the curves began, which is at 12 months after the start of study, approximately 9 months after the patients discontinued from study drugs, calls into question the potential impact of post-study therapy. Post-study therapy was not controlled by the protocol. The information was not prospectively collected, and the data was not completely available for all patients.

It is also important to note that the survival data in this study is not consistent with the efficacy parameters in the entire clinical program. All the other efficacy parameters, including survival, in the other two randomized studies were very comparable between the two treatment groups.

In summary, the results from this single-agent trial show that D-99 met the criteria for non-inferiority for the primary endpoint of response rates, as well as for duration of response and time to treatment failure.

Next I will present an analysis of the subgroups who were at high risk of cardiac toxicity, as defined by Dr. Alexander. Dr. Alexander had presented the meta-analyses for study 1 and 2 for this subgroup for the reduction in cardiotoxicity. For completeness, I'm now presenting for you the efficacy results for this group.

D-99 delivers very comparable antitumor efficacy for this group of patients who are at high risk of

cardiac toxicity. The hazards ratios are either very close to 1 or greater than 1, favoring D-99.

Now let's move on to study 3, the confirmatory trial for antitumor efficacy in combination regimen. Study 3 compares D-99 plus cyclophosphamide to equal doses of epirubicin plus cyclophosphamide. A question has been raised by the FDA regarding the dose of epirubicin, so I will take a few minutes to address this issue.

Epirubicin is a widely accepted anthracycline in Europe and in Canada for the treatment of advanced and metastatic breast cancer. In a combination regimen, epirubicin was approved at doses ranging from 30 to 75 milligrams per meter squared. Therefore, the dose of 75 milligrams per meter squared in combination with cyclophosphamide that is used in study 3 is actually at the high end of the approved dosing in Europe where this study was conducted.

On an equal milligram basis, epirubicin and doxorubicin were shown to be equivalent in antitumor efficacy. Randomized controlled trials were conducted using combination therapies including epirubicin and doxorubicin at equal doses.

The Italian trial was the largest of these studies, and as you can see there was no difference in response rates and no difference in survival at equal doses

of epirubicin and doxorubicin.

Similar findings were observed in the other three independent studies. There was no difference in response rates and no difference in survival.

Similarly, there was some single-agent studies that were conducted at equal doses of epirubicin to doxorubicin, and the findings were similar to the ones presented here. At equal doses, there were no differences in response rates or median survival.

The Ontario Cancer Care Guidelines recently published their assessments on the issue of dose comparability in patients with advanced breast cancer. A thorough literature review was conducted and a meta-analysis of six trials was performed totaling 983 patients being included. The results show that the hazards ratio for the efficacy parameters was very, very close to 1 and the Ontario group concluded that epirubicin and doxorubicin are equally efficacious in advanced or metastatic breast cancer.

Epirubicin was reviewed at the last June ODAC meeting and the application for the metastatic breast cancer was rejected and was not recommended for approval by ODAC. It is important to note that none of those studies submitted in that NDA compared epirubicin to doxorubicin. However, this issue of comparability of epirubicin to

doxorubicin was discussed by the FDA. Upon reviewing a meta-analysis that was submitted to the sponsor, the FDA concluded that the two treatment groups are comparable in the first-line therapy of metastatic breast cancer.

There were a number of studies conducted comparing escalating doses of epirubicin. Three studies compared doses of epirubicin at 50 to 100 milligrams. This study here was actually the study that was reviewed at the June ODAC meeting. While the results show that 100 milligrams per meter squared of epirubicin produced high response rates, there were no differences in survival with the higher doses.

This study here evaluated escalating doses of single agent epirubicin. The results show that there was a dose response up to 90 milligrams of epirubicin as a single agent, but there was no additional benefit beyond 90 milligrams per meter squared.

There was one study that was presented at this year's ASCO that compared 75 to 100 milligrams of epirubicin which showed that there was no additional benefit with 100 milligrams per meter squared.

Furthermore, there were also no studies comparing equal doses of epirubicin to doxorubicin where doxorubicin was shown to be superior.

Therefore, we are confident that the 75

milligrams per meter squared of epirubicin is an appropriate dose for comparison in study 3.

As mentioned earlier, the trial was terminated at 160 patients due to resource and administrative considerations. The responsibility for the clinical development program for D-99 was under the Pfizer Corporation until July of 1997. When The Liposome Company assumed the responsibility for TLC D-99 in 1997, this study was stopped due to resource considerations.

At the time of termination, there was absolutely no knowledge of the study results. The early termination reduced the sample size and resulted in a larger variance of the estimates, which actually makes it more difficult to meet the predefined criteria. As you will see later on, despite the reduced sample size, D-99 fulfilled the predefined criteria of non-inferiority in this study.

Now let's move on to the results from this trial. The endpoints in this study were identical to that in study 1, and the objective response and disease progression were assessed on a treatment-blinded basis by Dr. Noza Azarnia.

Patient characteristics were well balanced between the two treatment groups and objective response rate was 46 percent versus 39 percent, fulfilling the

criteria of ruling out the 15 percent delta as well as the more stringent criteria of ruling out a 10 percent delta.

The median duration of response was 10 months for the D-99 treated group versus 7.8 months for the epirubicin group. This difference was statistically significant at a p value of 0.03.

Time to progression was 7.7 months for the D-99 patients versus 6.3 months for epirubicin. The hazards ratio was 1.45 and the lower bound for the 95 percent limit was 1.03.

There was a statistically significantly longer time to treatment failure for the D-99 treated patients. The median was 6.8 months versus 4.4 months. The p value was 0.03, and the hazards ratio was 1.50.

There was no difference between the two treatment groups in overall survival. The median survival for D-99 was 18 months versus 16 months for the epirubicin treated patients. The hazards ratio was 1.15 and the lower bound for the 95 percent limit was .82.

Results from this study show that D-99 plus cyclophosphamide meets the criteria for non-inferiority compared to epirubicin plus cyclophosphamide at equal doses. It is important to note, once again, that the hazards ratios were all greater than 1, favoring D-99, and there is a statistically significant difference in duration

of response and in time to treatment failure favoring D-99.

In conclusion, the TLC D-99 clinical program provided two independent, well-controlled studies evaluating the antitumor efficacy of D-99 in combination with cyclophosphamide. Both studies fulfilled the criteria for non-inferiority for all of the efficacy endpoints. It is also important to note that the results from the D-99 plus cyclophosphamide arm in study 3 were very similar to that observed in the D-99 plus cyclophosphamide arm in study 1, hence providing an independent replication of the results for the antitumor efficacy of D-99 in combination with cyclophosphamide from the pivotal trial.

Dr. Jerry Batist, principal investigator for the pivotal study 1 and single-agent study 2, will now discuss the safety profile.

DR. BATIST: Thank you very much, Dr. Lee. Good afternoon, colleagues.

I've been an investigator with D-99 studies for over 10 years, so I have a large clinical experience with this novel formulation of doxorubicin, and I am pleased to be able to present the safety data to you.

For the most part, I'm going to be talking about one part of the database which is 323 patients, all of whom were in comparative phase III studies, in order to give you the context of a comparative trial with which to

compare the toxicities of D-99 as in study 1 to dogs and study 3 to epirubicin.

For some less frequently observed toxicities, I'm going to look at a larger number of patients which, as to these patients, patients who were treated on phase I and phase II studies, and that larger group all were treated at a starting dose of less than 100 milligrams per meter squared every 3 weeks. It's 542 patients.

Naturally in the development program of D-99, even larger, more intense doses were explored. Among those patients, there were no new toxicities or any different toxicities that you'll not see described in this discussion today. It's just that they were more frequent and more intense. Hence, the decision to focus our development on the 60 to 75 milligrams per meter squared range.

Now, you've heard from Dr. Alexander that D-99 results in significantly less cardiotoxicity than does doxorubicin, and the next few slides will summarize what I'm going to show you in terms of safety data this afternoon.

I'm going to show you that there's no increase in severity or incidence of doxorubicin toxicities and no unexpected toxicities; furthermore, that D-99 is associated with less mucositis and diarrhea than is doxorubicin; and also that D-99 patients presented with no grade 3 palmar-

plantar erythrodysesthesia, or hand-foot syndrome as we call it, which would result in dose modifications or reduction in the dose. In fact, among the very few cases of very low grade hand-foot syndrome, the incidence was 0.3 percent.

Then I'm going to describe the D-99 versus epirubicin study 3 toxicity. I want to put this in context a little bit. You've heard from Dr. Lee the published results of a number of studies in which epirubicin and doxorubicin at equivalent doses were compared. You've also heard that the overall response rates and the survival at equivalent doses were the same. There was no significant difference.

In the Pritchard analysis of all of these studies, what was also true was that in the doxorubicin group, there was more significant toxicity, more significant neutropenia, mucositis, fever, nausea and vomiting, and also importantly, more episodes of cardiac toxicity, both congestive heart failure and other indices of cardiac damage.

What you've heard today is that unlike doxorubicin at equivalent doses compared to epirubicin D-99 has a similar cardiac-sparing effect to that of epirubicin. On the other hand, like doxorubicin in all of those studies, it is more myelosuppressive and does result in

more mucositis.

This is the hematologic toxicities observed in study 1 and study 2 where D-99 is compared to doxorubicin, and they're generally very similar with the notable exception of grade 4 neutropenia, which is more frequent in the doxorubicin treated patients than in those who receive D-99. This is associated with a smaller number of patients who had this degree of myelosuppression for greater than 7 days. In study 2, looking at all grades of infection, there were significantly fewer in the D-99 group compared to the dox group.

Nonhematologic toxicities are shown here. The toxicities that are indicative of mucosal damage -- that is, stomatitis, mucositis, diarrhea -- in all of the cases, the numbers favor D-99 with less toxicity compared to doxorubicin. They reach statistical significance for stomatitis and mucositis in study 1 and study 2 and for diarrhea in study 2.

The treatment related deaths are shown here. 1 patient in study 1 died of sepsis in the D-99 group that was treatment related, and 1 patient in study 2 died of congestive heart failure in the dox group that was treatment related. There were no treatment related deaths in study 3.

The hematologic toxicities for study 3 are

shown here. As mentioned there is less myelotoxicity with epirubicin. There's a statistically significantly lower incidence of grade 4 neutropenia with epirubicin compared to D-99. There was not a statistical difference in the duration of the grade 4 neutropenia between the two groups and this did not result in any septic deaths.

Nonhematologic toxicities are shown here. They're generally the same with the exception of stomatitis and mucositis being less frequent with epirubicin compared to D-99.

Then finally I want to refer to that larger number of patients because I want to describe skin toxicities which are infrequent, though clinically very difficult for patients treated with doxorubicin. These are the patients, all of whom started at less than 100 milligrams per meter squared on a 3 weekly basis, 542 patients. In that number of patients, we saw only 3 cases of hand-foot syndrome, no grade 3 cases, which is the grade at which we would have to reduce the dose of the drug, 1 grade 2 and 2 grade 1. There was only 1 case of grade 1 radiation recall in the skin, and as it happened there were 7 patients who had accidental extravasation of D-99. In none of these cases was there necrosis or ulceration of the skin.

So, in summary, I've shown you that there was

significantly less cardiac toxicity of D-99 compared to doxorubicin. There is less mucositis and diarrhea compared to doxorubicin. Hand-foot syndrome is extremely rare, and in the few cases where it occurred, it was not severe enough to require dose modification. Where we saw extravasation, we saw no evidence of necrosis.

Now, I want to just take a second to respond to Mr. Cohen's concerns about safety that came up at the very beginning. I think we've made the case and our contention is that the dose of D-99 would be equivalent to doxorubicin in the 60 to 75 milligrams per meter squared range, so that there would be no requirement for a change in the practice of physicians, in contradistinction to what happens with Doxil. Therefore, there wouldn't be a safety issue in that regard. Where there might be a safety issue is in educating our colleagues that this is not Doxil.

Finally, as a practicing clinician who treats many women with breast cancer, I want to express my personal enthusiasm at the prospect of having this novel drug available. This provides dramatically reduced cardiac toxicity without bringing along with it any added new toxicities.

Thank you. Dr. Lee will provide a conclusion.

DR. LEE: Doxorubicin is one of the important agents for the treatment of breast cancer. All women

receiving doxorubicin as part of their treatment for breast cancer are at risk of the side effect of cardiotoxicity from doxorubicin treatment. Despite decades of efforts to reduce the cardiotoxicity of doxorubicin, there remains a continued and increasing need for a less cardiotoxic doxorubicin which could deliver a comparable level of antitumor efficacy.

TLC D-99 fulfills the objective of improving upon the therapeutic index of doxorubicin. On an equal milligram basis, D-99 significantly reduces cardiotoxicity while delivering antitumor efficacy that's comparable to doxorubicin. Both of these endpoints were demonstrated in the pivotal study 1 and these results were reproducible and independently confirmed in two separate studies.

Reduction of cardiotoxicity compared to doxorubicin is highly significant in both studies comparing D-99 to doxorubicin. D-99 reduces cardiotoxicity by 75 to 80 percent. The estimated dose at which a 5 percent risk of developing congestive heart failure was 780 milligrams per meter squared with D-99. Based on our own database, as well from the literature, the 5 percent risk of congestive heart failure for doxorubicin was 400 to 450 milligrams. So, compared to doxorubicin, this difference represents an additional 4 to 5 cycles of treatment with D-99 before reaching the same level of risk of CHF.

Preservation of antitumor efficacy is demonstrated in both studies using combination regimens with cyclophosphamide, the indication that we are seeking today. In both studies 1 and 3, hazards ratios are all greater than 1 in favor of D-99, and all the parameters met the criterion for demonstrating non-inferiority.

D-99 is associated with other safety advantages. Compared to doxorubicin, D-99 had significantly less mucositis and diarrhea. The D-99 formulation of liposomal doxorubicin is not associated with increased hand-foot syndrome. There were no reports of grade 3 or 4 hand-foot syndrome in the entire clinical program. TLC D-99 is also not associated with severe necrosis or ulceration upon accidental extravasation.

TLC D-99 provides clinical benefits and offers an important therapeutic option for breast cancer patients. Compared to doxorubicin, D-99 is a safer formulation while delivering comparable efficacy. The demonstrated patient benefits support the approval of D-99 for the first-line treatment of metastatic breast cancer in combination with cyclophosphamide.

Thank you very much for your attention, and we are ready to take your questions.

DR. SCHILSKY: Thank you very much.

Are there questions from the committee? Dr.

Nerenstone?

DR. NERENSTONE: Correct me if I'm wrong, but I think the way you defined time to treatment failure was a mixture of patients who progressed as well as those who had to drop out due to toxicity. One of the ways people could stop on the doxorubicin arm was if they reached a value of doxorubicin where the investigator thought that they might have some cardiac toxicity at a certain dose of doxorubicin, even in fact if they had no signs of cardiac toxicity. It was a nonobjective level, but it was at the investigator's discretion.

Can you tell me how many patients were actually stopped because of the cumulative dose of doxorubicin they had received and if any of those patients were actually responding at the time they were stopped?

DR. LEE: Dr. Nerenstone, can I clarify? You're asking the endpoint for time to progression or time to treatment failure?

DR. NERENSTONE: I believe it was your time to treatment failure.

DR. LEE: All right. May I have the slide?

This is for study 1. 5 percent of the patients versus 10 percent of the patients were stopped due to an ejection fraction drop, a documented cardiotoxicity endpoint. An additional 1 percent versus 5 percent of the

patients were discontinued from treatment because the investigator or the patient was concerned that there may be additional risk of cardiotoxicity and hence the patient was stopped from treatment.

DR. NERENSTONE: Do we know how many of those patients were actually responding when they were stopped? Or by virtue of the fact that they were stopped, they had to be responding at that point. Is that correct?

DR. LEE: By virtue of the fact that they were stopped due to a toxicity endpoint, they have not progressed. They could be either responding patients or stable patients.

DR. NERENSTONE: And I have one other question. Dexrazoxane was allowed on study, and there's no discussion of how many patients were on it and anything about their response rates as opposed to those patients who were not on it. Can you give us any more details about that?

DR. LEE: None of the patients actually was treated with dexrazoxane in this study because the protocol allowed patients to be treated with the amendment at the higher doses.

DR. SCHILSKY: Dr. Margolin?

DR. MARGOLIN: I'm going to sort of pursue the same pathway and hopefully break my set of questions down into not too many.

The first one I guess is the most straightforward one and maybe directed at one of the statisticians, which is, even though what appeared to be your posteriori look-back at the distribution of your identified cardiac factors looked like they were extremely well balanced between groups, at least as I recall in study 1, I'm not certain from a statistical point of view whether that's as good as prestratification which was apparently not done for preexisting cardiac risk factors. That's question 1.

DR. LEE: The only stratification factor in the study was prior use of adjuvant doxorubicin. That is the only stratification factor. There were no other stratification factors for other cardiac risk factors.

DR. MARGOLIN: I know that. That's why I asked the question.

DR. LEE: Could you please clarify?

DR. MARGOLIN: Maybe Dr. Lamborn could answer the question. You understand my question. Right?

DR. LAMBORN: As I understand, you're referring to slide 53, study 1, pivotal, cardiac risk factors where they list the balance of the risk factors.

DR. MARGOLIN: Exactly.

DR. LAMBORN: And the question was would there have been any additional benefit to doing baseline

stratification, and my sense is, as long as you ultimately have balance, which is what you hope for, that that's okay.

DR. LEE: Thank you very much.

DR. MARGOLIN: Thank you.

Then I guess the only other related question I'll ask is probably more rhetorical, but I'm also bothered by -- I guess you had probably some very good consulting cardiologists and people who specialize in this area, but it was somewhat bothersome that on slides 54 and 55, which I think are the same slide, you had quite a few patients grouped in the 400 to 499 milligrams per meter squared cumulative dose of dox, and even 7 patients 500 to 599, and 2 600 to 699. Some of us would probably stop sooner than that or would have added the dexrazoxane even without changes in the MUGA determined LVEF because it gets a little scary up there, and we know those tests are not perfect.

DR. LEE: I'd like Dr. Winer to address that please.

DR. WINER: Can you just clarify your question a little bit? Are you asking about whether those patients should have continued to be treated or --

DR. MARGOLIN: I guess the comfort level of allowing treatment to such high doses just because there wasn't a change in the LVEF or because the protocol said

you didn't you have to or because the dexrazoxane was considered p.r.n.

DR. WINER: Well, it's a tough issue. In the course of the study, every attempt was made I believe to try to make sure that patients would be treated as safely as possible in the sense that MUGAs were done every cycle after 500. They were shipped off to Yale. There was no treatment decision until the individual institutions and Yale had both read them. It's still a concern.

I think that in every situation, it's a matter of weighing the risks and benefits of continued treatment. I can tell you personally, having taken care of many of these patients, with each and every patient we debated whether it made sense to go on. There were patients who pushed to go on where I said, no, I really want to stop. Undoubtedly there were patients who came off the trial. I know for sure there were because I took some off because of cumulative dose. I think it's worth mentioning that those were patients who were not considered treatment failures. They were censored in that analysis because, in fact, it wasn't a treatment failure. They came off.

DR. SCHILSKY: Dr. Kelsen?

DR. KELSEN: Can I ask a follow-up to that? It's a question for your position and then a question about the survival curves.

What is your position on the addition of Zinecard to this agent should it reach approval? Are you going to recommend that it be used at a certain dose, or do you feel so strongly that cardiotoxicity is so unlikely that you believe there is no need for a cardioprotective agent?

DR. LEE: I would like to call upon Dr. Speyer who has a lot of experience with Zinecard to address this.

DR. SPEYER: Zinecard is a very interesting drug.

DR. SCHILSKY: For the record, could you identify yourself and state your affiliation?

DR. SPEYER: I'm Dr. James Speyer.

Zinecard is a very interesting drug. As you know, Zinecard is only approved for use after cumulative dose of doxorubicin at 300 milligrams per meter squared. It is an additional drug that has to be added, and therefore there is possible additional myelosuppression.

The possibility of interference with the antitumor efficacy was debated at this committee a number of years ago. It certainly led to the recommendation for waiting until patients had gotten 300 milligrams per meter squared. It also has led in the community to some reluctance to use it.

This trial simply didn't test that. What D-99

provides is an opportunity to use an anthracycline from the start with significantly less cardiotoxicity with doxorubicin.

DR. KELSEN: I do appreciate the answer but if I could just ask this again. Is your position that the cardiotoxicity risk is so small -- I believe I heard 1.5 percent on one of those slides -- that you believe another advantage of this is that you do not need, at any dose, to use a cardioprotective?

DR. SPEYER: My own personal view is that that's the case, and the only way I would think about doing that is if I were really seeing benefit and I were seeing a fall in ejection fraction and we weighed the risks versus benefits in that particular case.

DR. KELSEN: Then a second question. We're going to hear from the review board I believe a significant concern about survival in study 2, and I just want to make sure I understood the point that you were making. If I understood what you were saying correctly, the curves overlap during the period of time that patients in study 2 were receiving the study compound or the conventional control arm. The curves diverge in favor of the conventional arm after the median was well past, but you don't have data to address the issue as to whether salvage was different between the two arms. Is that correct?

DR. LEE: That is correct. One more point of clarification. Actually the curves overlap for the first 12 months. In that study, actually the median duration of treatment was 3 months. So, actually the curves overlap for an additional 9 months after the end of treatment for a majority of the patients.

DR. SCHILSKY: Dr. Raghavan?

DR. RAGHAVAN: A follow-up question to Dr. Kelsen's because I guess I missed the point of the answer. This is study 2, you slide 86. So, the curves are very nicely together for the first 12 months, but then at 18 months, there's a 15 percent difference in survival. At 24 months, there's a 15 percent difference in survival. The curves come back together again and then they diverge, although by then the power of the study is weak because there are 3 and 8 cases, although you could ask the question does that mean that there just weren't enough patients alive of the ones who were further out in the D-99 arm.

So, I always get very uneasy if I see a survival curve that starts to drop away. You've kind of dismissed it and I kind of want to bring you back to it. So, can you talk a little more about the one test of head-to-head efficacy of new drug versus old drug? Talk a bit more and let's forget about the first 12 months. Let's

talk about longer-term survival. What's the deal there?

DR. LEE: First of all, given the totality of the information, we really do not believe that this divergence in the curve is a reflection of treatment differences.

I'd like to call upon Dr. Eric Winer, who is principal investigator for this study, to give some clinical perspective.

DR. WINER: The data are the data, and in study 2 there is a trend towards a lower survival in patients who were treated on the D-99 arm. It's something that the sponsor has been very concerned about. It's something that I think all of the consultants have paid a great deal of attention to. I certainly have. I was involved in that study.

I think for a number of reasons I'm quite convinced that this is not a result of the inferiority of D-99, and I think there are a few issues to bring up.

One is that in that study the response rates are absolutely identical in the two arms. There are few treatments, unfortunately, that clearly have been demonstrated to change survival in metastatic breast cancer, and in almost all cases where that has occurred to my knowledge, it does not occur in the absence of a difference in response. So, it certainly doesn't go along

with the other parameters in the study.

Second, in the two other studies that have been presented, no difference in any efficacy parameter.

And third -- and I think this is the weakest of the points but I think it's worth noting -- the survival curves do stay together for those 12 months and separate after that. At least in my mind, it at least raises the question of the extent to which post-study treatment is playing a role. Post-study treatment was not controlled for at all, and even retrospectively going back and looking at it, the sponsor really was not able to sort that out.

So, I think for all of those reasons, I don't believe that the survival curves indicate that it's a less efficacious drug. But it's an issue.

DR. SCHILSKY: Other questions? Dr. Nerenstone?

DR. NERENSTONE: This is going to brought up I know by the FDA reviewers, but for your third study you make the argument that epirubicin is equal to doxorubicin on a milligram-per-milligram basis. However, when you look at the toxicity profile, in particular on your slides 117 and 118, clearly the D-99 is more toxic with ANC less than 500, with a p value that's significant compared to the epirubicin combination arm, likewise the stomatitis, mucositis, and greater than grade 3 side effects. Overall,

there's also nonsignificantly nausea and vomiting -- you can argue that's pretty close -- but infection and neutropenic infection, although those are not statistically significant.

This is a reversal of the toxicity profile that we saw in studies 1 and 2. How do you interpret those dissimilar results if you assume that epirubicin is equal to doxorubicin?

DR. LEE: I'd like to call upon Dr. Jerry Batist to discuss this.

DR. BATIST: Thank you.

You've heard reference to one of two meta-analyses that were performed to look at the question of what is the right dose of epirubicin relative to doxorubicin. The one that you heard most about was the one that gathered a group of eminent medical oncologists who practice principally in the treatment of breast cancer, chaired by Kathy Pritchard. They were charged with establishing practice guidelines for the use of epirubicin. Their conclusion was that it was absolutely equivalent in antitumor efficacy, and their recommendation, moreover, was in this dose range, 60 to 75 milligrams per meter squared.

Now, you're absolutely correct that we saw enhanced toxicity in some parameters. As I pointed out, and was observed in that meta-analysis as well, at

equivalent doses of doxorubicin and epirubicin, that's exactly what is seen as well: equivalent overall response, equivalent overall survival, increased nausea/vomiting, increased neutropenia, increased mucositis, but also importantly increased cardiac toxicity. That was observed even in a number of very small series of patients. And that's where there's a dramatic distinction.

So, this is equivalent to doxorubicin we can say, as it performs on a milligram-per-milligram basis in comparison to epirubicin, enhancement in some of these toxicities, some, but not all, but it's very different from doxorubicin because it preserves the cardioprotection that we see with epirubicin.

DR. LEE: If I may, I would also like to call upon Dr. Bob Leonard who is one of our investigators from study 3 and has a lot of experience using epirubicin in Europe. Maybe he can shed some light on the equal dose issue.

DR. LEONARD: Yes. I'm Bob Leonard, medical oncologist in Edinburgh, Scottland.

As you're probably aware, we have an enormous experience of using epirubicin in Europe as a standard drug in the treatment of advanced breast cancer and increasingly recently used in the adjuvant setting as well.

Now, the data you've heard in the comparative

analyses simply speaks to the experience that we've had using the drug in combination where we regard it as an equivalent drug to doxorubicin with probably a better safety profile in terms of the clinical toxicity and probably in terms of the cardiac toxicity as well. None of the data we've heard in the studies today would contradict that routine clinical experience that we have. It's very hard to use doses of epirubicin above the equivalent of 25 milligrams per meter squared weekly even as a single agent in advanced breast cancer because of clinical toxicities. So, it's a standard dose of a standard drug used very widely and in fact in Europe at a much higher level of patient exposure than even doxorubicin is these days.

I don't know if that answers all of your questions, but I think it answers some of them.

DR. SCHILSKY: Thank you.

Dr. Johnson.

DR. DAVID JOHNSON: I want to also come back to Dr. Kelsen and Dr. Raghavan's point because, as I've listened to the presentation and read through the material provided by the sponsor, I'm reasonably persuaded that there's less cardiac toxicity. I'm not persuaded that there's equal efficacy to doxorubicin.

What you've given us are three studies, none of which was powered or designed to really look at survival

endpoints. They looked at response endpoints.

DR. LEE: That's true.

DR. DAVID JOHNSON: Actually you gave us data about epirubicin where you can change response but you didn't change survival. In your own data, you showed us that.

The third study you showed us was actually an aborted study. So, I'm having a hard time even figuring out how to use those data.

So, I'm equally concerned with this difference in survival, and I think it would be surprising to me, frankly, if you did not have information about second-line treatment or the subsequent treatment for these patients and at least some differences that might have emerged that could have explained or given some possible explanation for the difference in survival beyond the fact that this drug is less efficacious than doxorubicin.

So, for example, if you could show me that the doxorubicin arm all received a taxane or 50 percent of them did or so and no one on the D-99 arm did, I would be much more comfortable with your conclusions about the comparability of the efficacy of the two agents.

DR. LEE: Dr. Johnson, I'd like to respond to the issue about the lack of information on post-study therapy, and then I'd like to call upon Dr. Steve

Piantadosi to talk about the validity and integrity of study 3 being a terminated trial.

First of all, the post-study therapy information was really incomplete for us to make any kind of inference or conclusion on. We simply didn't have information on 20 percent of the patients. Bear in mind that this study actually started in 1992, quite a while back, and we actually retrospectively tried, as hard as we can, to collect complete information on post-study treatment, but the information was simply not available for 20 percent of the patients, and even in the remaining 80 percent of the patients, there were a lot of patients that clearly had missing data. As you might imagine, patients move around. They move to other sites. They may have died some years ago, and it's very difficult to go back and collect information to the extent that we can address this post-study therapy issue.

Then I'd like Dr. Piantadosi to maybe address the study 3 early termination.

DR. WILLIAMS: Could I address the point you just mentioned before we go on?

From our review of your data, we saw that 69 percent in each arm, at least, had been documented to get a taxane. Now, there may have been, in that other 31 percent in each arm, others, but at least from the data you

submitted, 69 percent in each arm had a taxane.

DR. LEE: Yes, that is true, and actually we interpret that as sort of an indication of the lack of information also because during this period of time almost all metastatic breast cancer patients who had already been treated with a doxorubicin-based treatment would have received taxanes, and close to 30-40 percent of them, according to the records, didn't receive taxanes. So, we really do not believe that we have adequate information to draw any conclusions from this.

Dr. Piantadosi?

DR. PIANTADOSI: Thank you. I'm Steve Piantadosi from the Johns Hopkins Oncology Center, a consultant to the company. I'm sure the committee thought they had seen the last of me this morning, but I'm back.

(Laughter.)

DR. PIANTADOSI: I may be here tomorrow too, weather permitting.

I share the concern over study 3 and the illustration of an unfortunate consequence of having to do clinical trials in the real world, and that is resource limitations.

The essential question, as far as I'm concerned, is whether or not you believe that the study represents an unbiased estimate of the treatment effect.

The only way that the study could be biased is if a decision about termination were made with knowledge of the results, barring some other disaster with disclosed randomizations or something like that. But we're very confident in the methodology of the study. The only issue is whether or not it was terminated with knowledge of the results, and in fact that was not the case.

Now, as such, one could then imagine that the study had been designed with this particular sample size that is not an aborted study at all. Consider, for example, if the study had gone to twice its sample size rather than half its sample size. Would anybody then speculate, well, we should throw out half the data because the investigators didn't follow the original plan to conduct the study? And I think the answer is no. So, as long as you have confidence that the study is showing you an unbiased estimate of the treatment effect, it should be considered for the strength of evidence as it stands.

As such, it represents a higher hurdle for the treatment to overcome because the imprecision that results from the smaller sample size widens the confidence intervals on all of the outcome measures, primary and secondary, and as you saw from the results, even so, this study meets the prespecified hurdles.

DR. LEE: I just want to add and assure you

that we had absolutely no knowledge of the study results when we terminated the trial.

DR. DAVID JOHNSON: I believe you.

(Laughter.)

DR. DAVID JOHNSON: I actually have another question, and that has to do with your proposed indication. Why are you specifically asking to do this in combination with cyclophosphamide?

DR. LEE: We have two independent randomized trials conducted in combination regimens that show very comparable and replicated, reproducible antitumor efficacy, and on that basis, we are recommending the indication for combination treatment.

DR. DAVID JOHNSON: Yes, that's true, but you could also make the argument that "comparability" then is on the basis of the cyclophosphamide.

DR. LEE: I'd like Dr. Eric Winer to address this, please.

DR. WINER: You could make that argument. Cyclophosphamide was there. The question is in those studies is cyclophosphamide masking any inferiority of D-99 versus doxorubicin. Obviously, there's no experiment that can provide an answer to that. Cyclophosphamide is an active drug. It's a less active drug, I think we all believe, than doxorubicin. I personally think that it's

unlikely that it would be masking an effect, and in other trials that have included cyclophosphamide where other agents have been compared, there have been differences seen despite the presence of cyclophosphamide.

If I can just, for a second, come back to your question about the survival issue because it's clearly a very, very important issue, and I realize I already addressed this, so bear with me for a second.

But in the absence of seeing a difference in response in the two arms in study 2 or a difference in response duration or a difference in time to progression, it's very hard for me to attribute a difference in survival that's occurring many months out to the two treatments in question. I'm left, although it is less than a fully satisfying answer, recognizing that metastatic breast cancer is an incredibly heterogeneous illness and that there are differences in patients and that some of that can happen in studies that are of reasonable size but not huge.

DR. DAVID JOHNSON: In response to that, Eric, let me say that treatment for some other solid tumors -- and I'll use an example, say lung cancer, non-small cell. Treatment is brief and oftentimes the curves will follow along precisely the same for several weeks, to even months, after treatment is completed, then separate out well beyond the time of treatment administration. So, I don't know

that this is necessarily unusual or unique to this particular study. This may be something we've seen in solid tumors. So, I don't know that, just because it doesn't follow the trend of the other two studies, we should throw these data out.

I'm very concerned about these data actually. As I said, it troubles me about whether this is in fact comparable and efficacious. I suspect it is, but that's not what we're asked here. We're not asked to guess what we think. We're asked to render an opinion on what the data show.

DR. SCHILSKY: Let me ask one question since it hasn't come out yet and this has sort of been disturbing me, as I've been sitting here listening to the data. I'd appreciate any comments from FDA as well. It has to do with the choice of response rate as the primary efficacy parameter.

Typically the committee usually considers response rate as at best a surrogate estimate for efficacy. Here a decision was made to develop the entire clinical program based on a response rate as the primary efficacy parameter. I'm puzzled by how that decision was reached. I'm actually quite surprised that that was the agreed upon efficacy parameter. Perhaps we could have some discussion, and I'll ask you first, Dr. Lee, if you could just describe

to us how the decision was reached to use response rate as the efficacy parameter.

DR. LEE: Okay, and then perhaps Dr. Williams can add to that.

During the discussion of the protocol design, response rate was chosen as the primary endpoint at least in part because of the fact that we built in three interim analyses. In order for the three interim analyses to really be realistic so that, as you accrue the patients, you have sufficient information to act upon it, the information or the endpoints that you have to act upon should be rather readily available rather than having to wait for a long time.

However, in discussions with the FDA review team, they also said that you can use response rate as a primary efficacy endpoint but we need to make the commitment to follow up and collect sufficient information for time to progression. Hence, as you can see, in all of the three trials, we did follow up patients to provide sufficient information on time to disease progression. In fact, if you had chosen time to progression as the endpoint, we would have met the criteria for study 1 and 3, and for study 2, we just missed because of the duration in that study.

Dr. Williams?

DR. WILLIAMS: Rich, I was involved with these decisions, so I know what we were thinking because I was thinking them.

(Laughter.)

DR. SCHILSKY: Maybe you could share that with the rest of us.

DR. WILLIAMS: You're correct. This was novel. We had not done it before. The thinking behind it was that this was doxorubicin, and if you could ever believe a surrogate that tumor response was going to reflect a benefit you're going to see, that it would likely be in this surrogate.

I think we really stretched because we thought this was an important trial to do, and it was basically impossible to do an equivalence trial I think sized for survival equivalence, the number of years you have to go in breast cancer to see a death and then to do equivalence. So, because there was the hope that there was going to be a decrease in cardiac toxicity, we said we would take one-sided. We said we would take a ratio of response rates of .75, not the 15 percent absolute amount that was quoted earlier. And I think we really stretched. My feeling is that's as far as we should stretch to try to bring a drug with a decreased cardiotoxicity into the marketplace.

We had seen with dexrazoxane the actual

decrease in response rate in the largest trial of that submission. So, we had seen this occur.

It's still debatable whether that was the right decision, but that's the decision we made and that's why I don't feel very flexible about stretching those boundaries any further.

DR. SCHILSKY: Yes. I was just curious about the thinking because this morning we saw a drug which essentially is fluorouracil, and yet the studies were required to be powered around a survival endpoint and result in much bigger trials. So, I think it's important for the committee to understand sometimes the thinking behind --

DR. WILLIAMS: Of course, the response rate is higher in breast cancer. Maybe the response rate is a little more respected in breast cancer as a manifestation of effect, but I still think it is debatable.

DR. SCHILSKY: Other questions for the sponsor? Dr. Lippman?

DR. LIPPMAN: I'd just like to ask, I guess, Dr. Piantadosi about this. I don't imagine it would have an effect, but in study 2, which is the one that is raising the concern, in the design there were three planned interim analyses, but because of rapid accrual, I guess one or two of them weren't done. Does that affect the statistical

power? Does that change anything?

And then I guess the other comment that I have would relate to slides 87 and 90, this issue of looking at efficacy just in the patients at high risk for cardiotoxicity. But initially the interim analysis question.

DR. PIANTADOSI: Yes, thank you. My opinion regarding the interim analysis question is that having missed an interim analysis carries no impact or importance on the final outcome of the study. The study could easily have been designed without interim analyses and we'd be looking at the same data.

DR. LIPPMAN: I guess I was just wondering if it actually strengthened some of the findings because we know the more interim analyses you do, that you have to adjust for that I guess.

DR. PIANTADOSI: Well, you adjust typically the type 1 error level. No, I don't think it strengthens any of the findings.

DR. WILLIAMS: We didn't adjust for the one interim analysis that was done. Right? Wasn't there one interim analysis performed?

DR. LEE: For study 2 now, right? The second interim analysis and third interim analysis were performed and the boundaries were adjusted way ahead of time. We

stuck to those boundaries.

DR. WILLIAMS: And the final overall alpha is -- what is that that's presented? 95 percent or 90 something else?

DR. LEE: The overall type 1 error is still .05.

DR. WILLIAMS: And did you use 95 percent confidence intervals, though, at the end or adjust --

DR. LEE: Yes.

DR. PIANTADOSI: The interim analyses were conducted using very conservative O'Brien-Fleming kind of rules. So, the final analysis is probably so close to .05 it's not worth arguing about.

DR. SCHILSKY: Dr. Lamborn?

DR. LAMBORN: I'm just curious. Did you in fact look to see whether in study 2, using response as a time-dependent variable, you saw any relationship of response to survival?

DR. LEE: No, we did not specifically look at that. If I understand your question correctly, your question is whether responders have longer duration of survival?

DR. LAMBORN: With the appropriate analysis.

DR. LEE: No. Having read a lot of the recent publications saying that that is not appropriate, we did

not do that analysis.

DR. SCHILSKY: We'll have Dr. Raghavan ask the last question.

DR. RAGHAVAN: I'd like to ask Dr. Winer. I'd be concerned to do down a good drug that's less cardiotoxic, but to come back to the issue of study 2, when I look at the company's submission, table 13, page 39, you made the point that you thought that because response and time to progression and everything were identical, that therefore it implied that the change we were seeing in survival had to do with post hoc effects. And I don't want to make too much of this because the company has done under-powered trial.

But you've got 0 versus 2 percent CRs. You've got 32 percent immediate progression versus 27 percent progression. You've got 34 percent versus 39 percent stable. I don't know that that's inconsistent with a difference in survival based on worse efficacy. I really do understand that the numbers are small, but you seem very sure that all the other indices didn't fit with it. Can you just expand on that based on the data that the company has put in? Because it seems to me consistent with a survival deficit. It's really only the 26 percent partial response versus 24 percent that's the same.

DR. WINER: I think if we applied any

statistical test to these numbers, they would all be very comparable. These are all extremely small differences.

Again, just to come back for a second to this survival difference, I don't want to mislead anyone. I remain concerned about this. I think it's the single most important issue here. Because of the benefits of the drug, I feel that that outweighs this particular concern.

But just in terms of survival in metastatic breast cancer, it was just a few years ago that all of us would stand up and give talks and say that there was no evidence that anything we did affected survival in patients with metastatic breast cancer. Now, we all believed there was some impact, but it was impossible to show it in trials.

I know of three trials recently either published or presented that show survival differences: the Herceptin trials where there were clearly differences in all efficacy parameters; the docetaxel versus mitomycin, vinblastine trial published in the JCO this year as second-line therapy where there was a very dramatic difference in both response and survival; and finally, there was an Australian trial comparing CMFP versus Taxol where in fact there was no difference in response but there was a difference in survival. But in that trial, patients who received CMFP never received Taxol. So, the Taxol patients

were receiving an extra treatment.

So, again I don't want to mislead. I can't be certain that there is no difference, but given all of this information and the other studies, that's why I have arrived to the view that I have. I think these are all very good questions.

DR. SCHILSKY: Thank you very much.

DR. LEE: Dr. Schilsky?

DR. SCHILSKY: Yes.

DR. LEE: If I may because just now there was a discussion about the use of relative risk, and we would be --

DR. SCHILSKY: If you can be very brief.

DR. LEE: Yes. I'd like to call upon Dr. Gary Koch to perhaps give some statistical perspective on the interpretation of the two statistical approaches: one that's based on absolute difference in response rates versus the relative risk approach, particularly for study 2.

DR. KOCH: Gary Koch, a statistical consultant to The Liposome Company, and I'm with the University of North Carolina.

What has been emphasized in this discussion is excluding originally a 15 percent difference in response rate, but the studies actually showed with 95 percent

confidence intervals that you could exclude a 10 percent difference. Now, it was mentioned a few minutes ago that one could look at relative risk and that one would like the ratio of rates to be at least .75 or to exclude a 25 percent difference in whatever the response rate is.

Now, the response rate in study 2 was 25 percent, and 25 percent of 25 percent is 6 and a quarter percent. So, if you were designing the study to demonstrate exclusion of a 25 percent difference when the response rate was 25 percent, you would need a sample size of over 500 patients per group.

Now, you would be actually running 500 patients per group in a situation where you've already seen what the differences are in cardiotoxicity. You'd be producing p values that would have eight or nine 0's in front of a 1 for cardiotoxicity in order to have just a bit more comfort level with relative risk for a response rate that's down around 25 percent.

Now, usually when you're dealing with proportions that vary between 25 percent and 75 percent, almost always statisticians emphasize the response difference, and a response difference of less than 10 percent is very compelling. Relative risk is appealing when you're talking about rates of 2 percent or 5 percent and you talk about doubling the rate, but the risk ratio is

really not a very sound measure when you're talking about rates that are anywhere from 25 percent to 75 percent.

DR. WILLIAMS: Could I ask Dr. Koch a question? If what we're really interested in is preserving the antitumor efficacy of a drug in whatever way it's going to be used, is it not reasonable to consider what percentage of that efficacy one might lose? Certainly that was our thinking. And one must realize that in the combination studies, we're not looking at all doxorubicin efficacy. So, what you're saying is if 10 is okay with a response rate of 25, then 40 percent loss of the doxorubicin efficacy is not of concern.

DR. KOCH: Well, what I'm saying is that if you're going to talk about the difference in efficacy, it should be in terms of a straight percent. If you're looking at Kaplan-Meier curves, you should say the Kaplan-Meier curves never get farther apart than 10 percent in a confidence interval. If you're looking at response rates, then you're saying that they never get farther than 10 percent.

Now, if you're talking about death rates from treating patients with myocardial infarction and the control treatment has a rate of 9 percent and the test treatment can reduce that by half to 4 and a half percent, then it's useful to talk about risk ratios for that. But

when you're talking about rates that are between 25 percent and 80 percent, you should simply say how close do you want them to be. Do you want them to be within 5 percent, in which case you'll need sample sizes of over 500 per group, or are you satisfied that they're within 10 percent?

DR. WILLIAMS: I would agree with you except that we're using this as a surrogate for survival, and in that case we're talking about the percentage of efficacy that one might think is retained and thinking what that might do in a large survival study.

DR. KOCH: Well, I understand that, but then if you basically say that the sponsor not only has to provide a comfort level on response rate, they need to also provide a comfort level on time to progression and on overall survival. When you put forward a criterion on hazard ratio that you would like the lower limit of the confidence interval for the hazard ratio to be above a number, like .75, if the hazard ratio is above .75, the Kaplan-Meier curves never get farther apart than 10 percent.

So, you're basically requiring sponsors to get multiple wins. If a sponsor in study 1 and study 3 provide compelling results with respect to response rate, time to progression, and overall survival, that should be fairly reassuring. If in study 2 they do it for response rate and time to progression, the sponsor is demonstrating effect 8

out of 9 times. Even for overall survival they do it for the first year, and it's not until you're dealing with time after virtually all patients have had a progression -- that's one of the things you need to remember in study 2. Virtually all the patients have a progression by 1 year. So, what happens to them after 1 year is related to those treatments that they're getting after they've had a progression.

So, the sponsor has assured you with respect to antitumor efficacy that they fulfilled the criterion 8 times out of 9, and on the 9th one, they're not that far away except for the phenomenon that happens after 1 year for unexplainable reasons.

DR. WILLIAMS: But the criteria that they fulfilled are their criteria, not the ones that the FDA specified in 1994.

DR. SCHILSKY: Well, I'm sure we'll have additional discussion about this. So, why don't we take a brief break and reconvene at 3:45.

(Recess.)

DR. SCHILSKY: Dr. Cortazar will give the FDA presentation.

DR. CORTAZAR: Thank you, Dr. Schilsky, members of the advisory committee, colleagues, ladies and gentlemen. I'm going to present the FDA analysis of TLC

D-99. The indication under consideration is for first-line treatment of metastatic breast cancer.

I would first like to acknowledge all the members of the TLC review team.

I do not want to repeat things that you have already heard. I will try to emphasize the FDA's perspective in areas where there may be differences.

This slide outlines the critical times of interaction between the applicant and the FDA. I'm going to start with the end of phase II meeting.

Ddoxorubicin-based chemotherapy, which is considered the gold standard for the first-line treatment of metastatic breast cancer, is believed to convey a survival benefit of approximately 6 months. For the application of new drugs for this indication, FDA requests information about survival.

However, TLC, which is a liposomal doxorubicin, is a special case. It has the same active molecule as doxorubicin. In this case, FDA indicated that response rate would be an appropriate endpoint to demonstrate effectiveness when comparing TLC and doxorubicin.

The sponsor planned to power the studies to detect an absolute 15 percent increment in response rate and estimated the overall response rate would be 60 percent. The agency noted that the sponsor's plan would

not be appropriate because if the response rate for the doxorubicin-containing arm was substantially lower than 60 percent, the study would be under-powered. For example, if the actual response rate was 26 percent, as noted in study 2, this absolute increment of 15 percent will represent well over half of the doxorubicin response rate.

Also, FDA stated that the sponsor would need to demonstrate that the response rate of TLC is at least 75 percent of the response rate of doxorubicin by using a one-sided confidence interval.

These standards are less stringent than often required in equivalence trials but are appropriate because TLC was supposed to be less cardiotoxic than doxorubicin.

So, in summary, the agency would consider non-inferiority to be proven satisfactorily if in both trials the lower bound of the one-sided 95 percent confidence interval excludes 0.75 where R equals the ratio of the TLC response rate to the doxorubicin response rate.

During the June 30th meeting, the sponsor proposed to formally close the single agent trial after enrolling 224 patients. They stated that the accrual was very slow and that the results, after the third interim analysis, showed the study met the efficacy and safety endpoints. FDA agreed they had documented less cardiotoxicity with TLC D-99 but did not agree with their

claim of comparable efficacy. The FDA expressed concern about the survival curves for the single agent study which appeared consistently better for doxorubicin and the difference approached statistical significance. FDA recommended a confirmatory trial since the single agent trial appeared to show that TLC D-99 may be inferior to doxorubicin in its antitumor effect, especially survival, and wasn't likely to show non-inferiority.

The sponsor suggested as a supportive trial a randomized study of TLC D-99 versus epirubicin 75 milligrams per meter squared in combination with cyclophosphamide. The FDA questioned the epirubicin dose in the control arm as being too low and suggested that a dose of 100 to 120 milligrams per meter squared might be more appropriate. The FDA expressed the concern that comparison to epirubicin at this dose would not be interpretable in a regulatory context unless one could establish that epirubicin and doxorubicin are equivalent on a milligram-per-milligram basis.

On December 14, 1998, the NDA was submitted.

Again, I will not repeat what you have already heard from the applicant's presentation, just to remind you study 1 is a combination study, study 2 is the single agent study, and study 3 is the combination trial versus epirubicin.

These trials shared the following design features. They were phase III, multi-center, randomized, parallel, open-label trials. All of them enrolled metastatic breast cancer patients with no prior chemotherapy for metastatic disease, and patients had measurable or evaluable disease.

I would like to point out the differences in the trial design. Studies 1 and 3 are combination trials. Study 2 is a single agent trial. Study 3 was stopped prematurely after enrolling 160 patients.

Doses of TLC D-99 are lower on study 1, 60 milligrams per meter squared. They are higher in study 2 and study 3, 75 milligrams per meter squared. Only study 2 allowed dose escalations of TLC D-99 and doxorubicin.

Response rate was the protocol-specified primary endpoint for the three trials, while cardiotoxicity was not an endpoint for study 3.

Study 3 was not conducted under the U.S. IND, so the plans for statistical analysis were not reviewed with FDA. This was a European trial submitted to the FDA as a supportive trial for the NDA.

We believe the dose of epirubicin used in study 3 has not been established as equivalent to doxorubicin for the following reasons. Doses of epirubicin of 100 to 120 milligrams per meter squared have been used in combination

therapy and were more effective than the lower doses. The trials that supported the approval of epirubicin for adjuvant breast cancer compared doses of 50 versus 100 in combination therapy, and the epirubicin doses of 100 were associated with statistically significant improvement in both disease-free survival and overall survival.

In addition, trials in metastatic breast cancer comparing epirubicin doses of 50 versus 100 in combination therapy have shown a statistically significant improvement in response rate and time to progression and a nonstatistically significantly longer survival for the higher dose.

Epirubicin at 75 milligrams per meter squared has not been proven a standard treatment for first-line metastatic breast cancer. A recent ODAC considered an application for epirubicin as first-line treatment for metastatic breast cancer and even at higher doses of epirubicin, 100 milligrams per meter squared, there wasn't sufficient evidence for approval.

We believe epirubicin has not been demonstrated to be equivalent to doxorubicin on a milligram-per-milligram basis. The MTD of doxorubicin is about 90 milligrams per meter squared while the MTD of epirubicin is in the range of 150 to 180 milligrams per meter squared. Also, the cumulative recommended doxorubicin dose is lower

than the cumulative epirubicin dose, 450 to 500 milligrams per meter squared for doxorubicin and 900 milligrams per meter squared for epirubicin.

Therefore, we believe epirubicin at 75 milligrams per meter squared in combination with cyclophosphamide is not the best comparator to TLC D-99 75 milligrams per meter squared with cyclophosphamide because this dose of epirubicin may be suboptimal.

Of the two large studies comparing epirubicin to doxorubicin in combination, each at 50 milligrams per meter squared, a significant difference in survival was noted in one, the French multi-center trial. This survival curve from their report in JCO suggests there is a sustained difference in the survival curves. After covariate adjustments, the author stated that this difference was no longer statistically significant, but none of the details of the adjustment analysis were provided. This finding adds to the doubts we have that epirubicin and doxorubicin are equivalent on a milligram-per-milligram basis.

We do not believe the applicant has adequately established the non-inferiority of TLC to doxorubicin. Although the applicant's studies were powered assuming a response rate of 60 percent, the response rates observed for studies 1 and 2 were well under the expected 60

percent.

Although the agency stated it would use an odds ratio for the efficacy comparisons, we are presenting the relative risk, which is the ratio of the two response rates. This concept is easier to understand and is slightly less conservative than the odds ratio.

Using the ratio of response rates approach, only the combination trial demonstrates marginal equivalence, with a lower bound of the confidence interval of 0.78. The study comparing single drugs fails with the lower bound of the confidence interval of 0.62. The one-sided 95 percent confidence interval could not exclude in study 2 that the TLC arm was proportionally 38 percent worse than doxorubicin.

This slide summarizes the overall survival for studies 1 and 2. Survival was the secondary endpoint in both trials. You may note some subtle differences between the FDA and the applicant's survival analysis. The applicant submitted an updated survival analysis post-NDA submission. This update was not planned in the original protocol, and for this reason FDA decided to use the original submission data. The results of the applicant's updated analysis were similar to the original submission.

The differences of the median survival in study 1 are not reflective of the whole curve. As you will see

in the next slide, the curves cross. The hazard ratio was 1, indicative of similar overall survival. The 95 percent confidence interval lower bound for the hazard ratio was 0.71. However, demonstration of non-inferiority for survival was not explicitly discussed with the agency.

The median survival in study 2 was 5.5 months longer for the doxorubicin arm, with a p value of 0.07, and the 95 percent confidence interval lower bound for the hazard ratio of 0.54.

This slide shows the survival curves for the two treatment arms in study 1. Again, these curves cross at about 15 months. The hazard ratio is 1 and indicates no difference in overall survival.

This slide shows the survival curves for study 2. This near significant survival trend in favor of the doxorubicin arm is very concerning. The applicant argues that this trend might be due to an imbalance in baseline prognostic factors at the largest study center. The applicant excluded that center and reanalyzed the survival. FDA believes this approach is totally unjustified because the randomization was done centrally and there's no evidence of an imbalance for these prognostic factors for the whole patient population.

FDA selected prognostic factors based on a literature review and performed a multivariate analysis to

correct for imbalances in important prognostic factors. This slide shows the results of the multivariate analysis. You can see it did not correct the adverse trend in survival and in fact the finding became more convincing with a p value of 0.03 in favor of the doxorubicin arm.

FDA then also included progesterone and estrogen receptors, the two covariates that the applicant used to show the imbalance, and the analysis is still marginally significant in favor of the doxorubicin arm, with a p value of 0.05.

In evaluation of the case report forms and electronic data of tumor measurements, FDA detected a number of progression events in the applicant's original analysis from patients who died many months or years after the last formal evaluation for progression.

Therefore, FDA requested the applicant to reanalyze TTP. The purpose of this analysis was to exclude inappropriate late events that have had inadequate follow-up and to include legitimate early events. This slide shows the results of their reanalysis. The studies were not designed to show formal equivalence. The 95 percent confidence interval lower bound was 0.81 for study 1 and 0.66 for study 2.

I want to remind you study 3 is a combination trial comparing TLC to epirubicin, each at 75 milligrams

per meter squared. This trial was submitted as a second trial to demonstrate the efficacy of TLC D-99 in combination with cyclophosphamide. The trial was performed in Europe, outside of the U.S. IND, was stopped prematurely, and accrued 160 patients out of the 280 patients planned.

The overall response rates were 46 percent in the TLC arm and 39 percent in the epirubicin arm. The two rates were not statistically different with a p value of 0.4. The ratio of response rates is 1.19. The associated one-sided 95 percent confidence interval is 0.81.

Because this study was stopped prematurely, estimates of efficacy and confidence intervals must be viewed with skepticism. Moreover, unless one accepts epirubicin at this dose to be an established first-line treatment of metastatic breast cancer, these results are uninterpretable. The issue with this trial is not the results but the adequacy of the comparator arm.

The median survival was 18.5 months for TLC and 16 months for the control arm, with a log rank test p value of 0.35. The two-sided lower bound for the hazard ratio was 0.79.

As requested by FDA, the applicant reanalyzed TTP. TLC showed a trend toward a longer time to progression compared to the epirubicin arm.

In summary, study 3 showed similar response rates and a similar survival and a trend toward longer time to progression for TLC D-99 compared to epirubicin. However, problems with this study include a relative low dose of epirubicin was used, 75 milligrams per meter squared, instead of 100 to 120 milligrams per meter squared. This study does not test efficacy of TLC D-99 in combination against a proven dose of epirubicin and, therefore, does not independently substantiate the results of study 1. Furthermore, the statistical findings may be viewed with skepticism since this was a small study performed outside of the IND and was stopped prematurely. The results of this trial are not bad. However, FDA has doubts that the dose of epirubicin used in this trial can be used as a surrogate for the same dose of doxorubicin.

Cardiotoxicity was the primary safety endpoint for studies 1 and 2. It was evaluated with serial month dated measurements of left ventricular ejection fraction by MUGA scans and by the clinical evaluation of congestive heart failure. As you can see, TLC is statistically significantly less cardiotoxic than doxorubicin as measured by protocol-defined cardiac events: 6 percent for TLC, 21 percent for doxorubicin in study 1; 17 percent for TLC, 36 percent for doxorubicin in the single agent trial.

Also, there were more congestive heart failure

events in the doxorubicin arm for both studies, and the median time to cardiac event was significantly higher for TLC in both studies.

This is a very busy slide. I don't expect you to see all the details. This slide summarizes the toxicities observed in the three trials.

Neutropenic fever was similar between treatment arms on both studies 1 and 2. On study 3, neutropenic fever was higher in the TLC arm than the epirubicin arm with a difference that was near significance.

Significantly more thrombocytopenia for the TLC arm than epirubicin on study 3.

Stomatitis was significantly higher in the doxorubicin arm in studies 1 and 2. Again, significantly higher for the TLC arm than the epirubicin arm in study 3.

Diarrhea was higher in the doxorubicin arm in study 1, with a significant incidence in study 2.

FDA and the applicant agreed on the reported values for the cardiac toxicity primary endpoint. Non-inferiority in response rate of TLC compared to doxorubicin has not been established. Response rate estimates were similar but because of the study size, the confidence interval of the ratio of response rates included a lower bound well under 0.75 in study 2. Using the ratio of response rates approach, only the combination trial

demonstrates non-inferiority with a lower bound confidence interval of 0.78.

Overall survival was similar for the combination study with a p value greater than 0.9, hazard ratio of 1, and a 95 percent lower bound confidence interval of 0.71. The median survival on study 2 was 5.5 months longer for the doxorubicin arm, with a p value of 0.07. The applicant suggests the survival benefit in the doxorubicin arm is attributable to an imbalance in prognostic factors favoring doxorubicin. However, the FDA performed a multivariate analysis adjusting for prognostic factors, and instead of correcting the adverse trend, the adverse findings were strengthened, with a p value of 0.03.

The comparator arm on study 3 is not adequate because the epirubicin dose was suboptimal. Epirubicin at 75 milligrams per meter squared has not been established as first-line treatment in metastatic breast cancer.

The cardiotoxicity data looks very promising, and it is very disappointing that the studies are too small and did not have the power to show equivalence in antitumor efficacy. However, the FDA review team believes there is insufficient evidence to support TLC D-99 for first-line treatment of metastatic breast cancer. Using the ratio of response rates approach, only the combination trial, study 1, demonstrates non-inferiority to doxorubicin with a lower

bound confidence interval of 0.78. Even though comparison with standard therapy in study 1 is marginally persuasive, the findings are not replicated in study 2 with a negative survival trend, and study 3 cannot be interpreted in this context. The applicant has demonstrated that TLC D-99 is less cardiotoxic than doxorubicin, but this endpoint alone does not support the proposed indication.

DR. SCHILSKY: Thank you very much.

Are there questions for FDA?

Let me start off with a question while people are getting their thoughts together. One of the important issues I think is the whole issue of the epirubicin dose in study 3. Somehow, as you were going through your discussion of epirubicin, I had the sense that you were looking at the same literature as the sponsor but coming up with totally different results. You showed us a survival curve of what I thought you said was the French randomized study comparing FAC and FEC, and you indicated that there was a survival advantage I think you said for the FAC in that study. I believe that's the same study that was included on one of the sponsor's slides, and at least on that slide, the median survival results are said to be not significantly different in the two arms of the study.

You then I think said that there were a number of other studies, looking at various different doses of

epirubicin in combination, which appeared to show some sort of a dose-response relationship, with higher doses producing better outcomes. Yet, when I go back to the sponsor's slide in which they summarized a number of studies looking at different doses of epirubicin in combination, they did show the higher epirubicin doses associated with higher response rates, but none of those studies showed an advantage in survival for the higher epirubicin dose.

So, maybe you could just clarify again how it is that you've come to the conclusion that epirubicin at the doses used is not an effective therapy.

DR. WILLIAMS: I probably did the more recent review of epirubicin, as I was involved in the recent approval -- or recent evaluation. I've reviewed these literature in detail, especially after a recent interaction with the sponsor.

Now, remember, it's not our responsibility to show that they're not equivalent. It is the sponsor's responsibility to show us they are equivalent. We have shown a lot of data that there appears to be a different antitumor effect between 0 and 100. If we don't have studies large enough to show a survival effect, that isn't our responsibility.

I was shocked when I found this survival curve

within the article in JCO for 1988 because the abstract says it was not significant, but the actual text says it was significant. But we noticed the difference in number of sites and we adjusted and it's no longer significant. No p values are given before or after adjustment. It's amazing to me. So, that's where that comes from, and that's one of the two large studies comparing the 50 milligrams per meter squared dose in combination.

Again, that's not as good a test, I think, of equivalence if it had been 75 milligrams per meter squared because the more efficacy there is, the easier it is to show a difference.

Now, you talk about the dose-response trial that the sponsor cited. Again, it's under-powered so you can't say anything definitely, but in the point values for the tumor effects, both response and time to progression, you see a change between 60 and 90. Now, again, I don't know if 75 is above or below the plateau or if there is a plateau, but that isn't our responsibility. If we're going to use something as a surrogate for equivalence, then it needs to be proven to us, and it apparently wasn't proven to the sponsor of epirubicin well enough that they thought they could submit it to us and get an approval on that basis.

DR. SCHILSKY: I'm not trying to get to the

issue of whose responsibility it is to prove what. I'm just trying to understand why it appears that we have the two groups referring to the same literature and showing us apparently opposite results. I guess part of it is selective reading of certain parts of the literature.

Questions from other committee members? Dr. Krook?

DR. KROOK: I go back to the administrative history here. There was a meeting on June 30th, 1998. Study 2 was closed. It seems that most of us are looking at this and there was obviously a decision made, in combination with the FDA, to close the study early. Now we're looking at survival differences or whatever. I'm wondering if we could go back to there and why was the decision made at that point to close this. I suspect that might have affected the survival, less patients, and the numbers.

DR. WILLIAMS: Again, it might be appropriate if I answer since I was involved in those decisions.

Both from the minutes and from my recollection, we never advised to close the study. The company wanted to close the study. It seemed like the main reason was they were having accrual problems. We were concerned. So, we said you could add Zinecard to the other arm of the study if you're concerned that investigators don't think it's

ethical to continue. But we never advised them to stop. We stated if you want this to stop, that is your decision, but we're concerned about these values. So, I don't think there's any disagreement.

DR. KROOK: Again, I'm not a statistician. If we had the full complement of the study, would we do better statistically since that's where we seem to look at things? I think the answer is yes, but maybe Dr. Lamborn can comment. We obviously didn't go to the planned study accrual. Therefore, the statistical reasons change and the confidence intervals change, if I'm right.

DR. LAMBORN: As I understand your question, obviously if you have a larger pool of information, your confidence intervals are going to be narrow. Whether or not this would have affected the direction, for example, of the survival, there's no way of knowing. There's no reason to believe, as Dr. Piantadosi mentioned before, that an early stop that was not a basis by looking at something is going to affect that. But you definitely narrow your confidence intervals.

DR. SCHILSKY: Dr. Kelsen.

DR. KELSEN: Could you put back up the survival curve of FAC versus FEC? When you looked at this curve, it sort of looks a little bit like the curve from the second study that we're concerned about because it looks like the

survivals are absolutely overlapping for the first 200 days. I guess that's in days. So, that's about two-thirds of a year. And then they diverge and they don't come back together again, although they don't reach statistical significance. That looks an awful lot like the curve that you showed us or that we've seen from the sponsor for the phase II. Is your interpretation saying there's a salvage therapy difference? Is there any data in that paper as to a salvage therapy difference?

DR. WILLIAMS: There's a paucity of data to totally understand the study, and that's why I don't trust these results one way or the other. I think they raise doubt.

But you're not going to see a difference in curves unless you have deaths, and the fact is with breast cancer, it takes a longer time than it does with colon cancer or lung cancer. If you look how far you've come down on the survival curve before you see a difference, it looks like to me it may be 75 percent.

DR. KELSEN: I was struck. We had this real long discussion before as to whether or not there was a difference in efficacy of the liposomal preparation and the parent analog. Well, for the first 3 to 6 months or a year maybe, they're identical. But I'm struck that it almost looks like the same --

DR. WILLIAMS: But this is always going to occur with first-line breast cancer treatment.

DR. KELSEN: Yes.

DR. WILLIAMS: And you're always going to have a later part in the curve. So, if you ever have an effect on the curve, it's going to be sometime after you stopped your first treatment.

DR. SCHILSKY: Dr. Lippman?

DR. LIPPMAN: To clarify, I think, Grant, you were the one that mentioned this. Initially we got into this discussion with Dr. Koch about absolute versus relative changes in response rate and there's no way to win type of situation. But I think you were the one who said this, that initially when you target a 60 percent response rate, then a 10 percent absolute difference means something different entirely than if, unfortunately, your response rates are lower at 20 percent. So, obviously, the absolute difference means different things based on the overall response rate I guess is the point.

DR. WILLIAMS: You've got to remember that we were doing something different and using response rate for first-line approval of breast cancer here. This response rate is because we have the same drug, same molecule, and we're using it as a surrogate of what we think that effect is going to have on ultimate survival years down the line.

In that circumstance, we're interested in how much of the antitumor or of the beneficial effect imparted by doxorubicin are we losing, and I don't believe we ever conceived of using a response rate down in the 20's range to demonstrate equivalence for that effect. It was theorized that the response rate would be higher. As mentioned earlier, if it was going to be lower, it wouldn't have been practical to say that a ratio of .75 would be your goal. If you thought it was that low, then you probably should have designed the endpoint to be survival, but with the combination arm, the response rate turned out to be up in a higher range where it was reasonable to do equivalence trials.

So, what really happened was the response rate was much lower than planned. The study was much under-powered for this outcome, and we don't have the study showing equivalence that we wanted.

DR. SCHILSKY: Grant, the issue of using response rate as a surrogate to predict survival years down the road -- I guess one of my concerns about that is that, if we are willing to accept some reduction in response rate for this agent in exchange for the benefits with respect to cardiac toxicity, can we assume that reduction in response rate actually will translate into some decrement in survival down the road, and is it a one-to-one

relationship? I'm not comfortable with accepting that as -- you may say the response rate is a surrogate for benefit, but I'm not sure in my own mind that one can accept the notion that some change in response rate is a reliable predictor of a similar change in survival.

DR. WILLIAMS: Right, and that's why we asked for a demonstration of non-inferiority. If you're going to start extrapolating and say, well, I think maybe we've got 5 percent difference in response, then that translates. No, I don't think anybody could do that.

The responsibility of the new drug is to show that you have efficacy, and in this setting a non-inferiority design was chosen to show that you have the same efficacy as doxorubicin and response rate was chosen as a surrogate. I do believe that's debatable, but that's the agreement.

DR. SCHILSKY: Let me ask one other question to the committee. It has to do with the level of concern about the survival data in study 2 in the context of the proposed indication of using this drug in combination with cyclophosphamide because what we seem to have is study 1 which there seems to be a reasonable level of comfort with the notion that that may demonstrate equivalence in combination with cyclophosphamide. We have study 2 where there's concern about inferior survival in a single agent

comparison. We have study 3, which at best may be difficult to interpret, and we have a proposed indication not to use the drug as a single agent, but to use it together with cyclophosphamide.

I guess I'm just wondering if those people on the committee who have been concerned about the potential survival decrement in study 2 are equally concerned in the context of the proposed indication.

Dr. Raghavan.

DR. RAGHAVAN: I think that what this illustrates is that there are no shortcut paths to drug development, and if you have under-powered studies, you get to the FDA and you get into trouble.

Now, as was expressed by Dr. Williams, the third study, while it is sort of comforting in a general sense, in reality isn't comforting at all for all the reasons that were enunciated. It's almost valueless data to us because of the lack of controls and the early stopping and sort of breaking all the rules of good trial design.

So, that then leaves us with two trials, one of which, any way you cut it or slice it, gives the drug what appears to me to be the potential for inferior anticancer effect. That then leaves you with one study, which isn't a very big study, which leaves you with a comfortable feeling

that this is a much less cardiotoxic drug with equivalent activity.

The problem is what do you do with that. I'm persuaded by investigators that I really respect representing the company that this is a less cardiotoxic drug and they would like to use it regularly, single agent or whatever. Yet, I look at the data and I'm very troubled by study 2. Dr. Koch's statistical discussion, while I accept his points about the 75 percent lower limit, the rest of it I really found didn't address the issue very well.

So, I think we're left with a situation of almost equipoise. How much concern do we have that Adriamycin in an unrestrained environment will poison hearts? I was struck by the lady who spoke about here difficulties with cardiac toxicities right at the beginning, and I'm very sympathetic to that. On the other hand, I was also struck by the fact that she said, but I'm happy to be here.

So, I think when people look at us, if we're being harsh on this drug, I don't think we're being bureaucrats and being persnickety about trials. I think we're actually looking at a drug and saying is there evidence that to cut down toxicity, which can be avoided other ways, we're not sacrificing cure rates or response

rates. And we know that this will get translated down the line into the adjuvant setting.

So, I think we have the problem of just not enough information to be really comfortable.

DR. SCHILSKY: Dr. Margolin?

DR. MARGOLIN: I'd like to echo that and just sort of add a corollary which is that even though, according to data we were given today and what we all sort of know from being in the clinics, that the most popular combination with doxorubicin and its buddies is with cyclophosphamide. Things are changing fast in breast cancer. We're learning about how one can combine with various drugs. New drugs are coming up all the time. And if we have something for which we can only feel comfortable with the preservation of activity in this very strict setting of co-administration of these drugs, I think we're going to be stuck with more of a problem than a solution.

DR. SCHILSKY: Ms. Zook-Fischler?

MS. ZOOK-FISCHLER: Well, as a patient, I'm coming from a somewhat different perspective, and I'm a patient in treatment at the moment. My first concern for myself and other women dealing with breast cancer is always survival as the bottom line.

Nevertheless, it seems to me, for those particular patients for whom cardiotoxicity would be a

problem, it becomes a moot point because if you can't give them the full dose of doxorubicin, it would be good for the physician and the patient to have an option because they may not be able to get the survival advantage that you're speaking about with the doxorubicin. So, I personally from a patient's point of view, even though I'm most concerned about survival, I would like the option to be available to patients for whom the other drug is not really the best choice.

DR. SCHILSKY: Dr. Lamborn.

DR. LAMBORN: I'd just like to make a comment about study 3 from a statistical standpoint. I understand the issue of whether the epirubicin is a reasonable control, and I don't want to address that because I think that could stop the thing right there.

But if the issue is the fact that the study was stopped early, I do want to reiterate that the early stopping, if it was not done because of the effect, does not preclude looking at those confidence intervals and deciding if they are useful. So, just to make sure that we're making the decision on the right pieces of information.

DR. WILLIAMS: As a regulatory comment, as a reviewer I'm not sure that I can ever with certainty say why a study was stopped.

DR. SCHILSKY: Let me just bring us back to are there any other questions to be directed to FDA, because otherwise we're drifting into discussion again. Dr. Lippman.

DR. LIPPMAN: I just wondered if the FDA had looked at this type of subset analysis that was presented by the sponsor in slide 90 on the subset of women that were at high risk for cardiotoxicity, because the issue was just raised about women that are at high risk and this would be a good option. In this one, both studies 1 and 2 were lumped together and the survivals are equivalent. Has study 2 been looked at in this group?

DR. WILLIAMS: Well, I'll make a couple comments. First, I want to remind Ms. Fischler that there is a drug which is approved, which is Zinecard, which is available to be given after 300 per meter squared, which is the dose after which most people get congestive heart failure and cardiac toxicity. It is available. So, I can't see that this would be that helpful in terms of fulfilling a need after 300 per meter squared.

Regarding the post hoc analysis of patients at high risk, I really don't think that is quite the issue. First of all, I believe the data is under-powered for equivalence overall. It's certainly going to be under-powered for equivalence in any subset.

Secondly, that isn't the question. We believe there's a cardiac benefit. There may be a lot more in these patients, but we're dealing with the fundamental issue of whether we know there's equivalent efficacy. So, I think in the presence of an approved agent that should prevent or certainly decrease the number of events, if applied after 300 per meter squared, I don't see how that analysis is really relevant beyond the main question you have which is, is the benefit here worth the doubt regarding equivalent efficacy?

DR. SCHILSKY: Dr. Lamborn?

DR. LAMBORN: Dr. Williams, I just wanted to check that I had understood something that you said a little while ago. I thought I heard you say that if in fact you had anticipated a response rate on the order of 25 percent, that you would not have been comfortable with using response as a surrogate for survival in demonstrating equivalence. Is that what you said?

DR. WILLIAMS: I don't think it would have been practical, first of all, to power a study. We were comfortable with the concept that in doing an equivalence or a non-inferiority study, you need to rule out loss of not more than a certain amount of the efficacy of the drug. In this case, .75 was below the .8 that is often discussed. Powering a study to demonstrate that with such a low

response rate would probably lead you to a similar size study for showing equivalence in more ultimate endpoints such as survival.

DR. SCHILSKY: Any other questions for the FDA? Dr. Beitz?

DR. BEITZ: I just wanted to point out that if the committee does feel the need for additional information, that perhaps in your deliberations you could advise as to the nature of additional studies that could be performed with this compound to establish efficacy in relevant populations.

DR. SCHILSKY: Dr. Johnson?

DR. DAVID JOHNSON: Well, actually I want to discuss that too.

The question I wanted to pose to the FDA is putting myself for the moment into the sponsor's shoes, I'm trying to think of a way that I'm going to prove the comparability of my product, thinking that, as I've been told by my expert colleagues around the country, that it's becoming increasingly difficult to show survival differences based on a single agent because of the availability of other products.

I do a lot of work, of course, in lung cancer, and I hear this a lot, that the reason there wasn't a survival advantage with drug X over drug Y is because of

all the salvage therapy that went on. Now, that's always humorous to me in lung cancer, but it may be actually relevant in breast cancer. The only caveat that I would say about that is that a really good drug seems to be able to overcome that particular issue, Herceptin being a recent example of that.

My sense is this is sort of why you chose the response or at least agreed with the sponsor about using response as an endpoint, which I too like Rich was a bit surprised to see that we had thought that was acceptable. Is that something we need to discuss, or is it only because the product was perceived to be the same thing in just a different package?

DR. WILLIAMS: Of course, this was 1994 that this study was designed. The reason we accepted response was solely because it was two different forms of doxorubicin, and we thought that the antitumor efficacy would hopefully reflect the other.

We've had these discussions. I think you gave both sides of the argument about follow-up therapy. A good drug seems to be able to overcome it, and obviously the advisory committee took the position that it still could be shown and the time to progression wasn't sufficient in first-line breast cancer.

DR. SCHILSKY: If there are no other questions

directly for FDA, before we get into our discussion, we do have two additional requests from the public to speak to the committee. First is Dr. Marissa Weiss representing Living Beyond Breast Cancer.

DR. WEISS: My name is Dr. Marissa Weiss, and I'm a physician. I take care of breast cancer patients. I'm an oncologist. I'm also President and founder of Living Beyond Breast Cancer, which is a nonprofit educational organization. And I'm author of the book, Living Beyond Breast Cancer.

I invited myself here. Liposome did not invite me here. They do buy a table at our gala each year, as every other pharmaceutical company does do, and they do provide an unrestricted grant for our outreach, but so do other companies.

Living Beyond Breast Cancer's mission is to help women who have been diagnosed with breast cancer live as long as possible with the best quality of life. I think everyone is here for the same reason, which is that the whole point of finding breast cancer early and treating it effectively is to give life after treatment is finished. And that's why I'm here because I see Evacet as an opportunity for women who have breast cancer to live beyond their treatment with a good quality of life. And we're not just talking about quality of life. For those people here

who do treat women with breast cancer with Adriamycin-based regimens, you can die of that complication, the cardiotoxicity.

I chose to speak after the data were presented because I wanted to address some of the issues that were raised. I think everyone here knows that we want to have better treatment options to present to patients when they have metastatic disease, what that disease progresses, that that's what the purpose of this drug and other advances that have come forth for women with breast cancer.

I have to say that I've been following this discussion all around, and I'm really struck that everyone is taking the survival differences very seriously, which is terrific. But we are talking about one endpoint in one study that's observed after 1 full year of follow-up that wasn't even statistically significant. By many other criteria, these drugs did look like they were equally effective. We're not talking about proving beyond the effectiveness. We're just talking about equivalence here.

With respect to the comments of if we start using this drug, not just with Cytoxan but then maybe in the adjuvant setting and maybe you're going to start sticking it with Herceptin, that I think is irrelevant to this discussion because that's not what this indication is up for. We're talking about looking at this drug with

Cytoxan as first-line chemotherapy for women with metastatic breast cancer.

I think we have to keep some of this in perspective because Doxil has been approved for use in cancer. That's ovarian cancer not breast cancer, but it associated with the higher side effects of PPE and this drug is not. But probably more importantly, epirubicin was approved in the adjuvant breast cancer setting. However, the bar for efficacy was lower for epirubicin since it was compared to CMF and not to an Adriamycin-based regimen.

Also, if these data showed a trend to improved survival with Evacet over Adriamycin, we wouldn't have a discussion here. I know that the goal here is to prove at least equivalence.

I think that the data presented to ODAC in the spring on epirubicin was with the doses of 100 to 120 milligrams per meter squared, but this is not the standard dose in Europe nor is it the dose used in patients with metastatic breast cancer. It was the dose used in the United States for the adjuvant treatment, but you really can't fault this study for using the 75 milligrams per meter squared dose since it was the standard of care in Europe at that time, even though it was not used at that point in the United States.

So, I just don't want anyone here to lose sight

of the fact that as doctors taking care of people, women with breast cancer, that we need better options to treat the women that we have the privilege of taking care of. When you have a medicine that's working, when you've got somebody in front of you who has got metastatic breast cancer for which an Adriamycin-based regimen is working, you want to keep using the drug that's working. In this situation, twice the number of women who received the Adriamycin-based regimen had to stop their treatment for fear of the cardiac toxicity relative to Evacet.

So, the main thing is I see this as a physician, that this gives me a tool sitting across the table from a women in a blue gown who's got metastatic disease, who's trying to find something that she can extend her life with, I can keep going with the medication that she's already responding to. And I see this drug as giving women an option, and I think it's an important medication that should be approved.

Thank you.

DR. SCHILSKY: Thank you very much.

(Applause.)

DR. SCHILSKY: We had a request from Ms. Meeker who we heard from earlier today to address the committee once again.

MS. MEEKER: Thank you. I want to repeat that

I'm not here subsidized or connected with anybody. I'm here on my own. I'm here because this group was addressing the treatment of metastatic breast cancer, not initially diagnosed breast cancer.

In metastatic breast cancer, treatment may produce no response. The time of a remission is measured in months for all of the metastatic patients. The time to a second recurrence is measured in months. Ultimately, we still measure overall survival of metastatic breast cancer patients in months, not years.

I am really astonished and quite emotional I think because I hear a discussion that seems to me to assume that we're going to be alive for years, so it's very important whether the drug being discussed by you according to your rules -- and I won't attempt to have that discussion -- has months of less effectiveness rather than Adriamycin.

In my case, for example, when I developed my first recurrence -- and there's a mistake on the little thing I typed out this morning at home. I was given CMF because of the metastases to my left brachial plexus. During that time, I developed a second primary. I was not a candidate for anymore Adriamycin-based chemotherapy unless there would have been one that hopefully would not have damaged by heart. There might be one of those in the

future.

I now have metastatic disease. I'm in not a remission but at least only in a state with bone metastases. My future is measured, according to the median at least, in months. I don't believe it and I'm going to fight with all my heart. But what I want to do now is stay alive and healthy until there is something that will address my breast cancer.

Just as a couple of other people have stated, what we need is tools of choice for physicians to prescribe for their patients so that if one chemotherapy is not available or efficacious, there might be another one. In my case, the choices are constrained because of my reaction to the Adriamycin. Unfortunately, I had bad side effects even though I had very positive treatment effects.

It seems to me that whatever the rules for adoption of the drug, if you are down at the third decimal place to the right of the decimal point in discussing relative efficiency of the drugs, that you're describing a drug that might be useful for me and would be the only one that my body could possibly tolerate.

Thank you.

DR. SCHILSKY: Thank you.

(Applause.)

DR. SCHILSKY: I think we can go on to the

questions, and I'm sure they will engender some additional discussion. Yes?

DR. LEE: I know I asked for one. Can I --

DR. SCHILSKY: No, I'm afraid due to our time constraints, we need to go ahead with the questions.

So, we have a series of questions. We have a preamble describing the study results. And I will ask for a formal vote on each of these questions unlike what we did this morning.

Question 1. Do these studies demonstrate that Evacet is significantly less cardiotoxic than doxorubicin at the doses and schedules studied?

Does anyone want to take a crack at answering that one? Dr. Nerenstone?

DR. NERENSTONE: Yes. I think that that's the one thing that we could probably all pretty much agree on.

DR. SCHILSKY: All right. Well, any further discussion on that?

(No response.)

DR. SCHILSKY: Let's see if we do.

(Laughter.)

DR. SCHILSKY: All who would vote yes in response to that question, please raise your hand.

(A show of hands.)

DR. SCHILSKY: 11 yes.

All who would vote no?

(No response.)

DR. SCHILSKY: 0 noes. So, it's unanimous.

Moving on two question 2. Is study 1 an adequate and well-controlled clinical trial demonstrating the efficacy of Evacet in the first-line treatment of breast cancer?

Comments from anyone?

(No response.)

DR. SCHILSKY: No comments. Yes, Ms. Zook-Fischler.

MS. ZOOK-FISCHLER: I just wondered if a word was omitted. Was that to be first-line treatment of metastatic breast cancer?

DR. SCHILSKY: Yes.

Dr. Margolin?

DR. MARGOLIN: Actually I'd like to ask for another clarification which is, does the FDA intend for us not to have Evacet plus Cytoxan in that line?

DR. WILLIAMS: That's a point you can discuss. This was the proposed indication basically and you can make a discussion of that.

DR. SCHILSKY: Well, the proposed indication from the sponsor was Evacet in combination with cyclophosphamide. So, we could amend the wording to say is

study 1 an adequate and well-controlled clinical trial demonstrating the efficacy of Evacet in combination with cyclophosphamide in the first-line treatment of metastatic breast cancer?

DR. WILLIAMS: I think that's fine. That's really more appropriate. The final vote, if one voted that two of these studies were appropriate, then you could have the discussion whether it should be restricted to cyclophosphamide or not.

DR. SCHILSKY: So, let me just restate the question again with these modifications. Is study 1 an adequate and well-controlled clinical trial demonstrating the efficacy of Evacet in combination with cyclophosphamide in the first-line treatment of metastatic breast cancer?

All who would vote yes, please raise your hand.

(A show of hands.)

DR. SCHILSKY: 9 yes.

All those would vote no?

(A show of hands.)

DR. SCHILSKY: 1 no.

Any abstentions?

(A show of hands.)

DR. SCHILSKY: 1 abstention.

So, it is 9 yes, 1 no, 1 abstention.

Question 3. Considering the standards of

efficacy agreed to by the agency for this situation, is study 2 an adequate and well-controlled clinical trial demonstrating the efficacy of Evacet in the first-line treatment of breast cancer?

Maybe I could ask for a clarification here just so we're all clear. Perhaps the agency could restate what were the standards of efficacy the agency agreed to.

DR. WILLIAMS: That was the response rate, .75. That was the standard.

DR. SCHILSKY: So, you're asking us that in considering that standard with respect to response rate as the measure of efficacy, is study an adequate and well-controlled trial demonstrating efficacy for Evacet.

DR. WILLIAMS: We would expect you to take everything into consideration. I guess when it was written, some people might not actually buy the concept that the response rate was a surrogate, but taking in consideration the fact that we did commit to that idea, we wanted you to take that into consideration.

DR. SCHILSKY: So, it might be worthwhile before we vote on this to just refer one last time to the study 2 results that are shown in the table in the questions because this is the single agent study in which the response rates are 26 percent in each arm, the response duration is 7.8 months for Evacet, 6.5 months for

doxorubicin. The overall survival is 14.6 months median for Evacet, 20.1 months median for doxorubicin. That has a p value of 0.07, and the time to progression was 3.8 months for Evacet and 4.3 months for doxorubicin, also with a nonsignificant p value of 0.58.

Dr. Margolin.

DR. MARGOLIN: I think in the interest of fairness to the sponsor, since I think Grant is giving us a little bit the option to decide whether we want to stick by the FDA's very strict definition, it would be interesting to have, if somebody has at hand, what the confidence intervals were on those response rates.

DR. WILLIAMS: I'd like to challenge your statement. I consider our standards, the fact that we used response rates to start with, to be not at all strict and the .75 is actually lower than .8. If you use Dr. Koch's approach, they're lenient.

DR. SCHILSKY: Dr. Kelsen.

DR. KELSEN: If I can pursue that just a little bit. If I read this question right, the standard of efficacy that you agreed to was that the lower bound on the ratio of response rates, since we're talking about that now, would not drop below .75. Is it a correct statement that when you say considering the standards of efficacy that you agreed to, that the lower bound of the response

rate around this 1.0, .62 is below the .75 that you -- it's the ratio in this case that you required when you made this discussion with the sponsor.

DR. WILLIAMS: That's what I'm referring to. The other endpoints, I don't believe they're insignificant, but that was what I was referring to.

DR. SCHILSKY: So, we have that piece of information, and of course, we've been discussing for the last several hours the concern about the apparent decrement in survival, although it doesn't quite make it as a statistically significant difference.

Dr. Lippman?

DR. LIPPMAN: Yes, I just wanted to address that because it came up in one of the presentations recently that it wasn't significant, and we lived by the p value of .05. So, it could be by chance even though we know that there's a biologic continuum. But the concern I think that was consistent between the FDA and the sponsor was that the p value trended more towards significance after adjustment for covariates and was statistically significant in the agency analysis in the covariate adjustment.

DR. SCHILSKY: Dr. Kelsen?

DR. KELSEN: This is sort of a procedure question. The way that's written, how can one say anything

other than no since the way it's written, it says you said that they have to have .75 and it's .62. So, I'm a little lost as to how one --

DR. WILLIAMS: We just asked you to consider. We didn't say you're bound by them.

DR. KELSEN: Okay, thank you.

DR. SCHILSKY: So, I think what we're trying to get to the bottom line here is that this study doesn't appear to make it with respect to the response rate rules that had been agreed to previously, and there is a concerning trend in the survival, although not a statistically significant trend. I just wanted to be sure that everybody remembered the course of the discussion as we go through these questions.

David?

DR. LEE: I would like to provide the lower bound of the confidence limit that was requested by Dr. Margolin. For this study, the lower bound is minus 9 percent based on a difference in response rates.

DR. SCHILSKY: Thank you.

David?

DR. DAVID JOHNSON: Again, my sense has always been that we're advisory. We're not policy makers here. We've heard the fact that the FDA agreed with the sponsor for a particular endpoint doesn't mean we have to agree

with that endpoint. A large part of the discussion that has been ongoing today has dealt with that particular issue. I think that some of the discussion we've had has at least made fairly clear to me that some of us are uncomfortable with that as an endpoint to declare comparability and efficacy. I think this is an issue that we're going to be wrestling with not just this time but in the future as well, and I'm not sure there's an easy solution to it.

Based on the strict definition, if we accept what the FDA has said, then the answer to the first question in my view has to be yes, and the answer to the second question has to be no. There's not even reason for voting. I think that we have to remember that we're only advisory. We're not policy makers here.

DR. SCHILSKY: Dr. Raghavan.

DR. RAGHAVAN: I think the other point that I'd add -- two points. One is that the assumption was made -- and I think it was a mistake post hoc -- that response was an appropriate surrogate on the assumption that survival would be equivalent or potentially equivalent, and that assumption was not fulfilled. We can second guess all the reasons for it. I have yet to hear a compelling explanation that let's me feel that more patients would have been alive in that single-arm study on the new drug

than on the old drug, and dead patients don't get side effects.

The second point I would make is that when we talk about statistical significance, as Scott started to say, driven by a p level, even when we allow for prognostic variable adjustment of a 2 percent chance that we're making a mistake, so I don't know what we're really talking about. It seems to me fairly clear that in study 2 more patients died on the new drug.

DR. SCHILSKY: Dr. Lamborn?

DR. LAMBORN: In a way, I'd like to follow up on the same thing. I think that whenever you pick your primary efficacy measure, you do it on the assumption that the secondary measures do not have a strict p value criteria, but you expect them to be consistent or you worry. In this case I think we're in the situation where the primary efficacy was the surrogate to start with, and the secondary measure is certainly worrisome. So, I think that's the way the p value ought to be looked at, and I think that's what others in the group are saying as well.

DR. SCHILSKY: Although Dr. Johnson felt that we may not actually need to vote, we'll vote anyway. So, just to restate the question, considering the standards of efficacy agreed to by the agency for this situation, is study 2 an adequate and well controlled clinical trial

demonstrating the efficacy of Evacet in the first-line treatment of breast cancer?

All who would vote yes, raise your hand.

(No response.)

DR. SCHILSKY: All who would vote no?

(A show of hands.)

DR. SCHILSKY: 10 no.

Anyone who wishes to abstain?

(A show of hands.)

DR. SCHILSKY: 1 abstention. It's 0 yes, 10 no, 1 abstention.

Question 4. In the first-line treatment of breast cancer, can we assume that efficacy of epirubicin, 75 milligrams per meter squared, is equivalent to that of doxorubicin, 75 milligrams per meter squared, when given in combination with cyclophosphamide, 600 milligrams per meter squared?

Comments?

(No response.)

DR. SCHILSKY: No comments.

Shall we just vote it? So, can we assume the efficacy of epirubicin at 75 is equivalent to that of doxorubicin at 75 in first-line treatment of metastatic breast cancer? All who would vote yes?

(No response.)

DR. SCHILSKY: All who would vote no?

(A show of hands.)

DR. SCHILSKY: 8.

Abstentions?

(A show of hands.)

DR. SCHILSKY: 2 abstentions.

Let me just ask for those who would vote no again, please raise your hand.

(A show of hands.)

DR. SCHILSKY: 8 no.

And abstentions?

(A show of hands.)

DR. SCHILSKY: 3 abstentions.

My right visual field here is obscured by this projector. It's hard to see.

Question 5. Considering the standards of efficacy required by the agency in this situation, is study 3 an adequate and well-controlled clinical trial demonstrating the efficacy of Evacet in the first-line treatment of metastatic breast cancer?

All who would vote yes?

DR. LAMBORN: Can I ask a question first?

DR. SCHILSKY: Yes, please.

DR. LAMBORN: If the statement is that it is not equivalent, demonstration of equivalence has not been

made, which was the previous vote, under what guidelines are we then looking at the question of this as -- I'm not sure that question 5 is relevant if the majority vote was that the comparator was not demonstrated equivalent to doxorubicin and the indication is for -- I'd just like a discussion.

DR. DAVID JOHNSON: Now you're being logical.

(Laughter.)

DR. DAVID JOHNSON: You can't do that.

DR. SCHILSKY: Do you feel you need us to vote on that one?

DR. WILLIAMS: No.

DR. SCHILSKY: All right. No vote required.

Question 6. Do you recommend approval of Evacet in combination with cyclophosphamide for the first-line treatment of metastatic breast cancer?

Discussion on this one?

DR. DAVID JOHNSON: Yes. I'd actually like some discussion on this issue because as I recall the standards that we've adhered to in previous discussions, we've looked for a pivotal trial that we felt confident gave us the results we were looking for and then adequate supporting data to really back that up. Here I think, while I said this quite a long time ago, in my heart I believe -- no pun intended -- that this drug works and

potentially has a role, I don't think the data that have been presented to us are convincing of that. While the first trial the majority felt was adequate and well designed, et cetera, it's hard to imagine that we can vote anything other than no on this in light of the fact that we don't consider the supporting data is sufficient. Now, again, we're an advisory group. I guess we can do what we want to do, but it's difficult for me to understand how we would do that candidly.

DR. SCHILSKY: Other comments? Dr. Raghavan.

DR. RAGHAVAN: The other thing that has been kind of troubling -- well, there are two bits to it. One was that Grant Williams made the point that there actually are alternatives for those who want to continue to use Adriamycin. But there has been almost a tacit assumption -- and I'm unaware of data that support this assumption -- that treating forever is a good thing for patients with metastatic disease and that Adriamycin is the only drug out there.

The reality of the situation is that for the person who's had a major myocardial infarction before she gets her breast cancer and who never has access to Adriamycin, in addition to the fact that there's CMF which sometimes doesn't work, there must be 15 other drugs that can be either approved used or on some form of literature

approval used for breast cancer. So, it's not as if this is the only way that we can get this lifesaving drug for women with breast cancer.

The reason I make this point is that, therefore, I'm not sure that we should lower the bar by comparison with other drugs that we look at. The way the rhetoric has gone today, it sounded a little bit like the bad, old committee is keeping a very, very vital drug from people with breast cancer. I think that the offset of that is that we can make a worse mistake which is to forget that there are good alternatives that are proven and work in second, third, fourth, fifth line and then introduce a drug that may actually have less people alive at a time point, particularly if used as a single agent. I recognize that the approval indication is for in combination, but therein lies a problem in terms of trying to dissect out the various components of response. So, I think we just need to keep in mind what we're actually talking about here.

DR. SCHILSKY: Ms. Zook-Fischler.

MS. ZOOK-FISCHLER: I just have to reiterate what I said before. It seems to me approving this drug doesn't eliminate the physician -- approving this drug would not eliminate the other options. It would just add to the arsenal of options. I think that from my experience with the women I've known, breast cancer is such an

individual disease and each women's response to it on an emotional level, besides the physical level, is very personal. I think the more options out there, the better.

DR. SCHILSKY: Dr. Margolin?

DR. MARGOLIN: I think I agree much more with Dr. Raghavan, but there's a reason which is that the danger we're trying to avoid is getting a drug out there that hasn't been proven not to be inferior and will be used for the wrong reasons in too many people who will then not have been treated right and will miss their best chance to have their best long-term remission.

DR. SCHILSKY: Dr. Krook?

DR. KROOK: A little different opinion. I listened to Dr. Johnson here. I think the pivotal study does sway me towards the fact that this would be a drug -- we may argue whether Cytoxan adds something or not. Most of us perhaps don't use AC in the metastatic situation.

The other two studies -- and again it was brought up, are they supportive or not supportive. I think there's enough, at least in my opinion, in the other two studies to perhaps support it. I can look at the survival and say there's very few people out there that far. So, I guess I would say the pivotal study at least leans me towards saying that this is an option and would be viable and the others don't totally not unsupport it.

DR. SCHILSKY: So, you're prepared to?

DR. KROOK: So, I'm prepared to say that looking at the pivotal study, I see that at least it's similar with efficiency. I'm willing to say, as we've said, that it improves the lack of cardiotoxicity, and that the other two are somewhat supportive, although I see the survival problem.

DR. SCHILSKY: So, you'd be prepared to vote in favor of approval on the strength of one pivotal trial for which the primary efficacy endpoint is response rate.

DR. KROOK: You're going to ask for hands sooner or later.

(Laughter.)

DR. SCHILSKY: Dr. Lippman.

DR. LIPPMAN: I think that the concern is that the pivotal trial is in combination. If the pivotal trial had been single agent, you'd feel comfortable about that drug. I think Dave Johnson brought that up. So, that's one of the issues.

Of course, data in other trials could maybe not be supportive, but the fact that they're going in the wrong direction and very close to statistical worrisome levels is part of the issue. Your comment about the pivotal trial, the concern I have is that it's a combination.

DR. SCHILSKY: Dr. Nerenstone, a comment?

DR. NERENSTONE: I was just going to echo really what Dr. Lippman said. I think as somebody out in the community who is treating these women and you tend to use single agent treatment, the single agent study was not positive and, in fact, survival went in the same direction as the lack of response rate. That is, it didn't meet its primary goal, but also survival was in the wrong direction. So, that study is internally consistent, and that's the problem. So, I have a lot of concerns.

DR. SCHILSKY: I think we've had a good discussion. Why don't we go ahead and vote? I'll restate the question. Do you recommend approval of Evacet in combination with cyclophosphamide for the first-line treatment of metastatic breast cancer?

All who would vote yes, please raise your hand.

(A show of hands.)

DR. SCHILSKY: 2 yes.

All who would vote no?

(A show of hands.)

DR. SCHILSKY: 9 no.

Okay, that concludes today's session. Thank you very much.

DR. BEHRMAN: Dr. Schilsky?

DR. SCHILSKY: Yes.

DR. BEHRMAN: I know it's running late, but

could we get just a little more guidance on this concept of response rate as the endpoint? Is this something that the committee was very unhappy about seeing, mixed feelings, something you'd want to discuss at a later date?

DR. WILLIAMS: Right. If you have a discussion, it should be specifically about, say, liposomal doxorubicin, not all the different possibilities.

DR. BEHRMAN: In the setting where the compound -- where it's the same molecular entity.

DR. SCHILSKY: I think that's an important question. We can take a few minutes to discuss it. Well, Dr. Raghavan, do you want to start off?

DR. RAGHAVAN: I'm very unhappy with that concept.

(Laughter.)

DR. SCHILSKY: Dr. Johnson, anything to add?

DR. DAVID JOHNSON: Well, probably for the first and maybe the only time in my life, I agree with Derek.

(Laughter.)

DR. SCHILSKY: I think I raised the issue earlier today and I'm equally uncomfortable with having selected response rate as the primary efficacy parameter because even though we expect fairly high response rates in breast cancer, particularly with an active drug like