say that the people I represent want to have that option.

In terms of the subset analyses, I'm happy to see that the estrogen receptor negative patient who hasn't had the benefit from tamoxifen over these years and is very envious of the woman who's estrogen receptor positive who gets tamoxifen, but this is really good news for them.

But in terms of the subset analyses, you could really take that pretty far. For example, is there a

spectrum. You've shown us that the women that are hormone receptor negative, both estrogen and progesterone receptor negative, have the greatest benefit. If you look at the women who were either ER positive or PR positive, they don't see as great a benefit. There may be a continuous spectrum of benefit from starting from those patients who were both ER/PR negative having the greatest benefit and those patients who were both ER/PR positive who are also taking tamoxifen and stick with their tamoxifen, they're going to see the least benefit because those people of this group are going to do the best anyway. So, any incremental benefit is going to be hard to measure, particularly over this period of time. 3,000 patients is a lot of patients, but maybe not large enough.

So, these data are very compelling to me, and I am concerned about the subset analyses, and I think if you really want to put weight on these subset analyses, I'd like to see a spectrum of the differential effect that Taxol gives after AC for every combination of the hormone receptor positivity and negativity, starting from all ER/PR positive to the ER/PR negative and the different combinations, different numbers, and also if the patients stick to tamoxifen or they don't because we all have patients who are ER/PR positive who can't take it for some reason or who start taking it and stop taking it. Then

your hands are tied. What more can I do for this woman who's in front of me? We're talking about women whose lymph nodes are involved. You're talking about people whose long-term survival is 50 percent over long term, and we want to make things better.

So, as a physician and as an advocate for the 30,000 breast cancer patients nationally who are members of our organization, I think that Taxol should be approved and be available to the patient and the doctor with an up-front discussion. I really favor this being part of the package insert, where a doctor is guided by the package insert and says, we're in this situation now. You've had the benefit of this. What is your style of making decisions? Do you want to do everything possible today to make sure you never see the cancer again? And make sure that the decision to proceed with this is an informed one.

Thank you.

DR. NERENSTONE: Thank you very much, Dr. Weiss.

Now I'd like to open up the committee discussion. First, are there any general comments from the committee? Dr. Raghavan?

DR. RAGHAVAN: I think everybody has identified just how difficult one part of this is. I came in this morning thinking the FDA were absolutely wrong, and Grant

Williams is a thoughtful reviewer and I was surprised that he would actually do an about-face and allow subset analysis in with FDA blessing and, in fact, castigated him as I arrived in.

But listening to the discussion, the faster Larry Norton talked, the more confused I became and came out of it feeling that maybe he was wrong. He made one statement that troubled me a lot, which is the smaller the sample size, the broader the confidence interval, and that's not a generically true statement. It's only true if you have a scatter of points. If everybody has a similar survival with a small sample size, then the confidence intervals don't widen. It's a small point, but it just got me to thinking that it isn't that simple.

I listened to Dr. Weiss just now and I was thinking that she was oversimplifying things as well.

I think the reality is Taxol is a terrifically useful drug for some people, but it's a drug that causes side effects and people potentially have anaphylactic reactions. And we shouldn't just say this is an all or nothing thing in which it's either all good or all bad.

Now, I think everybody has conceded that in ER negative patients, there's a really substantial survival benefit, both overall and disease-free. That's terrific. It means that for ER negative patients this is a major step

forward, and Larry Norton's conceptual thinking has influenced us on this. And it's a huge step forward, and I think that's great.

Of course, what we're struggling with now is the fact that there's such a major impact on the outcome of that smaller group that it could easily have weighted the overall study. And it's pretty hard not to look at the survival curves and say they really sit one on top of the other, notwithstanding the fact that it's a subset analysis.

I think Dr. Temple's point is a little different because the subset that is being looked at is actually bigger than any other subset in the whole study.

So, in the discussion that ensues, I hope that the rhetoric that we've been hearing doesn't sway us. I think the reality of the situation is there's one group of about 1,000 patients that were ER negative/PR negative and didn't get tamoxifen or, for that matter, did get tamoxifen where the hazard ration clearly favors approval.

It's not quite that simple, I don't think, with the ER positives who got tamoxifen. The question, of course, is if a woman is having chemotherapy and is going through the tail end of it, which is normally when it's the toughest and the most wearing, if they're on tamoxifen, you want to be sure that you're actually giving them something

back for adding Taxol.

So, I'd like to hear the breast experts around the table and elsewhere talking a little more about it, not just to make a very simple one-liner that subset analyses are bad because I think this is one of the more difficult decisions we've had to make at the committee.

DR. NERENSTONE: Any takers?

I'll plunge in a little bit, Derek. I think one of the things as a practitioner that I agree the lack of significant effect is -- "concerning" is too great a word, and I think you're right. There is no question about the ER/PR negative patients.

The survival curves are very close, but there is an effect. The curves never cross, at least not from my non-statistical eyeballing of the curves, suggesting that it is very possible in the future that they will separate. Maybe we should have a statistical discussion about that. What is the likelihood that we will get an effect with more events and further follow-up because I think that's the question. Remember, this is a subset analysis and the study is very positive.

What I think clinicians want to avoid at this point is the denying of patients, possibly curative therapy, although everyone will admit the effect is going to be small, on the basis of a subset analysis where we

know the benefit is going to be small.

Dr. Lamborn, can you comment on that?

DR. LAMBORN: The problem, of course, as has been identified, is as soon as you go into subset analysis, you have to consider how much you believe this is based on prior medical judgment that these groups are going to be different versus you've just taken a whole series of subsets.

But the closest I can come, based on the information you have right now, is to reference back to I think it is the last slide that was in the FDA presentation where they looked at the ER positive and/or PR positive tumors and looked at the 3-year disease-free survival. And you asked did it cross. Obviously, it slightly crossed in terms of disease-free survival because that's 81.9 on the AC plus Taxol compared to 82.7 for the AC group. But they are so much on top of each other, what do you call "cross"?

But the other thing is your hazard ratio, which is a .98, which is pretty close to 1 -- when we were talking about equivalence yesterday, we would have said, .98, wow, they've really demonstrated equivalence. You do see a confidence interval. Again, you have to remember to interpret that in light of the fact that they've looked at multiple analyses.

But that's sort of the best I could do for you

in terms of trying to gauge the potential of what will see, and there's no reason, I guess, to expect that as you move forward, if you believe the modeling assumptions, that you're going to change that number. You would assume that this number is where it will about fit. The confidence interval would get narrower, but the estimate would stay about the same.

DR. NERENSTONE: The sponsor said, in their defense, that they thought the 1-year was more accurate because more patients had gotten to that point. Do you agree or not agree with that?

DR. LAMBORN: To the extent that we're describing where the value will actually be at the end of all the analyses, clearly the 1-year result is not going to change since everybody has moved beyond that point. I don't remember what the 1-year result was for this particular group of patients.

DR. BERRY: I don't think we gave it to you.

DR. LAMBORN: That's why we don't remember it.

(Laughter.)

DR. LAMBORN: Do you have it?

DR. BERRY: But you're talking about the ER positive.

DR. LAMBORN: That's right.

DR. BERRY: We didn't do that. You're talking

about the ER positive, and we didn't show that. We do have it I think.

DR. LAMBORN: I think it would be helpful if we could see it.

DR. HENDERSON: I did show those data, and the point was that I was trying to make was that as you go along, the confidence interval gets wider and wider. I will give you those numbers in just a second here.

There you go. So, you can see that there's a small benefit at 1 year, fairly narrow confidence intervals around each of the estimates, a slightly larger benefit at 2 years, slightly larger but still fairly tight intervals around the estimates, and then no difference at 3 years but wider confidence intervals around both of them. I think that's the data set that you're asking for.

DR. LAMBORN: That is specifically it because I think the issue we're being asked is what do we expect to see down the line. I think the only thing we can say is what we see now is our best estimate of what we would expect to see, and in some instances we're pretty sure of what we're going to see in terms of final data.

DR. BERRY: Excuse me. I want to point out that the reduction at 1 year is essentially what we see overall and, in fact, is better, if you go back to that please. Compare 97.7 versus 96.5. The reduction is about

a third in this ER positive group.

DR. NERENSTONE: Dr. Margolin.

DR. MARGOLIN: I'm sure everybody knows this, but I think we need to remember this business about ER and PR positivity and how positive in measuring, and the interaction with pre- and post-menopausal need to be kept in mind as well as we, those of us who are in the clinic treating patients, have to make a judgment every single time we make a recommendation to a patient about her adjuvant therapy.

The NSABP has tried, in some of their retrospective analyses, to look at their outcomes in various studies as grouped by level of ER and PR positivity, and they've taken the stance in many of their studies prospectively that they don't care. They just put everybody over 50 on tamoxifen.

So, I think that, again, rather than trying to say this is group A and group B, we really have quite a spectrum and it makes more biological sense to look at it that way.

DR. NERENSTONE: Dr. Lippman.

DR. LIPPMAN: Again, since we're about to discuss a recommendation based on a subset analysis and the issue of consistency from meeting to meeting -- it actually came up at the last meeting on another drug. But the issue

that Dr. Kelsen mentioned and I think Dr. Temple indicated his thoughts on this is that one of the things is the idea that one looks very skeptically on subset analyses from a negative trial.

I guess the question I have is, is there any reason to think that there's more importance or less importance or more validity or less validity to a subset analysis based on whether the primary endpoint of the study is positive or negative?

DR. NERENSTONE: Would someone from FDA like to answer that? Dr. Temple.

DR. TEMPLE: For what it's worth, one of the requirements that a sponsor faces in submitting an application is that we ask them to look at whether effects are similar in men and women, old and young, black and white, generally by looking at an overview of the data, pooling the available trials, and looking at those. Now, those are three demographic figures. It's not 20 subsets. It's three. And many people would condemn that and say that's just exploratory nonsense, and you really should pin it down.

But I think there's a feeling that it is worth looking for these things, and if the differences appear very large, you sort of do your best with them. I think most people would say that's the rule on subsets. You

should be skeptical. You shouldn't do it willy-nilly. You should be aware of how many things you're looking at.

So, one of the things you'd consider is how plausible, among the various things one is looking at, would it be to look at other therapy. Well, a lot of times the other therapy people are on is one of the first things you'd consider in looking at plausible subsets. So, a lot depends on whether there's 40 subsets out of which you're pulling it or only one, and medical plausibility and all that.

So, I don't think I could give a rule. We're generally skeptical about these things. That's our rule. But no one would ever say they're never credible.

DR. LIPPMAN: The comment that I was making, because it really did come up at the last meeting and you commented here, is just the issue of whether the study is actually positive or negative in terms of the primary endpoint and whether that changes the validity, statistically or otherwise, of subset analyses.

DR. WILLIAMS: Maybe I could add something. I believe our usual action that we take on the basis of subgroup analyses would be to put them in the labeling. Usually we have a positive trial and usually we would say, well, there seems to be less or more effect. So, we don't have that problem if we have a negative trial. There's

nothing to put in the labeling. So, I mean, maybe that's what we usually did.

DR. TEMPLE: One is also, let's say, more skeptical when the overall trial is negative because the urge to find a subset with an effect becomes overwhelming. Maybe there's less of an urge, maybe this is more spontaneous. These are all nuanced and no good rules.

But I think it's fair to say most people think you should. One of the great things about the overview analyses is there were so many patients in them that you can start to do credible subset analyses. So, Richard Peto who started both this and is a very powerful skeptic of subset analyses -- he's famous for showing that people with -- I guess aspirin doesn't work if you were born under certain zodiacal signs, which he did not consider support for astrology, but support for not doing subset analyses.

DR. JOHNSON: Do you remember which sign?

(Laughter.)

DR. TEMPLE: Gemini was one where it didn't work.

(Laughter.)

DR. TEMPLE: Libra and Gemini. So, those among you for whom that's relevant will know.

But at the same time as he's a known skeptic of these, one of the things you can do when you have 50,000

patients randomized is start to look and perhaps learn something. So, everybody who looks at this has mixed emotions. They all say don't do it, and they, every once in a while, find themselves to be persuaded anyway.

DR. LIPPMAN: But I think we can all agree this is a very intriguing finding. We talked about the value of hypothesis generation and so on. I think the issue really is -- I don't think any of us would disagree with putting this information in the clinical section of it. The question is whether to put it up front to really say that we're sure this should affect patient care as a hypothesis testing point, and I don't think that's what happened here.

DR. BERRY: Dr. Nerenstone, I don't know the protocol here. Can I address some of the things that have been discussed?

DR. NERENSTONE: Dr. Johnson.

DR. JOHNSON: Well, actually like yesterday, we've sort of gone back and forth between questions and discussion. I would like to just put forward some thoughts about this.

Like everyone else, I too am a little bit concerned about the subset analyses. I think had the study shown equivalence, let's say, and then a subset analysis had been done with 2,000 out of the 3,000 patients that was

positive, I'm not sure we would have accepted that as an indication to go forward and approve the product.

I agree with everything that Derek said. The reason I think we're a bit concerned about it is a point that has been made by others about the biological plausibility of the subset, which is a group that was ER/PR positive that got tamoxifen and obviously benefitted from that.

The biological facts are -- and we've known this for a long time -- is if there's difference in how the ER/PR positive tumor progresses, the growth if you will, the kinetics if you will, of that tumor, therefore the events may not be evident as early in the process as they would be with ER negative tumors. That may be what we're observing.

My personal preference -- again, I'm allowed to speak but not vote, like at home.

(Laughter.)

DR. JOHNSON: I'm talking about my home home, you know, with my wife and daughter.

It seems to me that we ought to accept the overall result of this large, powerful trial. And then I like Dr. Margolin's and Dr. Lippman's suggestion that we put forward the data in the package insert which guides the clinician and the patient as to what benefit he or she may

obtain from this.

I can tell you from having seen these data -- and I, like Derek, was wondering if Grant had lost his mind because yesterday he obviously lost his mind and again today he's lost it.

(Laughter.)

DR. JOHNSON: But I'm also very persuaded by the data as were shown, and I'm not sure how now I'm going to handle the patient that I have at home with positive nodes who's ER/PR positive. Candidly I've been going to using the sequential therapy, and now that I see these data, I'm a bit hesitant about that. But nevertheless, I like that as an option and I think these data prove that. I suspect -- this is my prediction -- that we will see a difference as time goes on, but related to the biology of the tumor types rather than just some sort of specific interaction with Taxol per se.

DR. NERENSTONE: Ms. Fischler.

MS. ZOOK-FISCHLER: Well, as the patient rep, I have to take a patient's position, and I think that is that patients need the options. As I'm listening, while I do see -- and it jumped off the page at me as well -- that the ER/PR negative women had the greater advantage, I didn't see that the women who were estrogen positive were at a disadvantage. They just weren't as at great an advantage.

But as the patient, I would like to be able to sit down with my doctor and decide what's best for me. I also know from working with women in SHARE, the group that I'm affiliated with, I've seen many women who can't tolerate tamoxifen. So, for those women, it would be a very important option to be able to have Taxol. So, I would like to see it go ahead with Dr. Margolin's proviso.

DR. NERENSTONE: Dr. Kelsen.

DR. KELSEN: These data now have been available for some time I think to the breast cancer specialist community. How has it influenced your studies? Larry, you just showed us a whole series of trials that are underway. When patients enter those trials and they are ER positive or ER negative, are they being treated differently in the Taxol-containing studies?

DR. NORTON: No, absolutely not. That was a very major consideration in the design of all these trials, and everybody felt that this type of subset analysis was inappropriate for guiding future decisions, especially because we want more data. And if we make that decision, then it's a self-fulfilling prophecy. We won't have the data and we won't have that kind of information. So, that's why you'll notice that there's a taxane and in fact Taxol in all the current and future plans in the cooperative groups.

DR. KELSEN: So that we will never have a prospectively randomized study in which women who are ER or PR positive or both are randomly assigned to receive T or not to receive T after AC with tamoxifen.

DR. NORTON: That is not currently planned. This is what Dr. Temple said. I think that with all of these trials involving taxanes and with patients with ER positive disease being treated as well as ER negative and patients getting tamoxifen or not, we're going to have a huge data set that we could then do some very reasonable subset analyses of in this regard, and that that's going to really give us the power for making that determination long term rather than the randomization.

In terms of the randomization, since you bring it up, it's an ethical consideration. It's exactly what we decided. Let's say we decided not to give Taxol to ER positive patients. Let's say 5 years from now we find out that indeed the curves start to separate as we get past 3 and a half, 4 years and the tamoxifen effect wears off and the curves start to separate. We've cost a lot of women their lives by making that decision.

If we decide, however, to give Taxol and it turns out long term not to be effective, what have we really cost them? We've caused some toxicity, but compared to what they've received with the AC and compared to many

other things we do in oncology, it's really very minimal.

So, balanced with that minimal toxicity versus the potential for saving lives, the intergroup decided to include Taxol for everybody regardless of hormone receptor positivity.

DR. NERENSTONE: There is another study. The NSABP study is closed. It was not randomized I don't believe. I mean, ER/PR was not in the randomization. I do believe it was in the stratification, and that study is now closed to accrual but did randomize AC plus or minus Taxol to stage II patients. So, there will be another group of patients along.

Dr. Margolin?

DR. MARGOLIN: Just as a point of clarification mainly to Ms. Zook-Fischler, I think we recognize that the data for the small number of patients who were estrogen receptor positive but didn't end up on tamoxifen is no more convincing of a Taxol effect than the whole group at large. So, I don't think for the patient who can't take tamoxifen, we can say that Taxol supplants that and it replaces the effect of tamoxifen.

DR. NERENSTONE: Dr. Temple.

DR. TEMPLE: I just wanted to be sure I understood. There are going to be further data on Taxol, yes or no, in the receptor subtypes, although not most of

the ongoing trials because everybody is getting Taxol there. But there are at least a couple trials where one will be able to look at it.

If they're stratified, that's more than sufficient. You can't randomize to receptor status, but whether stratified or not, both statuses are sufficiently common in the population that you'll get effective randomization I think anyway.

DR. NERENSTONE: Dr. Lamborn.

DR. LAMBORN: Could I ask that we hear Dr. Berry's additional comments or clarification?

DR. NERENSTONE: Yes, thank you.

DR. BERRY: Thank you.

I completely agree with Dr. Temple concerning looking at subsets and the strength of the subset. If there is something that grabs you by the hair or knocks your socks off, I look at it and I believe it.

The question is, does this knock your socks off? And the appropriate analysis is exactly what Dr. Lamborn suggested, namely we do a Cox proportional hazards model, adjusting for all the other covariates, and we ask if there an interaction between the use of Taxol in estrogen receptor/PR status. The answer was for overall survival, there's no significant interaction. For disease-free survival, there is a .036 p value.

Now, in doing interim analyses, we adjust for multiple looks. In doing subset analyses, we adjust for multiple subsets. How many subsets did I look at? I looked at nodes. I looked at tamoxifen. I looked at menopausal status. I looked at tumor size. How many? I don't know. A half dozen, 10? Even if I looked at two subsets and adjust this p value accordingly, it is not statistically significant. This is not an effect that knocks your socks off.

Two final points. One is Dr. Lippman's question. The vagaries of subset analyses are identical whether it's a negative study or a positive study. The same problems arise.

Another point about sample size and confidence intervals. If you take a random subset of a set of patients and look at the size of the confidence interval, it has to increase. So, Dr. Norton's statement I would agree with.

DR. NERENSTONE: Thank you.

Other questions from the committee?

(No response.)

DR. NERENSTONE: If not, then I'd like to go to the questions from the FDA. I will skip all of the preamble -- it just goes over the discussion and the data that we've already seen -- and go right to the questions,

which is the last page of the handout.

Do the results of this trial provide highly reliable and statistically strong evidence of an important clinical benefit from Taxol in patients with node-positive breast cancer?

Discussion?

(No response.)

DR. NERENSTONE: Okay, then let's see a show of hands. All the people who say yes?

(A show of hands.)

DR. NERENSTONE: That's 8 yeses. That's everyone who is voting.

The second question. Do the results of this trial provide evidence of clinical benefit from Taxol in patients with node-positive, receptor-positive breast cancer who also receive tamoxifen as adjuvant therapy?

Comments please? Dr. Lamborn.

DR. LAMBORN: I guess I have some problem with the question as it's posed because if you just were to say look at this subset and look at the data, then you have one answer to the question. If you ask the question of you have overall results and you've now done a subset analysis, do you have convincing evidence that in fact the result is different for the receptor positive group, then I think that it becomes a slightly different issue. So, I don't

know if others see this as a -- I think it's really the latter question that we can address from this data.

DR. NERENSTONE: Dr. Williams?

DR. WILLIAMS: I would suggest that you address it any way you want to. It's the decision I think someone is faced with when they have a women in this situation, based on anything you think is appropriate, including the evidence from this trial, whichever evidence you want to consider and what you've seen presented.

DR. NERENSTONE: I'm not sure I see the difference between question number 2 and question number 3, the first part. They really feed into one another. Maybe we should go to question number 3 which is really the crux of the discussion, which is, for which population with node-positive breast cancer -- all patients, patients with receptor negative tumors, patients with receptor negative tumors plus others who cannot receive adjuvant tamoxifen -- should this indication be approved? In deciding this, issues include the toxicity of Taxol, the size and the medical plausibility of the subgroup, and the unplanned nature of the subset analysis.

Discussion? Dr. Raghavan.

DR. RAGHAVAN: Well, I started by taking the devil's advocate view partly because I believed it and partly because I was asking questions. I think the

discussion actually resolved my concerns pretty comfortably. I'm a long-term opponent of subset analyses, and I think that even though this is a bigger subset than average, whoever made the point that the damage we would do by withholding the drug with the knowledge base we have is more than the damage we would do by letting it through.

I'm totally sympathetic to the position of the FDA. I think it's their job to raise questions like this and it's our job to deliberate on the data that are presented, not to do it in a trivial way, but in fact go through it very carefully.

Some of the early discussion I thought did trivialize the question, and I think now the discussion has been of a nature that when we look back in 10 years, my hunch is that once again Dave Johnson is wrong and the curves won't diverge. And he can't vote. So, who cares?

(Laughter.)

DR. RAGHAVAN: But I think his point is correct, which is that until we have data, then we should be conservative in favor of the patient. Therefore, these latter questions probably become moot. What we do is we advise the FDA. They've heard the clear sense of equipoise, but the jury moving towards feeling that the data support an approval for node-positive disease with caveats in the package insert.

So, I make the comment because I was the person at this part of the discussion that raised the questions, and I just want to comment that I'm pretty comfortable that my questions have been resolved.

DR. NERENSTONE: Ms. Zook-Fischler?

MS. ZOOK-FISCHLER: Yes. The question asks for which group of people it should be approved, and if it's approved for all patients, that doesn't mean all patients need to take that treatment. But it does open up all the possibilities for the patient and her physician, and I think that's what's really important here.

DR. NERENSTONE: Dr. Margolin.

DR. MARGOLIN: Well, just really a reiteration of what I said earlier. This is a very tiny point, but I would not leave in the package insert or any sort of subcomment that patients with ER positive tumors who cannot receive adjuvant tamoxifen -- we still don't know which makes you achieve less benefit with Taxol, the fact that you are receptor positive or the fact that you were receptor positive and received tamoxifen.

DR. NERENSTONE: Other comments? Dr. Blayney.

DR. BLAYNEY: I view statistics as a way to scientifically approach biology and the biology of breast cancer in this particular discussion. ER positive breast cancer is a slowly growing tumor. We don't eradicate and

cure some of those patients and the time that that makes itself manifest is longer.

I'm new to the regulatory advice arena, but I agree with Dr. Raghavan that I think, as presented, the data is persuasive to me that we should advise them to approve this for node-positive breast cancer patients, but with the caveat that the data is what it is in 1999, and the second caveat that I made earlier, that in over 65-year-olds, the data is what it is, and that should also be considered by physicians advising their women patients.

DR. NERENSTONE: Dr. Lippman?

DR. LIPPMAN: Yes, I actually agree with Dr. Johnson on both points. In terms of biologic plausibility, there certainly is biologic plausibility that with time we might see an effect in ER positive patients because the effect that we see will take longer to manifest, if it's really slow growing, based on Dr. Norton's kinetic argument. But we don't know, but there's biologic plausibility there.

First of all, people will see this published, and putting this information in the package insert will lead to deliberations like Dr. Johnson just mentioned. People will interpret this and it will affect, I think, the types of patients possibly and when it's being used. I think the information will be there and will guide us, and

with time, we'll have more information.

DR. NERENSTONE: Other comments from the committee?

(No response.)

DR. NERENSTONE: What I'd like to do then is we'll take the first question as all patients, and if it passes, then obviously we don't have to do a subgroup. For the population with node-positive breast cancer, starting with all patients, should this indication be approved? All those who say yes?

(A show of hands.)

DR. NERENSTONE: 8.

The second question and I think the sense of the committee was that a package insert should reflect the relative data that was presented here. Does that need to be voted on, or you have the sentiment of the committee?

DR. WILLIAMS: Could I get some more detail on that? Let me give you an example. Aredia package insert was altered because of an apparent different size of effect in hormone treated breast cancer patients versus chemotherapy treated patients. That was put in the indications section, a statement referring them to the clinical trials section. I don't think a lot of people read the clinical trials section.

It will mean a lot to the company. I think

they will not want it in the indications section. Most companies do not want their indications section to be cluttered with a statement talking about something somewhat negative.

So, I would wonder where you thought this would be appropriate, what level of concern should it be brought to, and if there's a statement that were to be put in the indications section, you might have some discussion about what it would say.

DR. NERENSTONE: Comments? Dr. Margolin, that was initially your suggestion.

DR. MARGOLIN: I'm not sure that I really understand what Grant is saying vis-a-vis the way the question reads. I thought the question was whether we want --

DR. WILLIAMS: It's a new question.

DR. MARGOLIN: Oh.

DR. WILLIAMS: This has to do with what kind of statement you want in the package insert, whether you want something in the indications section referring people to the clinical trials section where some data may be, or whether you want them, if they have the concern, to go find the indications section and look for the data.

DR. NERENSTONE: Dr. Temple, would you like to comment?

DR. TEMPLE: Well, just to illustrate. It could say for the treatment of patients following other therapy with node-positive breast cancer. It could also say, see clinical trials section for discussion of unbelievable difference between two --

(Laughter.)

DR. TEMPLE: Or some variation of that.

DR. NERENSTONE: Maybe relative clinical.

DR. TEMPLE: Yes. So, you flag it and that gives you some hope that someone will read the section although, as Grant says, who knows?

DR. NERENSTONE: Dr. Lippman and then Dr. Kelsen.

DR. LIPPMAN: Again, just in terms of consistency and setting a new precedent, I think if we do that, that kind of comment could be made on almost every drug that's approved. We could refer them in this case to people with a lot of positive nodes. So, I guess the question is, since there are subsets in a lot of these, this could be something that is put in, this kind of thing in a lot of approvals, and do we want to go there?

DR. TEMPLE: Well, you sort of have to trust me on this, but you don't see things this striking all that often.

Now, one of the things about subset analyses is

nobody pays any attention to them at all unless they're plausible and striking. So, there's a sort of self-fulfilling prophecy here and you can be misled and that's why people worry about it. But it's unusual to see anything that interesting in a large fraction of the patients treated. That doesn't happen every day at least partly because we don't pay any attention to them even if they're sort or large unless they seem credible and involve a large fraction of the population.

So, I guess I would say you don't have to worry that we're going to throw these every time because we're highly resistant to that suggestion. It's more a question of whether this is different enough or striking enough to merit unusual treatment.

DR. LIPPMAN: But again, just in terms of clarification of what Dr. Berry just said, if these data were presented in a different way, adjusting for the number of subset analyses, I understand that they would not have been even statistically significant.

DR. TEMPLE: Yes. That's essentially always going to be true. If you have 10 subsets and Bonferronize, you'll never overcome that. So, you have to do more subjective things like think how plausible it is and think how many subsets there really were that were that interesting. It's a very hard problem. That's why we

usually reject them.

DR. NERENSTONE: Dr. Kelsen.

DR. KELSEN: I think we should put something in the package insert about this difference. I'm not sure where it would go yet, but I wonder how we'll handle it -- how you'll handle it I guess -- at 2 years or 3 years from now when one of these two things is going to be true. One, there is a late effect. ER patients do benefit, that warning or whatever you want to call it, caveat should be removed. Two, we're wrong. Even though the toxicities are relatively acceptable for an increase in cure rate, there is no difference and therefore the package insert should be changed. How will that be handled?

DR. TEMPLE: Well, if the data start to look really good for that subset, I think that's going to be not a problem because the company will take care of reminding us of those data. If it poops along and looks sort of the same, I guess we might even come back to you. If it now looks really overwhelming, maybe we've learned something true or maybe other available data will contribute to that. So, we'll arrange with the sponsor to provide the follow-up. I'm sure they will be glad to do that.

DR. KELSEN: If it was going to be done in that way, then I would probably stick in the indications, see the clinical trials section, since it seems to me that

we're so uncertain at this point, rather than put it in the indications section.

DR. JOHNSON: Presumably in that section, you would have the very analysis that has been shown to us with those differences.

DR. TEMPLE: In the clinical trials or the indication?

DR. JOHNSON: No, in the clinical trials section.

DR. TEMPLE: Yes, that's exactly right.

DR. WILLIAMS: The question is exactly what sort of statement would be in the indications section that would be pointing you to the clinical trials statement. What is the sense of the committee? Should it be there's little data or preliminary data show, et cetera?

DR. JOHNSON: No. What I would do is based on what the trial was designed to do. I would say it's indicated for node-positive breast cancer. Then I would put, parentheses, see clinical trial data.

DR. WILLIAMS: Okay. So, from what I've heard from two so far is that you would not make a special statement in the indications section that would try to describe the sense of what's going to be in the clinical trials section.

DR. JOHNSON: We've had this same conversation

about toxicity issues in the past where we've allowed the sponsor or FDA has required that certain data be placed in there, and we've simply directed the physician to that area.

DR. WILLIAMS: The difference here is that oftentimes we will direct people to another section, but it will be in such a context, they'll know why they're looking. We might say, especially look because of the ER positive findings. Then they would know to look to the section.

Another that sounded like what you were saying is approve it and go look in the clinical trials section. Is that what you're saying?

DR. JOHNSON: Well, again, I think that what the study did was looked at node-positive patients. So, again, I would say it's approved for node-positive --

DR. WILLIAMS: The indication would be node-positive patients. That's no question. The next sentence might be to guide them to the clinical trials section for a particular purpose. The purpose of putting it in the indications section is to make it prominent.

DR. JOHNSON: No, I understand that, but it also suggests that the comment that you would like to put there would be, and especially pay attention to the ER positive/PR positive tamoxifen treated. And I wouldn't say

that. I personally would just say see the clinical data.

By the way, I'm stunned -- stunned -- actually that you think we don't read these package inserts.

(Laughter.)

DR. JOHNSON: And I want you to know, Bob, I personally trust you.

(Laughter.)

DR. NERENSTONE: Go ahead, Dr. Temple.

DR. TEMPLE: Well, I'd just like to hear a little more from everybody. There's a huge range of things one could say, but I think the assumption based on what you just said is it will say for node-positive patients. You can then say, see clinical trials. My bias is you tell people to do that and you don't tell them why, they don't pay much attention to you. So, one could say, see clinical trials and mention an unplanned subset analysis that suggested a possible difference based on receptor status. That's not as extreme as saying, don't use it, but it does point out what the area of problem might be, and then they'll go see it. So, unless you didn't think that was a good idea, that's probably what we would plan to do.

DR. NERENSTONE: Dr. Margolin.

DR. MARGOLIN: I was just going to suggest some wording to the effect of near the indications say, see clinical trials for important information about receptor

positive patients, and then in the clinical trials section, just before you show the graphs and the tables, just a statement that not that it doesn't work, not that we're waiting, but just say Taxol has not been proven to benefit patients with ER positive tumors who are receiving tamoxifen or just with ER positive tumors in overall survival and that the benefit in disease-free survival --

DR. TEMPLE: That's actually a relatively strong statement. I think other sense I get is that most people wouldn't want anything quite that strong, but those are the nuances. I think we have a pretty good sense of what people want.

DR. BLAYNEY: How easy is it to change the package insert in these various sections? I know Dr. Johnson would jump right on it when you did change it.

(Laughter.)

DR. BLAYNEY: So, how easy is it to change these inserts in the indications and clinical trial information?

DR. TEMPLE: Probably you've got to ask the companies that too. We think it's not very hard if you've got data that support it.

DR. BLAYNEY: In 3 years, for instance, if an analysis is published suggesting that there is benefit in ER positive patients, is that an easy thing for you all to

put into the clinical trials section?

DR. TEMPLE: If it's convincing, it's very easy. It could be changed in a very short order. We're familiar with the study. It's just the same analyses that have been done, extended by a little bit. It's a very easy change to make if the data are there.

DR. BLAYNEY: In that case, I would advocate putting a statement in the indications and including in that indication the phrase "unplanned subset analysis." I think that's fair warning and a fair statement of the data upon which we advised you today.

DR. NERENSTONE: Dr. Lippman?

DR. LIPPMAN: I guess I don't fully agree. I think the issue of unplanned, planned, secondary subset analyses -- we think a lot about that. I think one only needs to think about selenium and olaxafene and other issues to understand how that is accepted and understood elsewhere. If I were to say anything, I would say, see clinical trials section for detailed analyses and subset analyses. End of sentence without pulling anything out.

DR. NERENSTONE: Dr. Margolin.

DR. MARGOLIN: I strongly agree with that. I think the word "unplanned" is sort of meaningless. It's the numbers and the fact that it wasn't prestratified and things like that and not the fact that you didn't plan to

do it but now you did it. That's really irrelevant. It's a misleading word I think.

DR. NERENSTONE: Other comments? Dr. Blayney.

DR. BLAYNEY: I meant to convey the fact that a subset analysis is recognized not to be statistically rigorous. So, however you would want to flag that for people I think could be useful for practicing physicians.

DR. TEMPLE: I think what we'd probably try to do is mention in the indications section what area of the clinical trials is of interest, that is, it refers to receptor status, and then in the clinical trials section, one would discuss the nature of the analysis and all that stuff. If we put too much into the indications section, we're sort of taking away the indication, which is what a number of people have said you don't really want to do. So, we want to introduce a note of caution and get people to read that section, but we don't want to deny the indication because that was your recommendation.

DR. NERENSTONE: Dr. Lamborn.

DR. LAMBORN: I think that you've sort of hit on exactly what I get the sense is that we have here, which is something in the indication and something that might point them to where the area is that they would want to look for further information but not something that took the indication away.

DR. NERENSTONE: If everybody will turn to the last question, for the patient group designated by ODAC in question number 3 -- and that is all patients, which is what we voted on -- should Taxol be approved for use subsequent to standard combination chemotherapy or only for use after treatment with doxorubicin and cyclophosphamide, the chemotherapy used in the trial?

Comments? Dr. Kelsen.

DR. KELSEN: It would seem to me that if we did that, you'd sort of be saying that the standard of care for node-positive women is only AC, or at least you might be implying that the standard of care for node-positive women, as far as the non-Taxol part of treatment, is only AC and that no other regimen might be acceptable. If you approved it only for use with AC and with no other treatment, would that not be implying that that was the only acceptable standard of care with Taxol? That's a question.

DR. NERENSTONE: Dr. Margolin.

DR. MARGOLIN: Well, I think this is probably one of the toughest questions because we don't have the numbers to look at anything else, but you also don't want to be so rigid as to say that, even though this was a study that was done, this is the only setting in which it might work. I think the most important thing is the question of whether the interaction with Adriamycin is the compelling

thing and you don't want people to be using oddball regimens like melphalan-based regimens. So, perhaps a compromise to the effect of Adriamycin-based adjuvant therapy which you know 99 percent of regimens are going to include Adriamycin, Cytoxan with or without something else.

DR. NERENSTONE: Dr. Raghavan?

DR. RAGHAVAN: Yes, I agree with that. The one caveat is that I think we've spent the morning talking about data and what's presented, and we haven't heard anything about Taxol following anything else. So, I'm comfortable with what Kim said, which is Adriamycin-based regimens, but I don't know from anything I've heard in the last 4 hours what Taxol does after CMF. I know there are data that relate to that. They just haven't been presented. So, I think we should work within the confines of what the discussion was. If the company had wanted a broader indication, they might have presented data that related to it. So, I think flushed with enthusiasm for having done good work, we want to still remain within the bounds of sanity.

DR. NERENSTONE: Other discussion? Dr. Kelsen.

DR. KELSEN: I'm very comfortable with the doxorubicin-containing regimen.

DR. NERENSTONE: Dr. Lippman.

DR. LIPPMAN: Yes, I am too for the reason of

sort of biologic plausibility since it wasn't looked at here, but it certainly is consistent with the mechanisms.

DR. NERENSTONE: Just one comment. I'm also concerned about additive toxicities, certainly with CMF, you could have prolonged neutropenia, and how many doses of CMF? Would you get six? Would you get four? And the added toxicity of Taxol after 4 to 6 cycles of daily Cytoxan for 14 days I think your toxicity profile could well be quite different, and we don't have the data here to do that.

My question, though, is what about the dose of Adriamycin. Do we make a comment about that as well, or is that not necessary?

Would the FDA like to address that?

DR. WILLIAMS: I'd sort of like Dr. Temple's opinion on that. The study has three doses of doxorubicin, but of course this isn't the doxorubicin labeling. The study basically found no difference in effect with -- the lowest dose seemed to be acceptable. How should this label or especially dosage administration --

DR. TEMPLE: That's difficult, and Bob and I were just talking about this. The labeling for cytotoxic adjuvant therapy is grossly deficient. We just approved epirubicin, so we finally have one thing that's covered. None of the others are. So, the solution is not so easy.

I think what we usually do in that case is describe what was done, which takes care of the immediate problem. How to get the new doxorubicin finding into labeling is hard, given that it's not labeled for that use. I think we need to try to think about how to do it, and I don't know the answer yet.

DR. WILLIAMS: But your answer is that we shouldn't necessarily address it in this label in terms of indications section.

DR. TEMPLE: That would be most odd to basically label another drug, and it's not really the Taxol part of the study. But I'd be interested in hearing what people say. It certainly ought to get into the label somewhere.

DR. NERENSTONE: Dr. Margolin.

DR. MARGOLIN: Most people who are aware of these data are aware of the doxorubicin data from this trial and what the NSABP has done over and over again. I think if you just simply use the word "standard" doxorubicin-based chemotherapy, most people think standard and think 60 times 4, and you're going to have very little variation from that.

DR. NERENSTONE: Yes, Dr. Williams.

DR. WILLIAMS: As you know, I think just two days ago we approved epirubicin which is an anthracycline.

So, the question is, would you feel comfortable broadening this to anthracycline?

DR. NERENSTONE: Discussion from the committee? Dr. Johnson.

DR. JOHNSON: I'll talk about that in just a second, but we did hear yesterday in a survey, when we were talking about another product, that I think the figure was 86 percent of women currently receiving adjuvant treatment are getting a doxorubicin-based regimen. So, even if we summarily exclude CMF, it's not a high percentage of patients. Those were the data we saw yesterday.

But I have another concern. I actually think Kim's suggestion is the right one, with this minor concern, and that is 4 cycles of AC or 6 cycles of FAC or classic CAF? Again, there the issue about other toxicities, including cardiac toxicities, is another issue. I mean, it comes up. I think it's likely to be a relatively minor issue, but I don't know that we know that either. It goes back to what do we know, the data we have, and whether or not one should be willing to do this.

Again, my personal bias -- and we've repeatedly had these discussions around this table -- is that I believe we should leave flexibility for the physician treating the patient and the patient to make a decision, as long as we can provide appropriate guidelines and caveats.

In this case, if one were to use the language that Kim used, perhaps it might be appropriate to say standard therapy and then say the study was done with four cycles of AC, and then leave it to the treating physicians to interpret that data in an appropriate manner.

Oh, epirubicin. Personally again I would go back to the language that Kim used, doxorubicin-containing therapy, not to suggest that you shouldn't use epirubicin, but the study was done with doxorubicin therapy.

DR. NERENSTONE: Other comments? Dr. Lippman.

DR. LIPPMAN: I'd just like to clarify Dave's point. So, in the indication, you would put standard therapy. You wouldn't specify doxorubicin-containing, but you would put in parentheses the study was done with --

DR. JOHNSON: No. I would use the term "standard doxorubicin-containing adjuvant chemotherapy," but I would make it clear in the data set that it was 4 cycles of AC.

I do think that a lot of physicians use AC, but candidly, at least where I practice, in the region in which I practice, 4 cycles of AC is not what most of the physicians use. It may be what they ought to use, but that's not what most of the physicians use.

DR. NERENSTONE: Dr. Lippman.

DR. LIPPMAN: Well, if you're going to put sort

of in parentheses in the indication what the study used, would you want to, since we're basing this on sort of biologic plausibility of mechanism -- that's the doxorubicin-based therapy. Would you want to broaden it to anthracycline-based therapy? The study used 4 cycles of AC.

DR. JOHNSON: I'm less comfortable doing that personally. Again, if the committee and the FDA decides to do it, I'm fine with it, but again, I'd like to try to stick with the data at hand.

DR. NERENSTONE: Yes.

DR. JUSTICE: I think the number of cycles issue would be something we would address in the clinical study section normally, and we're already referring to it. So, I think we can cover it there.

DR. NERENSTONE: Other discussion?

(No response.)

DR. NERENSTONE: Do you need a vote on this, or do you have a sense of the committee?

DR. WILLIAMS: I think we have a sense. I'm not sure what we're going to do.

DR. NERENSTONE: Fair enough.

Well, thank you, everybody, for sitting through this. We'll adjourn now and reconvene at 1 o'clock. Thank you.

(Whereupon, at 11:52 a.m., the committee recessed, to reconvene at 1:00 p.m., this same day.)

AFTERNOON SESSION

(1:03 p.m.)

DR. SCHILSKY: My thanks to Dr. Nerenstone for standing in for me this morning.

We'd like to begin again with introduction of the committee members since we do have different people at the table at different sessions. So, Dr. Nerenstone?

DR. NERENSTONE: Stacy Nerenstone, medical oncology, Hartford, Connecticut.

DR. JOHNSON: I'm David Johnson, medical oncology at Vanderbilt University.

MR. McDONOUGH: Kenneth McDonough, Patient Representative, Pittsburgh, PA.

DR. PELUSI: Jody Pelusi, oncology nurse practitioner, Phoenix, Arizona, and consumer rep.

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

DR. BLAYNEY: Doug Blayney, medical oncologist, Pomona, California.

DR. SCHILSKY: Richard Schilsky, medical oncologist, University of Chicago.

DR. TEMPLETON-SOMERS: Karen Somers, Executive Secretary to the committee, FDA.

DR. LIPPMAN: Scott Lippman, medical oncologist, University of Texas, M.D. Anderson Cancer Center.

DR. LACHENBRUCH: Peter Lachenbruch, FDA, statistician.

DR. CARDINALI: Massimo Cardinali, FDA.

DR. KEEGAN: Patricia Keegan, Division of Clinical Trials, CBER.

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

DR. SCHILSKY: Thank you.

Karen has a conflict of interest statement.

DR. TEMPLETON-SOMERS: 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 for the meeting and all financial interests reported by the committee participants, it has been determined that all interests in firms regulated by the Center for Drug Evaluation and Research present no potential for an appearance of a conflict of interest at this meeting with the following exceptions.

Dr. Kim Margolin is excluded from participating in today's discussion and vote concerning Roferon.

In addition, in accordance with 18 U.S.C. 208(b)(3), a full waiver has been granted to Dr. Scott

Lippman which permits him to participate in all official matters concerning Roferon.

A copy of the waiver statements may be obtained by submitting a written request to the agency's Freedom of Information Office, room 12A-30 of the Parklawn Building.

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 FDA's invited guest, there are reported involvements which we believe should be made public to allow the participants to objectively evaluate his comments. Dr. John Kirkwood would like to disclose that he has an interest in Schering-Plough's interferon alpha 2b. He also has received grants, consulting fees, and speaking fees from Schering and speaking fees from Roche.

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

I'd also like to announce that Dr. Janice Dutcher was unable to attend due to weather problems and

that Dr. Scott Lippman has stalwartly agreed to take over the role of discussant.

Thank you.

DR. SCHILSKY: Thank you, Karen.

There's no one listed on the agenda as having requested to speak at the open public hearing, but is there anyone in the room who wishes to make a statement to the committee?

(No response.)

DR. SCHILSKY: If not, we'll move right on with the remainder of the agenda.

As Karen mentioned, the FDA has invited Dr. John Kirkwood from the University of Pittsburgh to make a presentation to the committee to help provide us some context in which to consider the sponsor's application today. Dr. Kirkwood?

DR. KIRKWOOD: Dr. Schilsky, Dr. Keegan, I'm delighted to have the opportunity to review with you the updated information on E1690, the intergroup trial of high dose and low dose interferon in high risk melanoma patients.

This trial was commenced based upon background data that I think everyone is well aware of, objective responses in approximately 16 percent of patients in large collected series treated with all varieties of interferon

alpha 2, durable responses in about 5 percent of these patients, which are very comparable to what we know from interleukin-2, subsequently approved for the therapy of metastatic melanoma.

A variety of antitumor effects in vitro and immunomodulatory effects, including up-regulation of MHC class 1 and class antigens, have been the focus of a variety of studies that I won't have time to talk about today.

The trial 1684, which was the pivotal basis for the approval of interferon alpha 2b at high dosage for high risk melanoma patients, included 287 patients, half randomized to high dose interferon for a year, the other half observed. As you all know, this showed very significant relapse-free survival improvements to a p value of .004, overall survival impact to a significance of .04, and a quality of life improvement, as well as cost efficacy, which is comparable to accepted therapies of adjuvant therapies of other solid tumor chemotherapies.

The trial data that I think you're all well aware of showed an impact which included durable response and now out to 10 years, no significant difference with the data that was published at 7 years, as you see reported here for the alpha 2b high dose trial 1684; survival impact which was also significant and which is also now updated to

10 years without change in this pattern.

The trial 1690 that I'll talk about today was designed in 1990 when the relapse-free survival benefit of 1684 was recognized, but certainly no survival impact had yet been observed. It was conducted between February of 1991 and June of 1995, and an important element that I didn't put in the chronology here is that in July of 1995 this committee considered the application for alpha 2b and approved it for adjuvant therapy of high risk patients with melanoma using the high dose regimen that we had developed in E1684.

In May of 1998, some two to three years before what we had anticipated would be the closure of 1690 at the scheduled number of 200 deaths or relapses, the data safety monitoring committee decided to unblind this trial because of the slowing number of events, the basis for this, the improved prognosis that I'll come back to discuss in the E1690 experience.

And over the summer of 1998, there were both external and internal audits of the data which corroborated all of the database that we had in ECOG.

In the fall of 1998, a statistical analysis was presented to the FDA and to CTEP on October 13th, and in November, this was placed on the web and summarized as an abstract presented at the European Society for Medical

Oncology.

Between March and April of 1999, data on salvage therapies, which I will review with you today, were collected, and this was all presented briefly to ASCO in May of 1999.

The trial 1690 included 642 patients, a third randomized to high dose interferon given for 1 year, a third to low dose for 2 years, and a third to observation.

The trial had one important difference in the eligibility in that patients who had primary cutaneous melanomas greater than 4 millimeters of Breslow depth were allowed with or without regional lymph node dissection, a key distinction from the E1684 trial such that 80 percent of the patients who entered this trial had clinically node-negative but not pathologically established node-negative disease. We included about 10 to 20 percent of patients who had regional lymph node metastases presenting as primary disease in the regional lymph nodes, but half of patients presented and entered this trial with recurrent lymph node metastatic disease.

The trial analysis that I'll report to you today included 642 patients in the intention-to-treat analysis, all patients who entered the trial. 34 cases were ineligible, and so all of the demographic analyses will focus upon the 95 percent of patients in this trial,

608, who met eligibility requirements.

The goals of this study were an endpoint first which was used for all monitoring committee decisions and for the decision to unblind, which was relapse-free survival; a second primary goal, overall survival analysis. And the design was to pick up 83 percent power for a 10 percent increase in cure or a 50 percent increase in either the median relapse-free or overall survival. And two two-sided log rank tests were specified for analysis.

I will also report to you Cox analyses, adjusting for all the prognostic variables that we recognized, and a comparison to the E1684 data as well as an analysis of the salvage therapies that have now been gone through in detail for 93 percent of the patients on the trial.

The demography of the patients entering this trial included 25 percent of patients who were node-negative, N0; 34 percent who had 1 node involved; 21 percent who had 2 or 3 nodes involved; and 20 percent who had 4 or more nodes involved. This contrasts with the E1684 trial which had only 11 percent of patients with T4 node-negative disease.

The analysis of the outcomes for relapse-free survival show a hazard ratio for prolongation of time to relapse or improvement in the fraction of relapse, 1.28,

with all of the 95 percent confidence intervals above 1, a p value of .05.

The low dose interferon impact was 1.09 hazard ratio, crossing the value of 1, with a p value of .17.

The surprise in this trial was that survival was not impacted at all on either of the therapeutic arms, and we'll come back to discuss that later.

The plots for the relapse-free survival illustrated with high dose interferon in all of these as yellow, low dose interferon as red, and observation as blue, revealed the data that's consistent with the hazard ratios I presented before, survival plots overlapping in all three of the arms.

Hazard function analysis shows, similar to the E1684 trial, an early impact of the high dose interferon illustrated in yellow here. The relapse risk of patients who were observed, somewhat less than we had seen in the E1684 trial, and the values for the hazard functions for the low dose interferon arm intermediate between the high dose and the observation plots.

Subset analyses, although I know these are somewhat fraught with problems, show a consistency of impact across all of the stratification groups that we analyzed both by stage of disease and by nodal category, the exception for this being the 1-node-positive group for

which the hazard ratio was 1.0. As you see, the node-negative population, hazard ratio 1.46, the node-positive populations also about equivalent, but this one group of single node-positive patients, clearly the outlier in the subset analyses.

I should back up to say that the one group that by itself achieved nominal significance was this one group of 2 to 3 node-positive patients, and for this group, the hazard ratio of 1.92 associated with the curves that I have on the next slide for this group achieving significance, as is shown here, in the subset alone.

The toxicity of interferon alpha 2b given at high dosage in this trial was about equivalent to what we saw in the E1684, the single exception being that we had no toxic deaths on the high dose interferon arm. In fact, the only two toxic deaths were observed both on the low dose interferon arm, one of a cerebrovascular accident, one of the myocardial infarction.

The toxicity required dose reduction during the induction first month of therapy in 44 percent of patients for toxicity reasons, not relapse in this particular case. Maintenance arm treatment associated with a requirement for dose delay or dose reduction in half of patients over the subsequent 11 months. And again a similar fraction to the earlier trial, 75 percent of patients were able to stay on

treatment throughout the period of a year of treatment.

The average daily dose delivered in the 1690 trial was above that which was delivered in 1684, in the induction phase, 18.5 million units per meter squared as the median dose; 8.2 as opposed to 8.1 during the maintenance phase.

Comparing the absolute and relative impact of 1684 and 1690, we have here the impact in terms of relapse-free survival for the high dose interferon arm. 37 percent over 26 percent continuously free of disease at 5 years in the E1684 trial; 44 percent as opposed to 35 percent in the 1690 trial. This increment in terms of absolute percentage points is 11 percent in the 1684 trial, 9 percent in the 1690 trial; the relative increment 42 percent in the 1684 trial, 25 percent in 1690. As we've earlier mentioned, there is no difference in the overall survivals at 5 years, as is shown here.

The conclusions we drew then at the first analysis of this were that the high dose interferon arm improves relapse-free survival with a hazard ratio of 1.28, a continuous relapse-free survival of 9 percent improved at 5 years, log rank p of .05, a Cox analysis, .03, as I'll show you in a minute, and is consistent with the 1684 trial.

Secondly, the subset data, which in the 1684

trial had showed no benefit for the node-negative population, were here refuted and the node-positive and node-negative populations behaved very, very consistently in this trial so that there seems to be a consistent effect across the risk groups that we studied.

Low dose interferon had a lower absolute reduction in relapse rate, a hazard ratio of 1.09, a log rank of .16, and a nonsignificant value by Cox analysis, and that none of the treatments tested in this trial had altered survival at 5 years, for which we will review some other analyses now.

The questions that we developed then were whether patient populations differed between the two studies or whether the treatment results differed between the studies. The conclusions we'll draw from data that I'll now show you are that there are major differences between these populations in terms of the observation arm outcomes, that the observation arm outcomes differ by .01 significance for relapse-free survival and .001 for overall survival, and that there is no study effect. There is no difference between the impact of high dose interferon in 1684 as it is compared to 1690 between the trials.

The Cox model analyses, adjusting for treatment, showed a significant study effect, as I mentioned already, .01 for relapse-free, .001 for overall

survival. The Cox model treatment by study analyses demonstrated consistency with the interaction term .55 uncorrected to .90 as it was corrected between the 1684 and the 1690 studies, saying that there was not a difference between the impact of interferon in 1684 and 1690.

Adjusting for staging and nodal stratification variables in 1690, the high dose treatment effect was significant in Cox model analysis to a p value of .03.

The differences in the aggregate populations studied in 1690, the solid line, and 1684, the dotted line, here are shown for relapse-free survival. So, this is all patients entered into the whole 1684 study here and all patients in the 1690 study here, and you see that this is the basis for the significance of .01 for the improvement in relapse-free survival between the studies.

Even greater is the difference between the overall survival of the 1690 population in solid white here and the 1684 population in the dotted white here, significant to a value of .001.

The largest discrepancy was already identified in the single node-positive population. Here you see the observation arm with 1 node positive, untreated in 1690, and the observation arm in 1684 compared where the value is almost the same even though it's a much smaller subset between the two studies. So, a radical difference in the

survivorship and the relapse-free interval for these populations.

Comparing the 1690 to the 1684 studies, within study arms, the hazard ratios that we can show suggest that consistent improvement in the relapse-free survival, 1.21 times better for the high dose interferon arm of 1690 compared to the high dose interferon arm of 1684; overall survival consistently better, 1.23, the hazard ratio for 1690 high dose interferon compared to high dose interferon 1684. But the observation arm compared within these two studies shows an improvement which is greater than that for the treated arm, and the greatest improvement of all is the 1.64 hazard ratio for the untreated arms of the two trials compared in terms of overall survival.

Looking at the stratification groups that we had entered patients into these trials and comparing again the two studies by subsets, we see that all of the subsets analyzed in 1690, whether by nodes positive on this plot or by the stage groupings that were used on the top plot, show a consistent of the 1690 or consistent improvement of the outcome for the high dose interferon in 1690 as opposed to 1684. The one discrepancy here, the single node-positive group that we've already talked about.

For the observation group comparing the two trials in subset analysis, we see that the one group that

does not show an improvement in the outcome for the 1690 trial is the node-negative group, and this group, you will recall, is the group that we entered into 1690 without node dissection so that we know this group is heterogeneous and contains perhaps 20 or more percent who had nodes involved. So, this is the explanation for the hazard decrement in that group.

Comparing graphically the outcome of 1684 on top and 1690 on the bottom, observation groups in blue and treatment groups in yellow, you see that the lighter bar is the relapse-free interval where we have an improvement in the relapse-free interval in 1684, which is about equivalent or even better in the 1690 trial. We have a post-relapse survival which is about 2 years in each of these after relapse for all groups, save for the observation group of 1690.

Displayed in a table, the numbers are 2.1 years, 1.8 years, 2.6 years for the post-relapse survival of the treated and the observation groups, except for this observation group of the 1690 trial where this is 4.34 years survival post relapse and an overall survival from time of entry to trial of nearly 6 years, really unheard of in trials that we've done beforehand.

So, how could this have occurred? The questions were, did this arise from entry demographic

changes between the two studies; stage migration, Will Rogers phenomenon; or changes in definitive surgery; or perhaps in post-relapse salvage therapies that were used for these patients?

The demographics of patients between 1684 and 1690 is here portrayed. The node-positive population in 1684 was 89 percent of patients who entered this trial. It was only 75 percent of the 1690 trial. The recurrent disease population was 65 percent of the 1684 trial, but it was only half of the 1690 trial.

Conversely, the T4 population, the most favorable subset of entry stratification groups, was 11 percent of the 1684 trial and 25 percent of the 1690 trial. Of this population, 80 percent were not dissected as they came into the 1690 trial, offering the frequent opportunity for surgical salvage and entry to treatment, as you recall, with July 1995 approval of interferon, through the back door off protocol with the very same agent that we were testing in the original trial.

In summary, of relapse sites of disease of the patients on all arms, there was no difference in the distribution of relapses between high dose arm, low dose arm, and observation. That is to say, the impact we saw was generalized across all groups in the trial. There was a significant fraction of regional, nonvisceral relapses

for which surgical salvage, as I've already mentioned, was a possibility and subsequent off-protocol therapy was feasible.

This is a graphical display of the regional, surgically salvageable relapses in 1690 arm A, high dose; arm b, low dose; and arm c, observation. You see here the 26 relapses, here the 37, 38 relapses that had the opportunity for subsequent surgical salvage and subsequent systemic treatment by a variety of routes.

So, we went back between February and April of 1999, analyzed those of the 642 patients in the trial for whom we could get data. Relapses constituted 357 patients at that time. 331, or 93 percent, of the data were obtained on these subsequent data sweeps: 228 by on-site audits, 103 by queries of institutions where 1 or less patients had been accrued to the trial. Only 26 patients had missing data, only 5 from the observation arm.

These are the systemic biological salvage therapies or biochemotherapy salvage therapies used for all patients in the high dose arm and all patients in the observation arm displayed. And I will go through these in detail, so I won't dwell longer upon this table, given the short time.

Interleukin-2 was approved in the interim period while this trial was unfolding. We surmised that

this might have been one of the therapies that would have accounted for the differences in outcome. Of the 114 failures from high dose interferon, only 13 received interleukin-2. Of the 121 from observation, 22 received interleukin-2. This difference is not a difference. It doesn't achieve significance, and we looked at the impact of this therapy and it also did not make a difference in terms of the outcome of these patients for their post-relapse survival.

Biochemotherapy was also in increasing favor. Biochemotherapy was given to only 7 of 114 high dose failures, where it was given to 20 of the 121 observation failures. This difference is a difference, but it didn't, in terms of post-relapse survival, have any further connotations. There were not longer survivals amongst the recipients of biochemotherapy than those who did not receive this, as I can show you later.

The interferon salvage of the patients who failed high dose interferon was 17 of 114. The numbers in parentheses here are just the high dose recipients. This contrasted against 37 of 121 patients who failed observation and this difference was the most significant that we observed to a p value of .004. The impact of the interferon treatment of these patients illustrated graphically was a 2.2 year post-relapse survival of the

treated patients as opposed to a .8 year median survival for the patients who were not treated.

We wondered what this had to do with the surgical salvage of regional disease. How did this differ between regional and systemic disease? So, the next plot shows you in the solid lines regional disease failures who received interferon as opposed to those who did not in the solid blue and solid yellow lines, systemic relapses who received interferon in the dotted yellow as opposed to systemic relapses who did not receive interferon in the blue. And you see that the impact was greater for those patients who had regional, salvageable, operable disease.

We wondered whether this was just a surrogate for treatability, the patients who looked better got treated and therefore did better. This is a plot of those who received chemotherapy or other forms of non-interferon-containing therapy illustrated here, as contrasted to the interferon, and there was a difference here as well.

So, the conclusions that I draw are that if we look at trials that have demonstrated relapse-free survival and overall survival impact, 1684 is what we have. If we look for continuous relapse-free survival impact, we have 1684 and we have 1690. I've not had time to date to talk much about the NCCTG 83-7052 trial that was reported in the same year as the 1684 trial, but in fact, for the subset of

node-positive patients, high risk patients showed exactly the same trend.

Pending we have a series of studies, the 1694 trial of ganglioside GM2 versus interferon, which will be completed within the next 2 weeks with 851 patients; the Sunbelt trial, a 3,000-patient trial, which is currently ongoing and about half done; and the EORTC 18952 trial which is being conducted in Europe testing two intermediate dosages. So, this data is coming in from a variety of new vantage points.

Of the data that is completed and in hand, we have the 1684 trial, the NCCTG trial that I mentioned already with 262 patients, 162 who had nodal involvement and who comprised the basis for this Cox analysis positive for the impact in that trial of 3 months of therapy, and the 1690 trial that I mentioned already in detail today.

These are the trials that are pending, and I don't need to spend longer on this since we're short on time.

But I think the conclusions that I draw or the implications that I draw from this are that we have established the adjuvant role of high dose interferon alpha 2b, and it is consistent with the findings that we have in 1690. We have salvage data for melanoma recurrences that I wouldn't have predicted and I don't think anybody else on

our committee would have predicted but are interesting and that suggest that for resectable nonvisceral as well as visceral disease there is an impact that I think we hadn't before anticipated.

The endpoints for future trials, I think a key point of consideration for this committee, because I think we have to worry from now on that any trial that focuses upon overall survival will have to deal with salvage of patients that is hard to constrain for trials conducted in the era when you have alternative therapies.

And we really need prognostic and response indicators that are much shorter time lines to data than any of the clinical endpoints that we talked about.

It's 1:30, Rich.

DR. SCHILSKY: John, thank you very much.

We'll take a few questions from the committee if there are any information items you want clarification on. Dr. Blayney.

DR. BLAYNEY: The 1690 trial included an observation arm. Is this an ethical thing to do given the results of the 1684 trial, or what figured into your deliberations?

DR. KIRKWOOD: Good question. 1690 was started before any survival impact was apparent, as I've shown in the chronology of time line. At the time that we first had

statistically significant survival and relapse interval data from 1684, we had already completed all accrual and all follow-up on all patients in 1690.

DR. BLAYNEY: How did you handle patients who had sentinel lymph node dissection in the 1690 trial?

DR. KIRKWOOD: As Rich said I was going to get my legs cut off if I didn't stop at 1:30, I took those slides out. Those analyses were all conducted. I actually expected we would see a significantly larger fraction of patients with sentinel node mapping done as a basis of entry to this trial. In fact, it turns out that less than 5 percent of the patients who were node-negative had any sentinel procedure done and less than 5 percent of patients in any of the other groups of 1 node, 2 to 3 node, or 4 or more nodes positivity had sentinel node procedure. So, it was a very small component of the surgical practice in this trial probably because it happened just before the wave of this hit the surface.

DR. SCHILSKY: John, let me just ask you two things. In the 1690 trial, what was the dose of the low dose interferon?

DR. KIRKWOOD: It was the exact same dose that you'll hear further about today given for 2 years. We actually deliberated, when we designed 1690, whether we should give 3 million units 3 times a day forever, and I

was the lone vote on our committee to actually push for that. We actually stopped at 2 years because people thought it was impossible to carry patients past 2 years of this therapy without knowledge about outcome.

DR. SCHILSKY: Just to be clear, the low dose interferon in the 1690 trial didn't demonstrate any benefit with respect to either disease-free or overall survival?

DR. KIRKWOOD: As I showed in the hazard ratio analysis and as we have in subset analyses that I didn't have time to present, it did show an impact and it showed an impact which was intermediate on average between the high dose and the observation.

DR. SCHILSKY: That was statistically significant?

DR. KIRKWOOD: It was not statistically significant in overview. The p value was .16.

DR. SCHILSKY: Thanks.

Any other questions for Dr. Kirkwood? Dr. Lippman.

DR. LIPPMAN: I just want to clarify. You went through the data pretty quickly because of time. I understand that. But just to clarify this good survival on the observation arm in 1690, the biggest difference between the salvage therapies involved the interferon.

DR. KIRKWOOD: True.

DR. LIPPMAN: Do you think that that was in part the explanation for the better survival on the observation arm?

DR. KIRKWOOD: There's a component that may have been played by biochemotherapy, but I think the interferon salvage is the only explanation we presently have for that greater survival of the patients in the observation arm.

DR. SCHILSKY: Dr. Simon.

DR. SIMON: Is there any documented randomized trial evidence for the use of effectiveness of interferon in recurrent patients commensurate with what you're claiming from this sort of nonrandomized comparison?

DR. KIRKWOOD: We have done a number of those trials and we've done them in small enough series that I think none of them has had the power required to detect this kind of an impact that we're seeing here. I think that there's not adequate data.

DR. SIMON: Well, what was the size of the trials you did?

DR. KIRKWOOD: 20, 30 patients. They were phase I/phase II trials.

DR. SIMON: They were randomized trials?

DR. KIRKWOOD: No. These are phase I/phase II trials.

DR. SIMON: So, there have been no --

DR. KIRKWOOD: There have been no randomized trials that I'm aware of that have tested the impact of this --

DR. SIMON: So, there's really no randomized documentation --

DR. KIRKWOOD: Right.

DR. SIMON: -- that that really is a real effect.

DR. KIRKWOOD: True.

DR. SCHILSKY: Okay, John, thank you very much. We'll proceed to the sponsor's presentation.

MS. da SILVA: Thank you. Good afternoon, everyone, ladies and gentlemen of the advisory committee and FDA. I'm Loni da Silva, Program Director of Regulatory Affairs at Hoffmann-La Roche, and this afternoon we'll be discussing Roferon-A for stage II treatment of malignant melanoma.

The proposed indication which we are seeking is adjuvant therapy of and prevention of recurrence in surgically resected stage II malignant melanoma, Breslow tumor thickness greater than 1.5 millimeters, in patients without clinically detectable lymph node metastases at a low dose of Roferon-A, 3 million units, subcutaneously 3 times weekly for 18 months.

Our presentations this afternoon will consist of two speakers. Our first speaker is Dr. Antonio Buzaid, the Executive Director of the Oncology Center, Hospital Sirio-Libanes, Sao Paulo, Brazil, who is also the former Medical Director of the Melanoma Unit at Yale and former Director of the Melanoma Skin Center at M.D. Anderson. He will be discussing the clinical overview of malignant melanoma and concentrating also on the difference in the staging between specifically stage II and stage III.

He will be followed then by Dr. Leon Hooftman, who is our Director of Oncology at Hoffmann-La Roche. He will be presenting our data on Roferon-A in the treatment of stage II malignant melanoma.

Specifically we'll be focusing on these key points. As I said previously, you will hear the differences between the disease stagings, specifically stage II and stage III, and that our data shows a prolonged disease-free interval compared to no treatment, that disease-free interval is our primary endpoint and is a good predictor for overall survival. There is a strong trend towards increase in overall survival, and with low dose Roferon-A, it has a well established safety profile.

With that, I would like to call Dr. Antonio Buzaid.

DR. BUZAID: Good afternoon, Chairman, members

of the committee.

My focus and task today is to provide an overview on prognostic factors of patients with melanoma stage I and II, briefly also in stage III disease, and finally provide a snapshot on adjuvant therapy of melanoma.

As you all know, the incidence of melanoma is growing markedly worldwide. In fact, in the U.S. by the year 2000, 1 of 75 Americans will have the diagnosis of melanoma.

As far as the staging is concerned, we currently have four stages for melanoma. Stages I and II pertain to patients with primary melanoma. Concerning the next presentation, clinical stage II disease are those with Breslow depth greater than 1.5 millimeters. Stage III disease was just presented by John, and it's basically patients with nodal metastases and also in-transit metastases, and stage IV is basically distant disease.

Most patients with melanoma present with stage I and II disease at the time of diagnosis. Obviously, the prognosis is very different otherwise it wouldn't be called stage I, II, and III. But it's important to emphasize a few things here.

First of all, in the stage I and II category, the slope of the curve goes down very slowly, while here, as you can see, stage III disease is a very rapid drop. In

fact, about 80 percent of the patients with stage III disease recur in the first 3 years, while only half of the patients with stage II disease. These patients probably have a lower microscopic tumor burden because imaging studies are usually negative in this setting. Although they recur, they recur in a much more slower fashion, while patients with stage III disease probably have a larger microscopic tumor burden because you can see that with CT-scans, but the curve drops reasonably rapidly.

Let's focus on the prognosis of primary melanoma, that is, stages I and II. Looking at one of the largest databases, about almost 5,000 patients, University of Alabama and Sidney Melanoma Unit database, the three most important factors is the Breslow depth or obviously tumor thickness, ulceration, the location of the primary, the pathologic stage, whether or not the nodes were involved regionally, level of invasion, Clark level, sex, and age. But the most powerful factor is obviously Breslow depth.

The Breslow, as you all know, is measured from the granular layer of the epidermis to the deepest melanoma cell that can be seen in the microscope, and there is obviously a direct correlation between tumor thickness and outcome. It's for patients less than 1 millimeter, 1 to 2, 2 to 4, and graded in 4 millimeters.

We know well that this correlation is direct but not linear, in fact, is relatively linear up to 5 millimeters or so, 4 to 5 millimeters, and then it flattens out somewhat. So, very thick lesions, if you have an 8 millimeter or a 6, it may not make a tremendous difference, but if you have a 2 versus 4, the jump is tremendous.

Now, let's focus a little bit on disease-free survival and overall survival. There are very few series in the medical literature that present data on disease-free survival in primary melanoma. This is the largest data set, 5,000 patients from Duke University, and the only one that actually has both curves clearly outlined. There are important messages here.

The first one is obviously -- this is shown by tumor thickness in groups between 0.76 and 1.5, 1.5 to 4 in blue, and finally orange, greater than 4 millimeters. The solid line is overall survival; the dashed line, disease-free survival.

First of all, there is obviously a direct correlation between disease-free survival and overall survival, as you would expect in melanoma. This is not testicular cancer, but you can salvage almost everybody with chemotherapy.

Now, on the other hand, there is about a 25 percent difference, absolute difference, that you see in

general, about 20-25 percent for almost each category, and you need to understand why this is happening here. So, you have patients that recurred but haven't died. These are patients with primary melanoma. The major element that explains the difference between disease-free survival and overall survival here is surgery because two-thirds of the patients with primary melanoma recur regionally, in general nodal metastasis, and about 40 percent of the patients that recur with nodal metastasis, you can salvage them with surgery. This gives you about 40 percent out of two-thirds, which is about 20 or so percent of the patients. So, the major difference between disease-free and overall survival is explained by surgery for regional metastases. Nonetheless, still the majority of the patients that recur eventually die, at least about 70 percent of them.

Sentinel node mapping is a novel technique for melanoma, although very old for other cancers. It started in melanoma in 1992. In sentinel node biopsy, basically we inject a blue dye and/or a radioactive material and try to find the first node the melanoma cells would drain to if they were to metastasize. That's the concept of sentinel node, and basically after the injection, you find the blue node and send it to pathology. We know that there was a strong correlation between this node and the remaining of the nodal basin. If this node is negative, there's about a

98 percent chance the rest will be negative. If it is positive, it's positive.

One of the largest databases in sentinel node mapping is from M.D. Anderson, Lee Moffit Cancer Center. It's about 500 or so patients recently published in the Journal of Clinical Oncology. As you can see here, there was a direct correlation between tumor thickness and the chances of having positive microscopic nodes. That's identical data to the elective node dissection in the past. As you can see here, pertaining to this particular presentation, greater than 1.5 millimeter Breslow depth has about a 22 percent chance of having microscopic nodal metastases. So, about 80 percent of the patients will be node-negative.

When you have such a database where all patients underwent sentinel node mapping, we've learned that the most powerful prognostic factor, if you do have that piece of information, is the sentinel node histologic status. In the multi-variate analysis, this is the most significant factor followed by Breslow depth. If you do not have sentinel node information, Breslow depth is the most powerful prognostic factor.

This is the actual Kaplan-Meier survival curve for disease-free survival. All patients studied. The negative patients, the curve goes up, so it's a more

favorable subset now, and those with positive nodes, obviously the curves do go down and go down relatively rapidly. This is disease-free. But not everybody has died yet. As you can see, about half of them have already died, and the majority of patients with sentinel node have only 1 positive node. That's why the curves look so favorable.

This leads to the next topic which is the prognosis of patients with regional metastases, primarily nodal metastases. Like all the other cancers in oncology, the number of positive nodes is the most powerful prognostic factor for patients with nodal metastases. Presence of extranodal extension is also an adverse effect, and also patients with dual nodal basin versus only one nodal basin as a more unfavorable group.

This is a Kaplan-Meier using an overlay graphic technique. What you can see from this slide here is that if you have nodal metastases, at least half of the patients will eventually die, and in fact, looking at all curves in general, about 70 percent of the patients will die. That is about 30 percent of the patients in general will be alive at 10 years, if you have nodal metastases.

Again, this difference pertains to the number of positive nodes. That is, patients with 1 node in general have about a 40 percent chance of being alive at 10 years. Patients with multiple nodes have usually about a

20 percent chance of being alive at 10 years. Patients with extranodal extension have about a 10 to 20 percent chance as well.

Now, as I pointed out before, if a patient has a primary in the back and this patient has 2 lymph nodes involved in one axilla, this patient fares a little bit better than a patient that would have both axillas involved in a primary in the back. It is 1 node on the left and 1 on the right. This patient will fare worse than one that has 2 nodes and one site only. This is single nodal basin versus dual nodal basin for the same number of nodes.

Finally, as far as subcutaneous and intradermal metastases, what we call in general in-transit metastases, the patients have a poor prognosis. Again similar to the patients with nodal metastases, about 70 percent of them in general will be dead at 10 years. This is similar to patients with local recurrences.

A snapshot on adjuvant therapy. As you all know, melanoma is the most serious type of skin cancer, which has a high chance, depending on the prognosis of the patient, to metastasize. Multiple attempts have been made in order to reduce this risk of recurrence. In the past -- this is all randomized phase III studies from stages I up to III -- chemotherapy has been employed, and the drug that has been most widely studied was carbazine. Other

regimens, some of them somewhat bizarre regimens, have also been studied and showed no impact in disease-free or overall survival.

Specific monotherapy, such as BCG, C. parvum, transfer factor, or gamma interferon, and levamisole, somewhat controversial but also considered negative definitely in this country, showed no impact in disease-free or overall survival. As you all know, when you combine things that don't work, they usually don't work well. We've done that in oncology as well. DTIC plus BCG is of no benefit in terms of overall survival or disease-free survival.

Vaccines have a tremendous appeal for the population. Whether it helps patients with melanoma, we don't know. What we know to date is there are two randomized trials reported. They're relatively small studies, but both were negative. The first trial is in the vaccine in melanoma, oncolysate, VMO. It was as negative as you can imagine. The p value was 0.99 and 0.88. The Memorial Sloan-Kettering program using a ganglioside had a very modest impact on disease-free survival and has been evaluated further in larger randomized trials, but again it was preliminarily negative. Other vaccine programs are ongoing and the results are not as of yet available.

Finally, interferons. John Kirkwood has

presented in absolute detail the ECOG 1690 and the ECOG 1684 data. He also alluded to the North Central Cancer Treatment Group protocol and WHO 16. It's important to emphasize that these studies were conducted in patients primarily with node-positive disease. The ECOG trials, about 80 percent of the patients had basically node-positive disease; the North Central, at least two-thirds have node-positive disease; and WHO was completely node-positive disease. So, these studies are really different, different population of patients compared to the trials that will be discussed today.

The trials that will be discussed today will be two studies, two randomized trials, which include patients with clinical stage II disease, that is, patients with primary greater than 1.5 millimeters and clinically node-negative.

And I will pass now to Dr. Hooftman. Thank you.

DR. HOOFTMAN: Good afternoon, ladies and gentlemen, members of the committee, and FDA. My name is Leon Hooftman. I'm one of the R&D directors for oncology for Hoffmann-La Roche.

It's my pleasure this afternoon to present you the data that form the basis of the license application that's under discussion. We are here today to get the

recommendation of the advisory committee with regard to the license application concerning low dose Roferon-A for adjuvant therapy of stage II melanoma patients, that is, clinical stage II melanoma, clinically node-negative melanoma.

I will do my job reasonably well if I am able to discuss four specific important messages that form the basis of this presentation.

Further to what Dr. Buzaid said, I would like to emphasize the fact that currently there's no recognized standard therapy available for patients with stage II melanoma.

Secondly, there's a distinct difference for disease prognosis, as well as disease state, between stage II and stage III melanoma.

Thirdly, we believe that low dose interferon alpha 2a prolongs disease-free interval in a patient population that consists only of stage II melanoma patients.

And last but not least, there is a robust and strong correlation between disease-free interval as a parameter and the important long-term outcome parameter, which is overall survival.

To come back to one of these points -- and I apologize for the reasonable simple nature of this slide --

we have studied a low dose variety of Roferon-A for stage II melanoma only. The ECOG 1684 and 1690 studies have a certain proportion of patients with stage IIb, but the main body of the study is about stage III, which is node-positive disease.

The Cascinelli study only studies stage III disease, but with a low dose, the same dose as we have studied in our trials.

What is also important to note is that a certain proportion of all patients with stage I/II and a certain proportion of all patients with stage II will develop stage III disease, a certain proportion of patients thereof will develop metastatic disease which is not curable.

I would like to discuss now the two large-scale, randomized, multi-center trials that form the basis of our license application, one pivotal, one supportive, that were conducted in France and Austria, respectively.

The first study we call our pivotal study. It's the French study performed by the French Melanoma Group that started in January 1990, and the lead investigator was Professor Grob. This study recruited 499 patients.

The study that we use for supportive purposes is the study performed by the Austrian Melanoma Group, and

that study recruited 311 patients, started almost at the same time, February 1990. The lead investigator here is Professor Pehamberger, and both investigators are here with us today.

These larger studies prospectively studied the usefulness of a low dose of interferon, 3 million units, given 3 times a week for a duration of 18 months, in order to be able to bring down the incidence of recurrence of disease, in other words, as adjuvant therapy, for stage II melanoma.

The design of the first study that I am going to discuss is as follows. This is the pivotal study as conducted by the French Melanoma Group in France. This well-controlled study started, as I said, in January 1990, and patients were recruited until January 1994, over a 4-year period.

The patient population of this study consisted of clinical stage II melanoma patients only, that is, patients without clinical, palpable lymph nodes, in other words, clinically node-negative.

The dose used was 3 million units subcutaneously given 3 times a week for 18 months.

Patients were randomized within 6 weeks after surgery. Stratification by center was applied, but not by risk factors. I will get back to that later.

Here you see depicted the conduct of this pivotal study. As I said, it was initiated in January 1990, and the primary efficacy analysis was done in January 1994 when all patients were recruited. We'll have to go back to that later.

246 patients went into the observation arm. 253 patients ended up in the Roferon-A arm. Treatment duration was for 18 months for all patients. Prospective follow-up, as per protocol, was for 36 months, meaning that all patients were followed up for 36 months, but the patients that had been in the study longer had a follow-up of up to 7 years.

At that point, the prospective part of the study finishes and a retrospective section of this study starts. Patients were asked to provide a second, new consent and were seen once by the clinician in order to be able to collect data for long-term follow-up.

The primary efficacy endpoint, as used in this study, was disease-free interval. This is the time between initiation of therapy and relapse. This primary efficacy analysis was conducted as a sequential analysis. This part of the study was conducted as a sequential trial. A triangular test was used. The alpha was 5 percent; the beta, 10 percent; in other words, with 90 percent power.

The assumptions for the design of this study

were as follows. At 3 years, the investigators expected that 60 percent of all patients in the observation arm would have relapsed, and what they wanted to do was increase this figure to 75 percent for the Roferon-A patients, an absolute increase of 15 percent. For that purpose, they needed 104 relapses, and all together at the time they thought they needed 452 patients.

Three sequential analyses were performed. At the last sequential analysis, a sample size adjustment was performed as well, and a sample size adjustment was used in this trial in order to be able to stop recruitment in the study at the moment in time that enough data would be collected to be able to answer the predefined question and show the predefined difference.

A first interim analysis was performed in July '92, when a total number of relapses existed of 59: 34 in the observation arm and 25 in the Roferon arm. A second sequential analysis in April '93, but the main efficacy analysis was performed as the third interim analysis, the third sequential analysis, in January '94.

At that moment in time, there were 134 relapses in total, 80 in the observation arm and 55 in the Roferon arm, a difference of 25.

The null hypothesis of this analysis of this part of the trial was that observation was the same as

Roferon-A. At that moment in time, this null hypothesis was rejected. A p value was reached of .038. This demonstrated that, at that moment in time, Roferon-A statistically significantly prolonged disease-free interval as compared to observation.

Quite separately from this main efficacy analysis, a long-term analysis was performed for all patients with at least 3 years follow-up. These were further exploratory analyses of the primary efficacy endpoint and analyses of secondary efficacy endpoints, as there are overall survival and safety. They were performed at the end of the study. That was the time when all patients had reached at least 36 months in the trial. And I remind you that treatment continued for 18 months.

For this long-term analysis, we used an eligible patient population. The total number of patients recruited was 499. The eligible patient population consisted of 489 patients. We think that this is very close to an ITT, an intent-to-treat, population.

As you can see here, these were the patients excluded from these long-term analyses. The reasons for exclusion, as listed here, are in fact violations and would have normally been considered exclusion criteria as per protocol. The 5 patients that had no injection initially agreed to participate in the trial, but then immediately