withdrew their consent because they didn't want to have the 3 times weekly injections.

One has to appreciate that at the time of initiation of this trial, it was not clear to the patients what their potential benefit of this therapy would be, and therefore the threshold, at least that was the risk a priori -- the threshold for withdrawal from the trial would be low. That has not materialized, fortunately, because withdrawal in total -- and I'll get back to that later -- was not considerable.

A little bit about demographics. As I said before, stratification by center was done, but not by the more relevant risk factors. However, with this number of patients, 499, it balanced out beautifully. Differences were very small. There was no statistical difference, for example, for the more powerful of the risk categories that we have used, which is Breslow thickness.

Further to Dr. Buzaid's presentation, you see here the categories of Breslow tumor thickness. Our patients in this stage II melanoma patient population consisted of patients with tumor thickness of 1.5 millimeters and more. This should be looked at in categories and not as a continuous variable because, obviously, these subcategories follow in some ways anatomical boundaries.

This is one of the busiest slides that I'm going to show this afternoon, and it will take me some time to guide you through it, but this is quite a crucial slide for the message of the presentation.

This is the long-term analysis on eligible patients for disease-free interval, and disease-free interval, the time from initiation of therapy to relapse, the difference remains significant. This analysis was done when a median time to follow-up existed of 4.4 years. That means that the first patients were up to 7 years in the trial and the last patient entered 36 months.

The time to 25 percent relapse -- and I do not show that on the slide here -- was 1.3 years in the observation arm and 2.1 years in the Roferon arm, a rather remarkable reduction of 25 percent, or 10 months.

The p value for the Kaplan-Meier estimates, as you see here, is .035.

The number of relapses in the Roferon arm in total was 100; in the observation arm, 119; a difference of 19.

Last but not least -- and that is perfectly justified by the protocol -- if one would do a cutoff analysis, something that most simple people like myself would understand better, if one would do a cutoff at 3 years, then the percentage of withdrawals here would be 32

percent and 49 percent in the observation arm, a difference of 17 percent. With stratification by center, that carries a p value of .005.

Breslow thickness, as presented before by Dr. Buzaid, is a powerful risk parameter or prognostic factor. We show this slide here today of the Kaplan-Meier estimates for the specific subsets of Breslow thickness only to show that the impact, the effect, for all categories is similar.

I also need to inform you that there was no interaction between this risk parameter and the outcome as disease-free interval, nor was there any interaction between age and sex and this outcome parameter.

Before I start explaining this slide, it's my task to bring across to you that this study was never designed to evaluate overall survival. I'll try and explain that.

A sequential analysis was performed and a triangular design was used. That means that discontinuation of recruitment into the trial was done at the moment in time that there were enough events to answer the question about disease-free interval. By nature of things, there will always be more events such as relapses than death. Therefore, it is a little bit unreasonable to expect that one would be able to show a difference for an outcome parameter which has less events like death.

As it happens, we come close with a p value of .059. But the only thing we can conclude from that is that there is a strong trend.

However, as I said before, there is a robust correlation between disease-free interval and overall survival, and I will get back to that when I conclude this talk.

There were 59 deaths in the Roferon arm in this analysis and 76 deaths in the observation arm. It's obvious that at 6 years, at the tail end of the curve, like with the other curve, there are few patients in the analysis simply because median follow-up time here as well was 4.4 years.

Dr. Buzaid showed a slide in his presentation where he put together disease-free interval or time to relapse and overall survival. Sorry. This is disease-free interval obviously for both and here is overall survival.

I would like to show to you what the difference is between the two with regard to events. 100 relapses in the Roferon arm, 119 in the observation arm. 59 deaths in the Roferon arm, 76 in the observation arm. One difference of 19. One difference of 17.

I think that the crux of my argument for this afternoon is that if we manage to delay or prevent recurrence in this disease, it is possible that we may

delay death as an event. I think that that is an important thing to keep in mind.

The shapes of these curves are similar, but that's the only thing I can say about them.

It's very important for a regimen that has to be continued for 18 months that tolerability is more than acceptable. We have looked at the adverse event pattern of this dose used in this study, 3 million units 3 times a week, and we have concluded that the pattern of adverse events that we observed is not different from the pattern of adverse events that we see with the use of this drug in other indications.

There are no surprises and there are no events that suggest the sort of toxicity that one would relate to a higher dose of this drug that we have also seen in other studies with our drug in the past.

So, here you see the percentages of the patients with flu-like symptoms, asthenia, headache, nausea/vomiting, depression, and dizziness being the most commonly reported adverse events in this trial.

If we then look at the percentage of patients with grade 3-4 toxicity, then these percentages are low. Again, this is a well-established safety profile that we know and have seen several times before with the use of this drug.

What is important to show, however, is that there is a certain withdrawal rate, and this withdrawal rate is 14 percent. 35 patients withdrew from treatment over the course of 18 months. The majority of these withdrawals happened around the 1-year time point. More importantly, they were for events such as asthenia, flu-like symptoms, dizziness, depression, usually grade 1-2. There were 9 patients, though, with grade 3-4 that withdrew, and you see them described here. There were 2 patients withdrawn for severe increases in liver enzymes.

I will now move on to discuss the study that formed the supportive data for this application, the study performed by the Austrian Melanoma Group. Recruitment took place between 1990 and 1994, roughly in parallel with the French study. This was also a prospective, randomized, multi-center trial. Patients had Breslow tumor thickness of 1.5 millimeters and more, in other words, clinically node-negative patients, exactly the same patient population as we had in the other study.

The primary efficacy parameter was also the same, disease-free interval, time from initiation of therapy to relapse.

The dose was the same, the regimen slightly different, and the treatment duration was different. 3 million units were given 5 times weekly, once daily for 5

days, for a duration of 3 weeks, sort of an induction regimen. The maintenance part was, however, the same as I described for the previous trial.

I base this part of the presentation on the publication database. The data that I've presented and present from the publication, this publication has a patient number of 311: 154 in the Roferon arm, 157 in the observation arm. There is currently a database that has 330 patients, as 19 CRFs were collected after the publication cutoff.

Demographics. Again, I show Breslow thickness as a risk parameter only, and here as well, whereas there was no stratification for this parameter, both arms are well balanced. There is certainly no statistically significant difference between the two. There are only small differences that are not clinically relevant.

These are the Kaplan-Meier estimates for this study, also for disease-free interval. Here you see the observation arm. Here you see the Roferon arm.

This analysis was done in September 1995 when patients had been in the study for at least 1 year and observed and followed up for at least 1 year. So, recruitment took 3 years, 154 here and 157 on the other side.

37 patients relapsed in the Roferon arm, 57 in

the observation arm. The p value was less than .05.

Here you see our overall conclusions. We have seen parallel efficacy in two independent studies with 800 and more patients in these studies all together.

The reduction in recurrence rates or time to recurrence of 25 percent in our view is clinically meaningful. This translates into prolongation of disease-free interval of 9 to 10 months. The time to 25 percent relapse in the French study, in the pivotal study, was 1.3 years in the observation arm and 2.1 years in the Roferon arm. If we cut off at 3 years, 32 percent of patients have relapsed in the Roferon arm and 44 percent in the observation arm.

We have seen a strong trend towards increase in overall survival that is properly correlated with the increase we have seen that is statistically significant for disease-free interval.

This drug has a well established safety profile. The withdrawal rate over 18 months in this study was low. It was 14 percent, but in view of the fact that patients did not know exactly what their advantage was going to be, this was very reasonable. The drug was therefore well tolerated. Patients could continue with work and lead an essentially normal life. This is important for a prophylaxis regimen and a regimen that

relies on compliance and has to be maintained for 18 months.

We designed low dose Roferon-A for a situation whereby there's a low tumor burden and an intermediate to high risk of recurrence. What this therapy does is it may prevent or delay the dreadful moment of disease recurrence. It may, therefore, delay death as visceral metastases directly lead to death within 12 to 18 months.

We, therefore, recommend low dose interferon alpha 2a, otherwise called Roferon-A, therapy as adjuvant therapy of stage II melanoma patients. These are patients with clinically node-negative melanoma. This translates into a Breslow tumor thickness of more than 1.5 millimeters. We recommend a treatment duration of 18 months.

This brings me to the end of my presentation. Thank you.

DR. SCHILSKY: Thank you very much.

Are there questions from the committee members for the sponsor? Dr. Raghavan?

DR. RAGHAVAN: These are two quite large sets of data and you're asking us to accept disease-free interval as a good surrogate of overall survival.

The one thing that troubles me and puzzles me is the time of recruitment to these two trials was for the

French trial January 1990 to December 1993, and the Austrian trial sometime in 1990 to 1994. By my calculations, you should have follow-up data conservatively to 9 years and maybe to 10 years, and yet the survival curves that you present show weak power out at 6 years. So, effectively you're presenting old data that haven't been updated and yet asking us to accept disease-free survival rather than overall survival. Could you clarify why that is?

DR. HOOFTMAN: I would not immediately agree with that. With this proposal for this therapy in an indication of stage II melanoma, median time to death is 7 to 8 years. Our median follow-up is 4.4 years. We are, however, getting closer to the moment in time where we could produce longer follow-up data.

DR. RAGHAVAN: No. I'm sorry. I guess I asked the question without clarity and I apologize.

I understand what you just said, but the reality of the situation is that even your disease-free survival curves, unless I'm misinterpreting them, don't go out to the full time that would be eligible for the duration of follow up. It looks to me like the data that you've shown us, whether they're disease-free or total survival, are old data. I can't understand if you had patients entered in 1990 who you propose are still alive,

which I hope is the case, why the survival curves have so few cases at 6 years that are still going. It doesn't make sense to me.

Why have you censored at 6 years? Why do the curves not go out at least to the 9-year point?

DR. SCHILSKY: Would you please identify yourself?

DR. WASSNER: I'm Elizabeth Wassner. I'm working in oncology in Basel.

The dossier has been submitted two years ago. These are the data that you reviewed.

Now, if we look at 5-year survival data, which is actually a reliable time point in the study, we've got a p value of 0.021, which is even more significant than what we've presented here.

DR. SCHILSKY: Can we just clarify that perhaps by hearing a brief summary of the registration history? You just said that the materials were submitted two years ago and that that's the data that we're reviewing today.

DR. WASSNER: Yes.

DR. SCHILSKY: Since you originally submitted the data two years ago, have you provided any update to those data?

DR. WASSNER: We haven't been requested to do that, but it is planned, of course, to look longer into

these data. But right now this is the data we have, and we're actually claiming overall disease-free survival and this is, I think, mature data. Overall survival, of course, would request 10-year follow-up in this population, and an end of recruitment, which is December 1993. 10-year data are still far away.

MS. da SILVA: Just to clarify the regulatory history of the submission, we originally submitted our application of September 1997 and the year time clock for acting on that with FDA was in September of 1998 when we received questions and responses from them. We then took into account their comments and resubmitted a response in March of 1999, which included a second study with the Austrian publication, and then we are here before you today, of course. We were notified in July, so we have not submitted an update as of yet.

DR. SCHILSKY: Thank you.

Other questions? Dr. Nerenstone.

DR. NERENSTONE: I'm not familiar at all with these clinical trials groups. We're usually given a little bit more information about frequency of follow-up or how patients are clinically staged. That's sort of important in a study where it's a disease-free interval difference that you're looking at. Can you tell me how often these patients are followed and what kind of tests are done,

whether liver function tests are done, CT scans, or clinical, and how often that interval is?

DR. HOOFTMAN: Can I please defer this question to Professor Grob who was the lead investigator of this trial?

PROFESSOR GROB: Jean-Jock Grob, dermatology, France.

Both groups were followed exactly in the same way. People were examined every 3 months and they underwent CT scan and x-ray explorations every 6 months, exactly in the same way in the two groups.

DR. NERENSTONE: And were laboratory evaluations done as well at every 3-month follow-up?

PROFESSOR GROB: Yes.

DR. NERENSTONE: Were CNS relapses considered relapse?

PROFESSOR GROB: Yes.

DR. SCHILSKY: Could I just pursue that before you sit down? Because, as I understand it, the follow-up was done for 36 months according to the protocol, and then there was an effort made I guess by the company to then ascertain again the clinical status of all the patients sometime after the protocol-prescribed follow-up was completed.

So, can you tell us something about what the

follow-up of the patients was in that interval of time from when the protocol-specified follow-up ended until the data were collected again from all the participating sites? Did the investigators continue to follow the patients on the same schedule? Do we have a way of verifying in fact that they were followed on the same schedule with the same tests being done at the same intervals on both arms?

PROFESSOR GROB: Well, I would say that we were out of the limits of the protocol, but most patients were followed exactly in the same way and some were followed more closely because the follow-up protocol is a little bit less tight than the usual process in France. The only way to check it would be to come back to the files because a point was made after.

DR. SCHILSKY: Yes. It is a bit of a concern because the ascertainment of relapse status in a sense could be very unbalanced in that interval of time when the protocol was no longer necessarily being followed. Since that's the primary endpoint that we're looking at here, I think we have some concern about whether in fact patients were followed exactly in the same way. It was an unblinded study. There could have been biases in favor or against the treatment that were in the minds of the physicians or the patients.

Okay. Other questions from the committee? Dr.

Johnson?

DR. JOHNSON: I think I read and understood Dr. Hooftman's presentation to say that the pivotal trial was designed without consideration of the usual prognostic factors being used for stratification purposes. I believe that was correct. Is that correct?

DR. HOOFTMAN: I wouldn't say without consideration, but there was no stratification for the more powerful risk categories such as Breslow, nor for age or sex. However, as I showed you on the slide, there was no imbalance between the two.

DR. JOHNSON: I won't be too melodramatic, but I'm very surprised that a study of this size undertaken at the time that this was would have done that, to be honest. I'm just very surprised. This is not new information really. I just don't understand why a trial of this size would be undertaken without proper consideration of known prognostic factors.

What you showed us was a Breslow depth. You haven't shown us the other prognostic factors I don't believe.

DR. HOOFTMAN: Can we call up these? We have some backup slides, with permission.

I can already start and answer the question. There was no imbalance at all with regard to the risk

categories of Breslow tumor thickness, age, sex, location of primary or pathology.

DR. JOHNSON: Do you have location?

DR. HOOFTMAN: Here you see depicted the sites of melanoma or location of primary.

DR. SCHILSKY: Anything else you want to see, David?

DR. JOHNSON: Yes. Well, I want to ask a couple of other questions.

You gave us the overall survival data and you mentioned the number of deaths, but I don't recall. Were all of those deaths due to melanoma?

DR. HOOFTMAN: No, they were not all due to melanoma.

DR. JOHNSON: Can you give us the causes of death on the two arms?

DR. HOOFTMAN: 4 deaths were not related to melanoma, 2 in each arm.

DR. JOHNSON: The other question I have, I was also surprised at the differences in the number of patients not eligible on the treatment arm. I believe there were 9 patients, if I'm not mistaken, versus 1 on the observation arm.

DR. HOOFTMAN: That's correct.

DR. JOHNSON: The skeptic that I tend to be, if

all 9 of those patients had, in fact, progressed, what would that have done to your DFI curves and the observation arm had remained the same? Would it still be statistically significantly different?

DR. HOOFTMAN: That is a perfectly reasonable question.

DR. JOHNSON: I thought so.

(Laughter.)

DR. HOOFTMAN: Can I defer this to my colleague, Sam Givens, the statistical expert?

DR. GIVENS: My name is Dr. Sam Givens. I'm a statistician at Hoffmann-La Roche.

Yes, that is a good question. Let me start off by answering it in one way, and that is that the sequential analysis that was done, which was defined in the protocol as the primary analysis to stop recruitment of the trial, was done on all patients. There were no exclusions in that analysis and that analysis was significant at the .038 level.

I think they naively did not include Breslow in their anticipated statistical analysis for that sequential stop. Their thought was that if they're balanced, they'll be okay, and the other aspect was, when we followed the patients longer, the expectation was to include that category into the final analysis.

As to the question of if all 9 of those patients had died, I believe that reduces the difference in survival by 9 and would drop it from 19 to 10. My expectation is certainly that that would have lost significance.

DR. JOHNSON: I'm asking also DFI. This is overall survival. I'm asking for DFI as well, which is the only endpoint that you showed a statistically significant difference.

DR. GIVENS: So, now you're saying in the hypothetical situation on DFI, if we had known all 9 of those patients had had a relapse.

DR. JOHNSON: Correct.

DR. GIVENS: Well, those 9 patients were included in the analysis with what we knew about them, but I think that had all 9 of those died that -- or had all 9 of those relapsed, I would anticipate that they would not be significant.

DR. SCHILSKY: Dr. Lippman.

DR. LIPPMAN: Actually I had a comment and a question, but before that, just following up on the last point, all 9 patients were included in an intent-to-treat analysis that was presented in terms of disease-free and overall survival?

DR. GIVENS: The sequential analysis that was

done included all patients. There were no patients who had been eliminated at that time that led to the stopping of the trial -- stopping of recruitment. Sorry.

DR. LIPPMAN: So, I think that answers that question, Dave, if they were included.

DR. JOHNSON: Well, actually I don't think that's what I heard. What I heard is that those 9 were not included in that analysis. Maybe in the stopping of the trial but not in the analysis of the DFI.

DR. SIMON: If I could clarify what I heard, it sounded like they were included at the interim analysis that led to the stopping of recruitment, but they were excluded in the analysis based on further follow-up.

DR. JOHNSON: That's right. That's what I understood, and the numbers reflect that I think there.

DR. GIVENS: You are both correct with that statement.

DR. SIEGEL: Can I get a clarification? Dr. Simon just referred to the analysis that led to the stopping of the trial as an interim analysis. If I understood the presentation, that's the analysis you presented as the primary analysis with the .038. This analysis is the analysis when everybody had 3 years of follow-up, which you presented as a secondary analysis, and then additional follow-up beyond 3-year data -- you haven't

presented those data. Is that a correct understanding?

DR. HOOFTMAN: It's almost correct. The primary efficacy analysis was for disease-free interval. It was at the same time the analysis that determined the discontinuation of recruitment in the trial. You have to set that apart from the long-term analysis that is an exploratory type of analysis.

The third analysis was solely -- it was done retrospectively, but to get more information with regard to overall survival. The trial and the protocol as such was written for a 36-month course. That means that the last patient entered reached 36 months and then the long-term analysis was performed.

DR. SCHILSKY: Dr. Lippman.

DR. LIPPMAN: I just have to clarify one other thing. Maybe I'm just missing the point. Hypothetically we assume what happened if they all progressed, and that's a big concern when they're eliminated from an intent-to-treat analysis. But we don't have to be hypothetical here. Right? You have follow-up on those and they were included in your analysis? We know as much as we know about those patients?

DR. HOOFTMAN: These are the patients that were excluded from this long-term type of analysis. 5 of these patients never received an injection because they, so to

say, got cold feet and they didn't want to be in the study once it was clear what was going to happen. 3 patients had the wrong diagnosis. The patients that you see at the top of the list had stage IV and died after a few days. The second patient had a Clark level I tumor. The third patient had lymphoma. The fourth patient had a previous melanoma, which was also an exclusion criteria, and the 1 patient in the observation arm had a previous melanoma.

DR. LIPPMAN: So that that would add 3 relapses, if they were included in patients that had the right eligibility criteria.

DR. JOHNSON: Well, no. I would say 5 at a minimum, the 5 who withdrew their consent. To me that's not an intent-to-treat analysis. That's a "I took out 5 people I didn't want to include" analysis.

DR. LIPPMAN: The question that I had actually is this issue of disease-free interval and the importance of that. Actually in the context of everything that we've heard this afternoon, the first presentation by Dr. Kirkwood and this, I actually was very disturbed by the finding of 1690 and the explanations for that in which you saw significant improvements in disease-free but absolutely nothing, not even a trend in survival. In this case there's a significant effect in disease-free survival and a .056 which translate to 59 deaths, if I read the slide

correctly, in Roferon, and 76 in the observation arm. So, it's certainly consistent and in the right direction.

But I want to get to the explanation that was given by Dr. Kirkwood, at least that I asked earlier, that the major aspect of that difference in survival he thought could have been explained by salvage interferon. So, the question here, have you looked at patients? Two issues. One, on the observation arm, if there as a drop-in rate on the interferon. Certainly it has been available and people have been talking about interferon and melanoma for a long time. And two, at relapse, the differences between the arms in terms of salvage interferon.

DR. HOOFTMAN: Would you please repeat the question?

DR. LIPPMAN: So, the question is, on the observation arm, of the patients that recurred, what was the salvage therapy? Were a substantial number of the recurrences on the observation arm treated with interferon at recurrence?

DR. HOOFTMAN: The only thing I can do in this situation is ask Professor Grob to answer the question. I think that the difference with what Dr. Kirkwood's group has done is that we have not formally retrieved that information in a retrospective fashion.

PROFESSOR GROB: If I understood you well, the

question is what kind of therapy did the patient receive after relapse. We do not have this information in our data. Of course, we can go to the files, but I think really that none of the therapy of metastatic disease, of distant metastatic disease, visceral metastases has shown any effect on the overall survival. So, this is my first answer.

And the second would be that it is highly likely that the treatment after recurrence were well balanced between the two groups. But the effect of the treatment on the overall survival, I would be happy to get one.

DR. LIPPMAN: The reason I bring that is up is I was surprised also by the presentation of Dr. Kirkwood that there as a major difference between the arms in terms of who had gotten interferon, and that that was the best explanation at least that exists, as I understand, for the fact that you see an improvement in disease-free survival but nothing in terms of survival. If that was even a potential confounder in this study, that might account for why your p value is .056 instead of .049. Could that have played an effect if what Dr. Kirkwood told us is correct?

PROFESSOR GROB: Well, this is an explanation and a hypothesis which was provided by Dr. Kirkwood. I would say I don't share this explanation because really I

don't think that either IL-2 or chemotherapy or interferon can really change the overall survival. At least this has not been established in the literature, neither in my experience.

DR. SCHILSKY: Dr. Simon?

DR. SIMON: I had a few questions. One is you indicated there were 35 patients who withdrew from treatment. How were they handled in the analysis?

DR. HOOFTMAN: You're asking a question about the 35 patients --

DR. SIMON: Yes.

DR. HOOFTMAN: -- the 14 percent who withdrew from treatment?

DR. SIMON: Right.

DR. HOOFTMAN: As usual, they were all included.

DR. SIMON: Their follow-up continued as for the patients who did not withdraw from treatment?

DR. HOOFTMAN: That's correct.

DR. SIMON: I would like to get some clarification about the database that was used for the analysis, not for the interim analysis because my experience is at a time of interim analysis, there are delays in reporting and that's really not necessarily a very accurate database, particularly in a multi-center

study with many centers involved and particularly when you're using something like a triangular test in which the protocol says you do analyses after every 20 recurrences. I don't really think that's practical in a multi-center study, and I have questions about the accuracy of the database in a situation like that. So, I would like clarification. So, for me, that's really not the definitive analysis.

I would like clarification of what additional follow-up was performed and what kind of auditing was done and how long each patient was followed and what proportion of the patients were lost to follow-up not for the interim analysis but for the subsequent analysis.

DR. HOOFTMAN: I understand the question. Can I give the work to a statistical colleague who was intrinsically involved at the time?

DR. RAMISIO: My name is Dr. Maurizio Ramisio, statistician, Hoffmann-La Roche, Basel.

The database that was used for the third sequential analysis is unfortunately not available anymore. We collected complete information on all the patients in the beginning of 1996 and, as Dr. Hooftman said, getting a new informed consent from all the patients. The follow-up analysis that has been presented is based on those data.

The triangular test analysis that has been

presented is based on the data of the 1st of January 1994, which are not available any longer.

We have simulated an analysis at the time of the 1st of January 1994 by putting a cutoff, using the data that we have to date, but putting a cutoff on the 1st of January 1994. The result that we have got with this analysis is still significant, is 0.035 on the log rank test. But again, we are not able to reproduce the analysis of that time.

DR. SIMON: So, the .035 represents an estimated significance level at the time that that interim analysis was performed?

DR. RAMISIO: This is what I'm saying now. What has been presented by Dr. Hooftman is the result which was obtained by Professor Chastung at that time doing the third sequential analysis on the data which was available at that time.

DR. SIMON: Suppose we forget about sequential analysis. Can you just clarify what is the most complete data available?

DR. RAMISIO: All right. The most complete data available is the data that have been collected in the beginning of 1996, and this is the data that have been presented as follow-up analysis by Dr. Hooftman.

As I said before, if we do a cutoff on that set

of data, which has been quality controlled, and source documents verified, and we do the analysis as it would have been done on the 1st of January 1994. We get a log rank test with 0.035 percent.

DR. SIMON: Suppose you don't do a cutoff and you just do the analysis with all of the data.

DR. RAMISIO: If we do the analysis with all of the data -- I don't remember what was the significance. If we do the analysis on disease-free interval, including all the patients, so intent-to-treat, including all the 499 patients, we have to exclude 2 who had no follow-up visit at all. They went into the study. They were randomized but had no visit at all. So, if we analyze that -- I'm sorry. I must find the right page.

Here. The disease-free interval -- the significance, stratifying by center, is 0.074. If we do the analysis on the eligible patient population, so excluding the 10 patients that we have discussed about before, we get a p value, which is 0.035. This is including all the data available up to the beginning of 1996.

If we do the analysis as it was prescribed by the protocol, we said an analysis will be performed at the end of the study, which could be interpreted as when all the patients will have had 3 years follow-up. The p value

becomes 0.005.

Is this answering your question?

DR. SIMON: What was the last point? If you do what?

DR. RAMISIO: The protocol prescribed a primary analysis, which was the sequential, and said, unfortunately a little bit unclearly, a further analysis will be performed at the end of the study. So, it is a matter of interpretation what is the end of that study.

In another place, the protocol says the patients will have to be followed for 3 years. So, an interpretation of the end of the study might be when all the patients will have been followed for 3 years. So, if we do an analysis cutting all the data following the 3 years, so treating is censored all the patients who had a relapse after the 3 years, we obtain a log rank test with a p value of 0.005.

If we do not do that, if we take all the data considering a median follow-up of 4.4 years, where some patients have been followed up for 3 years and some have been followed up for 6 years and more, then we get, on the eligible patients population, a p value of 0.035 and, on the ITT population, a p value of 0.074.

DR. SIMON: One other question. You didn't present any data on sites of recurrence, which ones were

resectable, which weren't. Do you have that data?

DR. HOOFTMAN: Yes, we have that information. We just have to find it.

As you can see here, the recurrences were mainly regional or local as opposed to visceral.

DR. SCHILSKY: Dr. Blayney.

DR. BLAYNEY: Thank you. I have three questions.

As has been alluded to earlier, in an analysis where you're looking at disease-free interval, there's a potential for bias introduced into the ascertainment of the data points because patients may be lost to follow-up, the ones that recur may die without knowledge of the investigator. Without a prospective plan for follow-up, this is of some concern in trying to interpret the data. I guess I would have some more comfort if you could tell me how many patients were lost to follow-up and how these were handled in your analysis.

DR. HOOFTMAN: Please bear with us until we find that information.

Can I defer this question to Dr. Sam Givens?

DR. WASSNER: We only lost something like 6 patients to follow-up in the long-term follow-up in the no-treatment arm and 8 patients in the treatment arm over the 7 years of the trial.

DR. BLAYNEY: So, since those numbers are equal, I'm understanding that there's probably a -- or roughly equal, there's no bias, likely there would be no follow-up bias in that.

DR. WASSNER: No. And less than 2 percent of the patients have been lost to follow-up over this period.

DR. BLAYNEY: In your slide number 111, you have a p value of .038. Now, maybe Dr. Simon's question got to this issue, but is that p value adjusted for multiple analyses?

DR. WASSNER: Yes. This value has been adjusted only for that, only for the multiple analysis, not for any prognostic factors.

DR. BLAYNEY: Thirdly, why did you choose or why was it chosen to give patients 3 million units and not adjust based on body surface area or some other measure of size?

DR. HOOFTMAN: The decision by the clinicians separately for the French study, as well as for the Austrian -- they made that decision separately and not knowing from each other what they exactly were going to do -- was based on the fact that they were looking for the dose that could be maintained for a long time and the lower dose that was effective, which was 3 million units, as used in other indications, for example, hairy cell leukemia, at

the time.

DR. SCHILSKY: Let me just make a comment to the committee. I'm bound and determined to keep us on schedule this afternoon because I know that some committee members will have to be leaving. So, we have about 3 minutes left for questions. So, let me just ask you to just keep your questions very focused.

Dr. Raghavan, do you have a question?

DR. RAGHAVAN: I just wanted clarification of one quick thing. I think I understood somebody from the sponsor to say the database is no longer available. What does that mean and why?

DR. GIVENS: What that means is that they did not save the database when they did the publication. They kept adding to the database and making corrections. So, the database as of today is the most up-to-date that we have, but we don't have a copy of precisely what they used when they did the sequential analysis, which is why we went back and said, let's cut off all data that should have been collected on visits up until the 1st of January and do the analysis again.

DR. SCHILSKY: Dr. Nerenstone?

DR. NERENSTONE: Very briefly, first of all, was there central pathologic review?

DR. HOOFTMAN: No, there was not.

DR. NERENSTONE: We've heard about how many patients were withdrawn because of adverse experiences. However, you have no information about what actual dose was given, what kind of delays there were in the patients who were on treatment for specific toxicity or even for the asthenia, depression, and flu-like symptoms. Do you have any other data available about that?

DR. HOOFTMAN: Yes, we have. We have information with regard to dose reductions. About 83 patients, 33 percent, in the Roferon arm had their dose reduced temporarily.

DR. SCHILSKY: Any other questions from the committee?

(No response.)

DR. SCHILSKY: If there are none, then let's break for about 14 minutes and reconvene promptly at 3:15. Shorter if we can.

(Recess.)

DR. SCHILSKY: We'd like to continue with the FDA presentation.

DR. CARDINALI: Good afternoon. My name is Massimo Cardinali. I will introduce the FDA perspective on this application.

First, I would like to acknowledge the review team that worked on this application. Dr. Neeman did the

bulk of the statistical review, and Dr. Tiwari also participated in the review. Dr. Gupta in the last week or so did some additional analysis.

This slide is to remind the approved indication for this product. The indication for the hairy cell leukemia has the closest dosage to the one that the company is seeking for this application.

This is the indication that the company is seeking for this product as presented in the submission.

I'll briefly go over the events that took place. You see in white the company and in yellow the agency. The supplemental application was submitted in 1997. The company provided us with the translated protocol and statistical plan and database for the Grob study, as well as the available literature at the time on the subject and an unpublished report. This was the study WHO 16, the Cascinelli study.

We finished our review in March of '98, and Dr. Neeman asked the company for some additional information on the Grob study and that was received in May of that year.

The monitoring of the French centers was completed in May of '98.

We issued a complete review letter in August of that year. The database and data that the company provided was perceived to be not sufficient for approval by the

agency, and we requested a database for the other study with Roferon that was available, as well as some additional clarification on the Grob study. The information was provided in November of that year, and the paper for the Pehamberger study was submitted to the application in March of '99.

We received about a month ago the translated study protocol for the Pehamberger study and early this month the data set that Dr. Gupta analyzed.

I will go briefly to the structure of the two studies. The Grob study was conducted between 1990 and 1994.

The inclusion criteria, essentially patients with AJCC stage II and no previous therapy was in the provision of the protocol. And the performance status was set as ECOG less than or equal to 2.

The endpoint specified in the protocol, disease-free interval, and as secondary endpoints, overall survival and tolerability of the treatment.

The dose administered was 3 million units 3 times per week subcutaneous for a total duration of 18 months.

The study conducted in Austria was started approximately at the same time and the same duration than the French study. The inclusion criteria were almost

identical in terms of the staging of the disease. There was no systemic therapy within 3 months of inclusion in the study and the performance status was a little more stringent.

The material that we received did not specify the endpoint, and there was no statistical plan in the protocol.

Again, the studies are very similar. The difference that we can observe is the duration of the treatment. The study had an induction phase of a 3-week duration and then it was continued at 3 million units 3 times per week for a year.

I'll leave the floor to Dr. Lachenbruch that will summarize the results and the statistical analysis.

DR. LACHENBRUCH: Thank you. I'm almost an imposter up here in that the primary analysis was done by Dr. Neeman at the FDA and then later Dr. Tiwari did this work.

The study by Grob, M 23031, is the primary trial that was submitted to the FDA. This trial was planned to have sequential looks every 20 events. However, the timing was not adhered to and three looks were done.

As you can see here in a triangular test, a score Z is computed, and if the null hypothesis is true, that will be around 0, and a variance V is also computed

which is proportional to the number of events at the time of analysis. If the points exceed the upper boundary, the null hypothesis is rejected, as you see. On January 1st, '94 when the analysis was done, it did exceed the null hypothesis.

During the FDA review, we requested that the sponsor submit more mature data from the additional follow-up that they have, and our analyses are all based on an intent-to-treat at this time of final analysis.

This is a graph you've seen before. The medians are indicated. Because the number of relapses at and before this time of the medians, the estimate of the medians may be somewhat variable. This again is based on the ITT population and not the per-protocol population. This results in an additional 9 patients being added to the overall population, and the significance level that we see here is .095 as opposed to the .038 from the sponsor's analysis. This is no doubt due to both the additional data, more mature data, and the additional patients.

The overall survival is shown here, again with the ITT population. We came up with a .09 p value.

We also decided to examine some additional analyses which are exploratory, and these are, indeed, post hoc but I think they are of some importance. This slide shows the effect on relapse-free survival of the covariate

alone, and that's important to realize. Thus, the Breslow thickness has a p value of less than .001. That is for the effect of Breslow thickness on survival. It is not a p value for Roferon given Breslow thickness.

Among these data, the p value for Roferon is larger, i.e., less significant, than for any of the others. Also, I should point out that Dr. Neeman used the Breslow thickness as a continuous rather than as a categorical variable.

We also attempted to find a best model for using the covariates, and in this case we found that Breslow thickness, age, and sex gave the best model. Adding Roferon treatment to those three led to a p value for Roferon of .25. The sponsor, Roche, did do a similar analysis. They dichotomized age as greater than 50 or less than 50. The differences may be due to more mature data, the use of age, or the additional patients.

The results are marginal significance. The p value at the time of the termination of the study is .038, but after the data had matured, it was .095.

We received the Pehamberger data last week, and we have been unable to do a detailed and rigorous analysis of the results. We received a translation of the protocol about a week earlier.

We attempted to reproduce the analyses that

appeared in the article and will present some comments. The inclusion criteria, of course, are essentially the same as for the Grob study. The analytic plan was not presented in the protocol and endpoints were not specified. We used relapse-free survival and overall survival, and we've also done some adjustments for Breslow depth and did a corresponding analysis including age and gender as we did with the Grob study.

Here we see the relapse-free survival, and we found a p value of .04 and median for controls is 4. The Roferon group did not reach a median.

In doing the same proportional hazards model, we find quite similar results. Breslow thickness is highly significant; age, significant; sex, somewhat less; and Roferon as, of course, .04.

At the same time we did the adjustment for Breslow alone, which is what was reported in the Pehamberger article, and found a p value of .1, and if we adjust for Breslow thickness, age, and sex, we had a p value of .22, quite similar and comparable to the p of .25 that was seen in the Grob study.

Again, our conclusions seem to show that there was a moderate effect of Roferon by itself, which is the primary analyses that are presented by the company. However, adjusting for Breslow thickness and other

variables does seem to reduce the effect.

Based on this, we felt that it was appropriate to begin planning an overview of the published literature. So, we are doing this to combine the evidence. What we want to do is substantiate the evidence of efficacy from known studies of adjuvant interferon in melanoma, and for this purpose, we will use studies of both Roferon and Intron. These are exploratory and we want to emphasize that the data support from Roche will be the only material that is used in any decisions regarding this product. We will be using relapse-free survival and overall survival, as they are the generally accepted outcomes. And we are in the process of obtaining data from investigators.

We will be looking at Roferon and Intron trials. We want them to be randomized, concurrent controlled trials, and so far all have an observational control and are for adjuvant therapy.

We have searched a number of databases seen here. The trials that we have identified and the studies come from North America, Europe, Australia, and New Zealand. We will be looking to get estimates of the odds ratio by means of ratio of medians, and that's very nice if you happen to have exponential survival. That's for the statisticians. And the Peto method is basically a log rank type method.

We will also be looking for estimates of survival, either relapse-free or total survival at 3 years. We'll be looking at Kaplan-Meier estimates, 95 percent confidence intervals, and so forth.

So far the studies that we have found are those from Dr. Creagan, Dr. Cascinelli, Dr. Grob, Dr. Pehamberger, which all were using Roferon. We've seen five studies from Kokoschka, Kirkwood, Cornbleet, Rusciani, and the Kirkwood ECOG 1690.

This slide provides estimates of the percent improvement and confidence intervals for relapse-free survival that we have seen thus far. A square is placed at the estimate for the difference in proportions. The whiskers are the 95 percent confidence intervals. A positive value is favorable for interferon. So, if the whiskers cross the line, it is not possible to rule out a difference of 0 between observation and interferon.

The size of the box, that is the area, is proportional to the sample size. These generally indicate a consistent improvement of about 8 to 9 percent over observation. We don't have reliable 5-year data at the present time to conduct a similar display.

In overall survival, we see the same picture. As you can see, there's a bit less of an impressive difference in these. We did not have the data from Dr.

Pehamberger for survival. The difference is around overall about 4 to 5 percent.

Our next steps will be to get individual data from studies and perform the analyses that we have indicated above. The information contained in the literature does not permit sufficiently detailed analyses.

To summarize, for relapse-free survival, all studies do point in the same direction. These are marginally significant or barely not significant, and there's a moderate early effect. But we don't have a lot of data for longer term effects.

For overall survival, there is a consistent trend toward improvement but evidence is not that strong, and I have in my notes, parentheses, "yet" with a question mark. We did not show it, but there do seem to be fairly similar results with high and low dose and with node-positive and node-negative disease from the material that we've seen.

Thank you.

DR. SCHILSKY: Thank you very much.

Questions for the FDA? Dr. Raghavan?

DR. RAGHAVAN: I'm totally mystified as to why you went through that statistical exercise because the best data points come from a product that isn't even up for submission. So, I just wondered why you spent all your

time doing this and what the point was.

DR. LACHENBRUCH: The purpose here was to really look for evidence combining all of the Roferon data. Over here, we see that there are four studies, and so what we would like to do is be able to draw information from all of these. So, what we see is overall there does seem to be a significant improvement in 3-year survival.

DR. SCHILSKY: Other questions? Dr. Simon?

DR. SIMON: I guess I wouldn't put much credence in a meta-analysis based on literature data. There may be exclusions. There are all kinds of biases in published reports. The fact that they're published may be publication bias. If you're planning on doing an individual case meta-analysis, I would say go ahead and do it, but I don't find it useful to present a meta-analysis based on publications.

DR. LACHENBRUCH: These are very preliminary results, and we are trying to get the data at the present time. So, I would agree with you.

DR. KEEGAN: I think to some extent the reason why these data were presented was that up until very recently, the only information we had was from a single study. So, this was our attempt to see what other information was available in support of this application. We're not saying it's optimal information, but it was all

that we had available.

DR. CARDINALI: As a note, the Pehamberger and Grob study data is from the publication not from the data set we have analyzed.

DR. SCHILSKY: Dr. Simon.

DR. SIMON: Do you have any insight for the French study as to why the significance level, say, for relapse-free survival, after adjustment for thickness, age, and sex, changed so much? Were there any imbalances?

DR. LACHENBRUCH: No. For a covariate analysis, as you know, the purpose is not necessarily to adjust for imbalance, although that can be one use of it, but these happen to be important prognostic factors for survival. So, what we're saying is we'd like to look at these after we have adjusted for these.

DR. SCHILSKY: Dr. Lippman.

DR. LIPPMAN: Just a quick clarification. In your last conclusion slide, you said that there were similar results with high and low dose. Is that what we just saw from Dr. Kirkwood with Intron or is that with Roferon?

DR. LACHENBRUCH: I believe that was the for the Roferon, the study of Dr. Creagan and the Grob and --

DR. SCHILSKY: Other questions from the committee members?

(No response.)

DR. SCHILSKY: Okay, thank you.

Let me point out to the committee members that there's a slightly different set of questions than the ones that were in the blue folder, and those should have been put at your place right after lunch. It looks like this. It's a two-page thing. It has only one of these meta-analysis charts. I think the content of the questions is largely the same, but these are the questions that we should be focusing on at this point.

Before we get into the questions, actually I'd like clarification of one point from the FDA because most of these questions are posed in such a way that they ask us to consider the results of the sponsor's data in conjunction with the overview analysis that was just presented. Now, I was quite sure I heard the FDA presenter say that the overview analysis would not be taken into consideration by FDA in assessment of the sponsor's application. So, could we get some clarification on that?

DR. LACHENBRUCH: Yes. What I said was no Intron data would be taken into account.

DR. SCHILSKY: I see. It's a little bit difficult for us to sort out from those meta-analyses which ones had Intron data and which ones had Roferon data.

DR. SIEGEL: Let me clarify something. First

of all, the Roferon data were the top part of all those slides and are on the second page of the questions.

The FDA has a policy regarding use of literature in support of applications for new indications for already approved drugs. The gist of the policy says that literature data, especially if consistent and compelling from multiple sites, can be important, but the value of the data is largely dependent on the ability to substantiate it through finding protocols, data sets, ensuring that there were intent-to-treat analyses, and the normal things. So, these are things I think that, as a matter of policy and procedure, should not be ignored, but I think that the weaknesses or concerns that have been highlighted are important ones to take into account.

DR. SCHILSKY: Okay, thank you.

Maybe we'll just get on with the questions then. Yes, Scott.

DR. LIPPMAN: I know that we're not considering Intron here, but I think the data are relevant in the sense that -- two issues. One is the biological plausibility mechanism and the other is consistency within the committee in terms of approval.

Again, we talk about the fact that there's very little data. So, we have one study of 500 patients which, at least in the FDA presentation, we've talked about those

mysterious 9 cases and how that would affect. But at least in the FDA presentation, it was significant. Every one of the boxes is -- it's modest, but it's positive both in terms of disease-free and overall survival, and the whiskers come very close, just past the survival curve of 0, as opposed to another situation where we're using interferon where it's approved and where you don't see that pattern even with a very high dose in terms of survival. And we've heard some explanations of that. It's really a question of whether we should take that issue, the consistency, the biology, the mechanism, into account in some of these discussions.

DR. SCHILSKY: I don't think we should ignore the universe of information that we're aware of and we have available to us.

I just want to get clarification on this again. First of all, the meta-analyses with respect to the Roferon data, which is what's on our question sheet -- so, there are four studies listed for disease-freed survival and three listed for overall survival. Of those, only the Grob study would appear to show a significant benefit with respect to disease-free survival as it's listed here. However, as the more detailed analysis of the study was presented to us, there are questions as to, in fact, whether even that study shows a significant difference in

disease-free interval. So, although the trend appears to be in favor of interferon in each of these examples, there's very little in the way of a statistically significant benefit for interferon.

Further, it's fair to say that, I guess, in a sense these are at best incomplete meta-analyses for the reasons Dr. Simon mentioned, that this information is just based upon data you could glean from published reports in the literature, not from the actual patient data that's contained within those reports. Correct? Okay.

Scott?

DR. LIPPMAN: Just to clarify, because with all the discussion, I guess I was sort of surprised when I look at this. I'm not talking about the meta-analysis, just the big box of 500 patients under Grob. It is significant, doesn't cross the line. I haven't read the recent set of questions, but one of them was should we recommend approval based on one large randomized trial. So, I'd like to clarify maybe from the FDA if they're going to stick with this box. In that case, that is statistically significant and survival is close and the other studies corroborate that. So, I'd just like to clarify.

DR. SIEGEL: Well, I guess a lot of people have addressed different parts of this question. I'll take my turn.

That box was an endpoint that was chosen in part because it was, I think, the easiest endpoint to get on all of the trials, and it's endpoint data truncated at 3 years. That's the endpoint that the Grob data looked the best at because, in fact, the curves have maximal separation at about 3 years and start coming together after 3 years. As noted, that studied had 3 years of planned and prescheduled follow-up, so it's not an irrelevant time period for that study. But at best, let's say that the primary time for follow-up is ambiguous in the protocol and difficult to determine. As we determine it, the intent-to-treat analysis of the most complete available data set was at the .095 level and with covariate correction at the .25 level.

We'll stand behind that analysis. It's one of several analyses. We won't stand behind it as like the one that tells the story. I don't think, given the ambiguities of the protocol and the flaws and strengths of different analyses, that there's probably not one p value that you can hang your hat on and say this tells you the statistical significance of the trial.

DR. SCHILSKY: Are we ready to go to the questions? Let me just read the first question. There's a two-paragraph summary. Then the question is, does the committee find that the results of a single multi-center,

randomized, controlled trial, in conjunction with the overview analysis of the three randomized, controlled trials of Roferon-A, provide substantial evidence that Roferon-A prolongs the disease-free interval in patients with surgically resected melanoma?

Is there discussion on that before we vote? Dr. Lippman.

DR. LIPPMAN: I will just say that the real fundamental issue that I'm having a problem with is the floating p values. Given that we've heard a lot of discussion on this and still know real consensus, I don't think, in terms of what is either reasonable or meant or intended, that's going to fundamentally affect how I vote anyway on this.

DR. SCHILSKY: Well, I think we've seen the data as presented by the sponsor. We've seen the data as presented by the FDA with the adjustments to the p value, if you will, based upon the other covariate prognostic factors. We've seen, for what it's worth, the preliminary meta-analysis. So, is there anything else you would like to know before you vote on this?

DR. LIPPMAN: I think fundamentally if we knew exactly in the design what the primary endpoint was -- was it a 3-year? I think that's where the debate is.

DR. SCHILSKY: It appears that we don't know

that because it wasn't well specified.

DR. KEEGAN: That's correct. The protocol really is open to quite a bit of interpretation as to when that final analysis was to have occurred and exactly what it was to consist of.

DR. SIMON: I will say, however, that my experience is if you have an endpoint, that your most accurate analysis is the one based on the longest follow-up and that's what you should hang your hat on and not one that was simulated based on what might have happened some years ago. So, anyway, I guess that's one issue.

The other issue is for myself I guess I just have some basic uncertainty about the quality of the data from that trial, the potential biases in follow-up. It looked like there was too much of an emphasis that the main analysis would have been the one that was essentially an interim analysis that stopped the recruitment. Then there were sort of ad hoc attempts to increase follow-up. I just am left with some uncertainty as to how accurate that additional follow-up was. So, I myself, in addition to the variable p values, just have some uncertainty in the credibility of that data.

DR. SCHILSKY: Dr. Keegan.

DR. KEEGAN: I would say that the protocol did not specify what the continued follow-up should be after 36

months, and when we requested the additional data, it was necessary for the company to go back to the investigators, who then reconsented patients to get the information. From the monitoring inspections of some of the sites, it's clear that there wasn't a rigidly adhered to schedule for follow-up.

We did also ask the company to analyze the data to determine whether or not there was a systematic bias in terms of the follow-up, and it didn't appear that the follow-up was systematically biased towards one or the other arm. It was equally -- I won't characterize it as haphazard, but definitely not done according to a rigid schedule. But that seemed to be present in both arms.

One other point I'd like to make in terms of the policy is that for a single study in support of effectiveness, one of the criteria that FDA uses is that the trial have a statistically significant result that's fairly robust such that we would have confidence that the result would be reproducible. At best, the p value here is .04, and our concern at the time of even the review of the data with the most up-to-date follow-up that we could get through 1997 suggested to us that that result, although statistically significant, would not meet that condition of being so robust that we were convinced that it was a reproducible result, which is why we encouraged the company

to go back and obtain additional study data.

DR. SCHILSKY: Dr. Johnson?

DR. JOHNSON: Yes. I didn't realize this was going to take a lot of discussion, but since Scott seems conflicted, let me go through a number of reasons why I think this is a poor study.

First of all, I'm not sure I accept the endpoint as one that's therapeutically efficacious. DFI, in the absence of a survival benefit, is of uncertain benefit in my view. We can debate that but there are plenty of diseases where DFI can be prolonged and survival is not. And we don't do the therapy that prolongs the DFI. Small cell lung cancer immediately comes to mind. There are 10 randomized trials out there showing DFI is prolonged, survival is not. No one uses maintenance chemotherapy in that disease.

If they had shown me some quality of life benefit to that DFI, that symptoms had improved or some other meaningful patient benefit, then perhaps I could have accepted that as an endpoint of value, but I don't. And I didn't see that data.

Thirdly, again, I find it shocking -- and I think that's the word -- that a study of this size would be undertaken without appropriate stratification for known prognostic endpoints. That being said, even more

importantly, there was no quality control of pathology. We have no idea whether these patients were equally balanced other than what they tell us. There was no central review of the patient pathology. They could have all been one stage in the Roferon arm and quite another in the other, just on the basis of that inequity. All we have is a report. They've told us there was no central pathology review.

Candidly, I just think that the overall data are highly questionable. I agree with Richard. I think these are not the quality of data that we see come to this agency that generates approval by this body. That's my perspective on this, and personally I don't see how we can vote anything other than no on this question.

DR. SCHILSKY: Dr. Raghavan?

DR. RAGHAVAN: Yes. I think I always feel sorry for the FDA because they're victims and they get beaten up by everyone, but as a taxpayer I really have to say that I don't think you've done as well as you usually do this time. You've left it to the committee to identify a whole series of very bad statistical concepts and poor quality data. I shouldn't have to remind you: garbage in, garbage out no matter what the p value. I just feel very disappointed that we've had to go through this exercise.

Dr. Lippman has tried very hard to be fair, and

I recognize and respect that. For those of us who are crusty veterans who have seen outstandingly good data over the years, this is not an example of that. And bending over backwards to bring in Intron data that were approved based on good quality data and then tainting that information based on very poor quality information with bad follow-up sets up a precedent that that I think is kind of disappointing. And I would hate people to leave here starting to question decisions made in the past based on good data when we've now added a bunch of information that's out-of-date, hard to quantify, irreproducible, et cetera.

And I just felt I wanted to make that comment. I apologize for beating you up, but you deserve it.

(Laughter.)

DR. SIEGEL: Allow me to respond in part, although I don't want to take up too much time with this.

First of all, I think it's a mischaracterization to suggest that it took the committee to identify the flaws in this data. I don't think there was a flaw discussed here that was not identified by the FDA. The FDA did an intent-to-treat analysis from the beginning. We carefully inquired and investigated about the relevance of the follow-up data, the quality of the follow-up data, and the choice of the endpoints, and made a

presentation of the data, I think, that accurately reflects our perception.

As to the question of why these data were brought before the committee, perhaps this requires a bit of understanding of time lines. At the time we need to make a decision about scheduling a committee, it's usually a couple months before the committee. As we have made clear in the presentation, we had felt that based on the Grob study alone, there was no reason to discuss or consider approval of this application.

What we had available to us at the period two months before this committee was a published report from the Pehamberger study that showed a p value of .02 and new information from the company that they were, in fact, going to be able to get the data set and the protocol. Those, as you've heard, I'm sure for a good reason, took longer than anticipated to get. So, they arrived within the last week or two. You've seen the preliminary analyses of those. The study did not look like what we expected it to look like, but I think with that perspective, perhaps you can better appreciate where we've come from.

DR. SCHILSKY: All right. Thank you.

In the interest of time, I'm going to call for the vote. I think we're probably ready. Let me just restate briefly the question. Does the committee find that

the results of a single multi-center, randomized, controlled trial provides substantial evidence that Roferon-A prolongs the disease-free interval in patients with surgically resected melanoma?

All those who would vote yes, please raise your hand.

(No response.)

DR. SCHILSKY: That's 0 yes.

All those who would vote no?

(A show of hands.)

DR. SCHILSKY: 7 no.

Abstentions?

(A show of hands.)

DR. SCHILSKY: 1 abstention. Sorry. 2 abstentions.

DR. SIEGEL: I think we're done.

DR. SCHILSKY: That's what I was about to ask because the second question says, assuming that the answer to question 1 is yes, well, we know now what the answer to question 1 is. So, I think that completes the committee's deliberations. Thank you all very much.

(Whereupon, at 4:02 p.m., the committee was adjourned.)

- B -

- C -

- D -

- E -

- F -

- G -

- H -

- I -

- J -

- K -

- N -

- Q -

- R -

- T -