we increase the levels within the complete liver or the tumor by giving histamine injections.

So, I think actually we have an environment that has the possibility to be inhibited, the H2 receptors are being blocked by histamine, and this hypothesis would work.

If we do the experiments with tumors in the liver at sub-Q, the potential of giving histamine in preclinical experiments, we have a larger growth rate reduction in the liver than in other tissues.

DR. GEHLSON: I did have an answer for Dr. Dutcher's question. According to our database, we had no patients in either arm that had performance status of 2 either in the intent-to-treat population or in the population with liver metastases.

DR. NERENSTONE: Dr. Albain.

DR. ALBAIN: I have two questions directed toward Dr. Atkins, please. In your very nice review, Mike, you mentioned the lack of real efficacy data for the low dose IL-2 regimens.

Could you comment and your thoughts as to what could be going on here since it was a low dose outpatient regimen, could this potentially be histamine effect alone if indeed there is an effect? That is my first question. I will let you do that one first.

DR. ATKINS: I think the low dose IL-2 control

arms in this particular study supports the data that I showed with low dose IL-2 on that particular table with the response rate in the 2 or 3 percent arm.

I think from listening to this presentation, I wasn't involved in the study, that the theory behind histamine is that it might take an immune reaction that is being inhibited by some other factors like oxygen free radicals, and allow it to go on even with less immune stimulation than is typically used with high dose IL-2.

DR. ALBAIN: The other question to you is are you aware of any multivariate analyses done specifically for patients with liver metastases in the various historical databases from the large groups? In other words, are there indeed other independent factors present predicting adverse outcome in the face of known liver metastases?

DR. ATKINS: I wouldn't be able to comment on that definitely although I know that these are multivariate analyses and I assume that means when you look at patients with liver metastases, and it comes out as an independent predictor, it takes into account all the other variables.

DR. ALBAIN: I am just trying to get at the possible explanation. What is puzzling me the most here so far is that at least if you look at the FDA analysis, there is neither a response or a time to progression benefit yet there is a survival benefit.

I wonder if either of you could comment on the FDA analysis that was presented in the briefing documents for time to progression because you did address the response issues in your presentation.

DR. NERENSTONE: Kathy, I think I would like to hold that question until after FDA has a chance to present their data.

DR. ALBAIN: The reason I asked is because he did some rebuttal of the response analysis and stated that the reason there was a survival advantage was perhaps because a response would not bear out with stable disease as another important outcome.

DR. NERENSTONE: I think it is going to be important for us to hear the entire FDA presentation and then open it back up, and if it is appropriate for the sponsor to react to that, I would like to save it for that time.

Dr. Taylor.

DR. TAYLOR: I have several questions. Number one, all of this is done at home. Has anybody counted bottles, do you keep diaries, how do you know how compliant they were and how much drug was actually given by the patients?

Number two, did you control for other drugs that they might be taking, such as antihistamines that might

interfere with your histamine, and if you did so, did you control so in both arms?

Number three, on your quality of life study, a common problem that we have is that as people get further along in their disease, they answer fewer and fewer of the questions, and that is a lot of questions. What percent of people actually completed your Quality of Life studies?

DR. GEHLSON: Thank you for those questions.

Yes, in the beginning of the study there weren't injection logs. There was certainly a lot of drug accountability that went on as the patients received their drugs each week, they were evaluated each week in the clinic with their physician and their nurse.

As the study progressed, we included injection logs that actually would allow the patients to document, and we also had, in many of the patients, home health care was provided, so the home health care nurse could actually, in fact, administer the drugs and fill in the injection logs.

Then, there was a lot of accountability. They had to return all the syringes and the vials each week, so it is an at-home treatment. We would hope that these patients were sufficiently motivated to administer their treatment at home, and we had the best reliable information we could have by their injection logs.

For the Quality of Life for the number of patients

who actually completed all of the forms, as they progress and drop off the study, they aren't going to be filling out any more forms, and I will actually ask the statistical group who put together the Quality of Life analysis to address that question.

But on your second question, can you ask me that?

DR. TAYLOR: What about concomitant drugs?

DR. GEHLSON: Oh, yes, a very good question. Of course, we didn't want our patients to have H2 receptor blockers because we know that this is an H2 receptor phenomenon.

They could have had H1 antihistamines for very short periods of time, but we know that the effect on the phagocytic cells H2, so they were not allowed to take any of the antacids that are H2 blockers, but instead we asked them to take the proton pump blockers, which is Prilosec and Prevacid, and then we also listed the other concomitant medications that they could have had in the protocol, and all that was monitored in great detail.

If I could have the people who worked on the Quality of Well-Being, they can actually better address that question than I can.

DR. ACKERMAN: I am Stacy Ackerman from Covance and I am a consultant to Maxim. I would like to help address the question about the amount of missing data.

Overall, in the histamine plus IL-2 group, there was 33.4 percent missing data in terms of the QWB self-administered questionnaire. In the IL-2 alone group, there was 34.9 percent, and tests for difference in terms of missing data, the chi-square value was 0.4, suggesting that there wasn't a statistical difference in terms of the degree of missing data between treatment groups.

Does that sufficiently address? Okay. Thank you.

DR. NERENSTONE: Dr. Santana.

DR. SANTANA: I, too, like Dr. Lippman, don't quite understand whether there has been clinical proof of principle of how this drug works, particularly in the group of patients in which it would appear that it works, but having said that, I also recognize--and this is a comment that there are a lot of drugs I don't understand how they work, but I still use them--my comment is primarily to safety.

Does histamine cause a tachyphylaxis both in terms of side effects with repeated dosing, and a correlate to that, translated to its potential physiologic activity as it relates to this combination with IL-2?

DR. GEHLSON: Thank you for that question.

I am going to ask Dr. Agarwala to actually address that question since he is one of the investigators that treated more patients in this study than anyone else.

DR. AGARWALA: Clinically, patients who would react to the histamine would start off at a 20-minute injection or so and modify that time based upon how they reacted. So that is one variable that came into that from a clinical standpoint as to their flushing and headache, and so on.

In general, patients tend to tolerate the treatment better as they go along. I am not exactly sure if that is from tachyphylaxis or is it from just them getting used to finding the right timing for administration of the drug.

However, there really wasn't a situation where patients would have a severe reaction on the first day of treatment and then by the third cycle were not having any reaction at all. They would all tell us they were having some flushing, some redness of the skin.

My feeling is that it got a little bit less, but did not completely disappear, if that helps to answer that, but I don't know of any biologic way of measuring that.

DR. SANTANA: Do you know if there is any preclinical data that would carry that further in terms of the effect of histamine on its reported physiologic effect and how it interacts with IL-2?

DR. AGARWALA: I am not aware of any. Peter, do you know of any preclinical data that would suggest

tachyphylaxis?

DR. GEHLSON: We have actually done in humans some pharmacokinetic studies looking at the administration of histamine before or after interleukin-2, and there is no interaction between the two drugs. So, we don't change the pharmacokinetic profile of interleukin-2, nor does it change the pharmacokinetic profile of histamine in patients.

DR. NERENSTONE: Dr. Simon.

DR. SIMON: One final clarification. Could you clarify was there a formal independent data monitoring committee? Your terminology and your description of the safety monitoring seems a little ambiguous to me. Was this something that was done by the CRO or was there an independent data monitoring committee?

DR. GEHLSON: That is an excellent question. Yes, we actually had a disinterested independent Data Safety Monitoring Board comprised of three clinicians in the United States plus one statistician that were not at all involved with the CRO or in the conduct of this study.

They independently evaluated safety on a monthly basis, and they did, as per the protocol, did one interim efficacy evaluation. That was a closed evaluation that we actually have not had access to. So, there was, in fact, a disinterested Data Safety Monitoring Board.

DR. SIMON: I didn't see in the protocol a clear

statement of when the final analysis would be done. How was it decided, was that decided by the independent Data Safety Monitoring Board or what? We heard Dr. Whitman mention in his statement before the presentation that he had analyzed the data from his clinic independently and noticed what he thought was a difference in outcome for the patients.

Timewise, when were those kinds of analyses being done by individual investigators?

DR. GEHLSON: Thank you. That is also a very good question. Relating to the DSMB, they actually had developed a very specific charter for when they would actually do an interim efficacy evaluation, and that charter is actually also included in our statistical analysis plan. So, they prespecified the time that they would do it and how they would do it.

As far as other people analyzing their data, I do not believe that Dr. Whitman analyzed his data before the study was complete or reported in aggregate, and I would have to ask--I mean I would hope that he hadn't, and I am sure that he did not, because we don't know of any other investigator who did.

DR. SIMON: But in terms of the final analysis, you were addressing the interim analysis, what about the timing of the final analysis or the analysis leading to the NDA submission?

DR. GEHLSON: That is a very good question. We had actually specified that we would analyze the data 12 months after the last patient was enrolled. We could have looked at a number of events that would have also given us a guide, and we assumed that 240 events or 12 months after the last patient enrolled would actually give us the level of significance we needed to show between the two groups, but we ended up with 238 events at the 12-month mark, so that is exactly when we specified.

DR. NERENSTONE: Dr. Przepiorka.

DR. PRZEPIORKA: You showed two slides which indicated that the majority of Grade 4 adverse events and deaths within 28 days were due to tumor, but when you do the subtraction it appears that there are still more events in the histamine arm than in the control arm.

Could you go through the nontumor-related deaths in the study, please?

DR. GEHLSON: Sure. I think we actually have those patient narratives, so I can go through those.

There were 4 deaths in the histamine group versus 2 in the intent-to-treat population, and 2 in the liver met population versus 1.

If it is okay, I will just give you some quick highlights on those patients. Is that what you would like?

DR. PRZEPIORKA: Just the causes of death.

DR. GEHLSON: The first patient in the IL-2 group died of influenza and respiratory failure, and it was considered not related to study drug.

One patient in the histamine/IL-2 group was considered to have liver failure, and that was considered not related to study drug.

Another patient in the IL-2 group had an upper GI bleed with cardiopulmonary arrest, and it was considered not related to study drug.

For those that were considered related to study drugs, we had one patient in the histamine group that it was considered a seizure leading to death, and that was considered related to the study drugs.

We had one patient in the histamine group with liver failure that was considered related to study drugs.

We had one patient with myocardial cerebral infarction that was also considered related to study drugs.

Those are the patients.

DR. NERENSTONE: Just to take the Chair's prerogative for a moment and to continue with this, the FDA's analysis actually, they gave us descriptions of 10 patients of whom they said there were 30 patients who died within 28 days from the last dose of study medication.

The thing that concerned me in terms of safety is that there seemed to be a significant number of patient

deaths in patients who had poor performance status, who by your own admission criteria should not have been treated on study. There is another patient who died who had received prior IL-2, again who should never have gone on study.

So, I guess the devil is in the detail, and this may be addressed by FDA in their presentation as to what exactly is disease progression and what is toxicity of the drug.

Would you comment on the number of ineligible patients who were actually on study, who had major violations of the entry protocol out of your study?

DR. GEHLSON: Yes, we can do that. I think we have a slide that will summarize the types of exclusions that were granted for patients on this study.

The medical monitor, just so you know, we had a medical monitor that was independent from the sponsor, and so the investigators, if they felt that they had an exception or a patient that they wanted to enroll, that may not have met all the criteria, they were to go through the medical monitor and it was the decision of the medical monitor, together with the investigator, that would allow such patients on.

I apologize for not having a slide for this particular one.

In the IL-2 alone group, there were a total of 42

patients that were granted exclusion or they were granted the eligibility into the trial, and there were 45 patients in the histamine arm.

Two patients in the histamine arm had received prior adjuvant IL-2 at some point in time in their course. There were 3 patients in the IL-2 alone arm and 1 patient in the Maxamine arm that did not meet the measurable disease criteria.

There were 10 patients in the IL-2 group and 11 patients in the histamine group that had scans that were greater than 3 weeks from study entry, but usually that was about a day or so.

There was 1 patient in each group that had a CT head scan in place of the MRI brain scan. There was 1 patient in the histamine group that actually did not meet adequate marrow, kidney, cardiac, or liver function criteria.

There were 10 patients in the IL-2 group and 2 patients in the histamine group that had PT or PTT abnormalities, 3 patients in the IL-2 group had hemoglobin abnormalities, 1 patient in the histamine group had a platelet abnormality.

There are some that are characterized as Other, and I might have to ask, do you know what is in the Other?

DR. NERENSTONE: Could you come to a microphone,

please, and identify yourself.

MS. WOOD: Diana Wood. Basically, in the Other category, contained general items that really couldn't be classified according to the protocol itself for which exceptions were granted. There were situations where an exception was granted, quote, unquote, however, that wasn't necessarily an exception that was required, and that was typically falling in that category.

DR. NERENSTONE: And so performance status, according to your records, was not an exception that was granted on the basis of performance status.

MS. WOOD: Not in our review of the data, no.

DR. NERENSTONE: Have you looked to correlate patients' own reporting of their performance status and correlating it with the investigators, because there at least seems to be some discrepancy in the patients who died, that performance status was not accurately assessed certainly in some patients.

I think this is a very important point because we are going to have to talk about the safety of this medication especially in patients who are at home, and the argument is that this is safer than high dose IL-2 and this is tolerable in the community.

I am very concerned that that may not be true.

MS. WOOD: I think that one of the things that we

can show is that the patients did have multiple therapies before they came into the study. Many patients were rapid progressers, and we did allow patients into the study that had brain metastases that were controlled.

I would assume that there are patients during the prestudy period, which is three weeks in length, that during that time, they may have had a lot of their work done within that two- to three-week prior to first dose.

Once they were randomized, they had two weeks before they start study drug, so it could be during the prestudy period, yes, the performance status was adequate and it met the inclusion criteria for the protocol, however, once they came into the protocol prior to the first evaluation, they would have deteriorated very quickly.

DR. NERENSTONE: Dr. Lippman.

DR. LIPPMAN: I would just like to follow up on a comment that Dr. Santana made on one of my comments. The reason I was focused on the biologic plausibility, I mean there are a lot of drugs that we don't know how they work, but they work, and in a large randomized study that is acceptable even though we do work to try to improve on that.

In this case, since we are being asked to consider as a primary indication a drug based on a subgroup analysis. In my opinion, the issue of biologic plausibility becomes much more important, and that is really the context that I

raised, and we will discuss obviously more about this, the whole issue of the subgroup finding even preplanned.

DR. NERENSTONE: Dr. Redman.

DR. REDMAN: Back to the issue of safety and tolerability, one-third of your patients dropped out before first analysis. How many of those patients had documented progressive disease?

DR. GEHLSON: Do you have a slide that can address that question? These are patients that had documented progressive disease, but were continued on therapy. That is not the question that he is asking. He is asking of the patients that were not evaluable at 12 weeks, how many of those were not evaluable due to progressive disease.

DR. REDMAN: Documented progressive disease.

DR. GEHLSON: Documented progressive disease.

DR. REDMAN: I guess while you are looking for that, I have another question regarding this endpoint of biology, because the information is confusing.

You have a treatment at least in the median number of cycles that you gave me per therapy were no different between the two arms, and I am assuming that that is the same for intent-to-treat liver metastases also, that the number of cycles between IL-2, IL-2/histamine are not different.

DR. GEHLSON: That is correct.

DR. REDMAN: The time to progression is not different using standard criteria for time to progression, the response rate is not different, it's inactive in both arms using response rate.

Why do you think other than data that is going to come out with the FDA review, patients lived longer with IL-2 and histamine?

DR. GEHLSON: I think it is a very reasonable and fair question. As you have heard earlier, there have been a number of studies that have improved the response rate without any improvement in survival, and so there really isn't a direct correlation between the improvement in survival and response rate to date, at least in this population, however, the opposite could be true.

We made survival the primary endpoint in this study because we knew that we could not reliably document responses. We did not set up a tumor board to look at responses in progressive disease because if we really are having lymphocyte infiltration in the tumor, can you determine a tumor that is actually progressing or one that is actually being attacked by the immune system. You can't do it.

So, we have always said that we could possibly better design a trial if we did survival as the primary endpoint. Of course, we have to look at response rates as a

secondary endpoint and look at time to disease progression, but if it can't be absolutely reliably measured in the patient population, we are left with survival, and we are.

We are asking you to think somewhat differently. I think the reason that there is an improvement in survival is the number of patients who have stabilized disease, and stable disease is a clinically meaningful benefit for the patient especially if you can do it in a setting like we have done. It may also allow them to get other additional treatments down the road if we could stabilize the rapid progresser.

I don't know if Sanjiv or Dr. Glaspy want to add to that. John, would you like to add a little bit?

DR. GLASPY: I am John Glaspy from UCLA. I think we need to add one more thing in terms of the delivery subset question, which we all have, because we want to know why if it works, it works in liver.

There is a second mathematical explanation which would be that the liver metastases patients represent a much more homogeneous group of patients, and where you have the power in a much smaller number of people to detect survival differences.

It may be that the biology isn't different, we just don't have the power in the intent-to-treat analysis, and that is another way I think about this.

DR. REDMAN: John, the converse could also be true, though. It is not a homogeneous population, but the population itself in the trial is heterogeneous.

DR. GLASPY: Yes, I agree.

DR. NERENSTONE: Dr. Taylor, did you have a question?

DR. TAYLOR: Well, I wanted to clarify the numbers you gave her were 42 and 45. Are those ineligible patients?

DR. GEHLSON: No, those are patients that were granted protocol exceptions for the reasons that I listed.

DR. NERENSTONE: Dr. Simon.

DR. SIMON: One last clarification. You indicated that the patients who were not evaluable for response or progression evaluation were included in the survival analysis. So, how many patients either died or you have follow-up information on, say, up to within six months of either the September update or the previous? I am trying to get an idea.

I mean it is one thing to say that the patients were included in the analysis, but if they were sort of included but censored a long time because you lost track of them, that is including them, but, you know, you can't really have confidence in the results.

So, I want to get an idea for the patients who have not been indicated as having died, how up to date the

follow-up is on those patients and what proportion of them you sort of lost contact with.

DR. GEHLSON: That is another excellent question. All patients that were randomized have been followed up for survival regardless of when they died, and those patients that were still alive on the data cutoff date, were actually identified and followed. There were no patients that were censored or lost to follow-up.

So, the same thing happened on September 8th, on the reanalysis, every single patient who had died between the previous analysis and the new analysis were actually followed, and all patients alive today will continue to be followed for survival, and we have no patients lost to follow-up in either one of those analyses.

DR. NERENSTONE: Dr. Przepiorka.

DR. PRZEPIORKA: It appeared from some of the tables that a substantial proportion of the patients had been treated before coming on the study, so there was an opportunity to observe them prior to study entry.

Could you tell us--and I don't know if you have collected this information either at the time of randomization or start of drug--the proportion of patients in each of the arms who had progressive disease versus stable disease?

DR. GEHLSON: That is a very good question, but

our entry criteria specified that they had to have progressive disease upon entry, so we assumed that all patients that entered this study had actually had progressive disease, and they had to have discontinued their prior therapy 14 or 30 days prior before being eligible for this study.

DR. NERENSTONE: Dr. Redman.

DR. REDMAN: Did you find an answer to the question about the third of patients who are not evaluable?

DR. GEHLSON: They are looking and we will find that answer, I promise.

DR. NERENSTONE: Dr. Albain.

DR. ALBAIN: You had just mentioned no patients had been lost to your follow-up, so what happened in that first three months is really critical to understand. The same question.

DR. GEHLSON: You mean the first three months before they were in the study?

DR. ALBAIN: No, the first three months from time zero to month 3, those who went off study.

DR. GEHLSON: Yes, and all those patients, of course, are included in the survival analysis as we mentioned, but we will find out if they have discontinued. I would imagine the majority of them discontinued for progressive disease, not for an adverse event.

DR. ALBAIN: And were they then scored as progressive disease--

DR. GEHLSON: Yes.

DR. ALBAIN: --if they had done so, even though they hadn't reached the first evaluation point?

DR. GEHLSON: Yes. In fact, I think that is evident here because if here is the first scheduled evaluation, you can see that--all patients, of course, are included in the analysis unless they were, well, in fact, in time to disease progression, all patients are included, so if they had progressed here or died, or died, then, they are included in here.

I think the majority of those patients actually progressed and died before the first scheduled evaluation. I think this is also why you see this artifact because the first scheduled evaluation was after two cycles, and you see the majority of the patients actually had the diagnosis of progressive disease at that first evaluation, but the patients were allowed to continue if they did have progressive disease, and we can tell you that there was no investigator bias because those patients that were continued with what might be progressive disease, and were allowed to continue on therapy, is essentially identical to the group, in fact, favors the IL-2 group.

More patients were continued on IL-2 with the

first diagnosis of progressive disease than on the histamine/IL-2 group. They were equally distributed between the two groups, so we can show you those data.

Actually, you could also just tell me the number of deaths before 12 weeks, because the number of patients are on the slide.

DR. DUTCHER: Just help me with this. The not evaluables are a subset, I mean are included in the progressive disease, some of them, or not some of them?

DR. GEHLSON: No.

DR. DUTCHER: They are not evaluated because they didn't reach the 12 weeks.

DR. GEHLSON: Yes, they didn't get their first response evaluation.

DR. DUTCHER: But you are saying they came off study because they progressed, most of them, or you think that most of them did.

DR. GEHLSON: You have to have a scan to know if the patients had progressive disease, if they are rapidly deteriorating and then they die before they actually get to the first evaluation, but that is why we need to know.

MS. WOOD: This is Diana Wood. When the patients are removed from study, they are required to fill out a study completion page. On that page, the reason for study discontinuation must be checked. If the patient was removed

from study due to progression of disease, that check box would be ticked.

Does that answer the question?

DR. REDMAN: And you don't have the raw data, you don't have the documentation of their x-rays and scans showing that.

MS. WOOD: That would be monitored by the site monitors as part of the monitoring conventions to ensure that if progression of disease was checked, that there should be some documentation either in the form of a scan, physical exam, or some other portion within the case report form that the patient is indeed progressing.

DR. REDMAN: So, you do have that data.

MS. WOOD: That data would be located within the case report form.

DR. REDMAN: It is an important issue if it comes down to safety, because metastatic melanoma patients with performance statuses of zero to 1, dying within 12 weeks of going on trial sort of negates the fact that performance status is an important criteria.

MS. WOOD: I think what we would like to do, if it's acceptable, is to provide you with that 12-week data of the patients that died and the patients that died due to progression of disease after the break, if that is acceptable.

DR. REDMAN: Yes.

DR. NERENSTONE: Dr. Sledge.

DR. SLEDGE: A question about the eligibility. This is a fairly astoundingly high rate of patients with enrollment violations, I mean in excess of 40 percent. I saw the listing of why they were granted exceptions, but I guess my question would be if you have such poor control over who enters the study, why should we have any great deal of confidence about the conduct of the study past that point.

DR. GEHLSON: Well, it is a fair question. I mean I can't directly answer that. I know that U.S. clinicians are very compassionate people, and when you identify a patient who may benefit from this particular treatment, if their scan is one day out of the three-week window, I don't think granting an exception for a patient like that is much of a problem.

Many of these patients had failed all prior therapy, so we found that our medical monitors and our clinicians were rather compassionate and wanted to include these patients, but I don't know if we should be penalized for those because if we would have had more forethought on the types of inclusion criterias, we could have broadened those earlier on, but we allowed all patients into the study that met the criteria if they were granted an exception by

the medical monitor, and it appeared to be appropriate by the medical monitor and the investigator, these patients are included in the intent-to-treat analysis.

DR. SLEDGE: I can understand why your physicians were being compassionate for their patients. I don't quite understand why your medical monitors were. I mean this is a truly extraordinarily high rate of ineligible patients by cooperative group standards.

DR. GEHLSON: Yes, but the reasons for, though, are relatively insignificant I would say clinically. If they missed a scan by a day or so, which is the majority of those ineligibility criteria, or they have one lab value off, I think that those things don't represent looseness of the medical monitor, but actually working together with the investigator to identify and try to help these patients and put them onto the study. We weren't afraid of having them in our study.

DR. NERENSTONE: Dr. Temple.

DR. TEMPLE: I suppose ideally, your medical monitor would have told you this was happening, so you could, if necessary, adjust the entry criteria. I mean you may be right, that being a day off is no big deal, but it sort of could suggest no one was watching closely.

DR. GEHLSON: I think you are absolutely right, and, in fact, we actually have seven amendments to the

protocol, and most of the majority of the amendments were related to criteria that the medical monitor and the investigators identified that needed to be changed.

So, during the course of the study, we made adjustments to the eligibility criteria to allow for these particular changes, and you would see that most of these patients who may have been granted exceptions were granted early on in the study before many of these changes took place. So, in fact, we were in direct communication with our medical monitor at all times.

DR. NERENSTONE: There being no more questions from the committee, I would like to call a break. Please be back at 3:50 to start the last part of the afternoon. Thank you.

[Recess.]

FDA Presentation

DR. CHIAO: Good afternoon. Dr. Nerenstone, members of ODAC, ladies and gentlemen, it is my pleasure here today to be part of the members of the FDA team presenting to you the FDA review of the New Drug Application of histamine dihydrochloride.

[Slide.]

I will review the clinical studies in the NDA. Following my presentation, FDA statistician Dr. Sridhara will go over some important statistical issues in the major

study. Finally, Dr. Donna Griebel, medical team leader, will address the regulatory standard for using a single study to support an NDA.

This slide lists the members of the FDA review team.

[Slide.]

The applicant is seeking marketing approval of the histamine/IL-2 combination for the treatment of patients with melanoma that has metastasized to the liver.

[Slide.]

This slide outlines the FDA presentation. The FDA review of the major study MP-US-MO1 will focus on the following three issues: Is survival difference in the intent-to-treat population a persuasive finding? Is survival difference in the liver subgroup a persuasive finding? Is the histamine/IL-2 combination a well tolerated regimen?

[Slide.]

As we have heard earlier, metastatic melanoma is a relatively chemo-resistant disease with very limited treatment options. This disease is known to have a very variable clinical course.

Publications over the past 20 years have consistent demonstrated that survival in this disease is influenced by prognostic factors as the ECOG investigators

pointed out in the recent analyses of 1,362 patients on ECOG melanoma trials. The prognosis for patients with Stage IV melanoma still seems to be influenced by disease factors than any particular therapy.

[Slide.]

This slide lists the major prognostic factors. Although patients with liver metastases have poor prognosis, the literature suggests that these patients also have a variable clinical course.

Two studies, one from Memorial Sloan-Kettering and one from ECOG, showed that the presence of liver metastases at study entry is not a significant predictor for poor survival in multivariate analyses.

[Slide.]

The Division met with the applicant in April 1997 to discuss the design of registration study for the histamine/IL-2 combination. The dose and schedule of histamine/IL-2 combination in the proposed MP-US-MO1 has never been tested in Phase I/II setting.

The Division felt there was a lack of safety data on the tolerability of the proposed treatment regimen and a lack of efficacy data to support the estimates for the proposed sample size.

Not convinced that the results of MP-US-MO1 alone would be compelling to support an efficacy claim, the

Division asked for two trials stating that additional studies might provide supporting evidence in case the results of MP-US-MO1 was weak.

The Division emphasized that the registration studies should be stratified by key prognostic factors to ensure the two arms were balanced in patient characteristics that might affect survival.

[Slide.]

Regarding the design of a second well-controlled study, the applicant proposed to do an international Phase III study which uses a different treatment regimen, that is, histamine/IL-2 plus alpha-interferon versus DTIC alone.

The Division stated that this trial could not support MP-US-MO1 because the treatment regimen was different. The Division also stated that the single arm study MA-0103 could not be used as second well-controlled trial because the added benefit of histamine to IL-2 could not be demonstrated and survival data would be difficult to interpret in the absence of a control arm.

[Slide.]

This slide shows the study design for MP-US-MO1. It is an open-label, randomized study which was not stratified by any prognostic factors. There was no stratification based on the presence of liver metastases at study entry and therefore, randomization within a stratum,

that is, in the liver subgroup was not carried out. The primary endpoint of this study was survival.

With 300 patients, the study had an 80 percent power to detect a 50 percent increase in median survival from 7.8 months to 11.3 months.

The following three slides summarize the statistical analysis plan concerning the liver subgroup.

[Slide.]

The original protocol dated July the 1st, 1997, stated the primary objective of the study is to evaluate the efficacy and safety of the histamine/IL-2 combination in patients with metastatic melanoma who have not been treated or have failed other first-line therapies.

Patients will be stratified in subgroup analyses, liver versus no liver mets; prior treatment with DTIC or no DTIC.

[Slide.]

The last patient was randomized on March 26, 1999. Revised statistical analysis plan in June 24, 1999, still stated that the intent-to-treat subset will be used as the primary subset. All efficacy endpoints will be summarized for non-exploratory subgroup of patients with liver metastases at study entry.

[Slide.]

The null hypothesis testing for the liver subgroup

was introduced in the final statistical analysis plan dated November 18, 1999, which was about a month after the study finished accrual and five months before the data was analyzed.

It is our view that survival analysis in the liver subgroup is not a prospectively defined primary analysis.

[Slide.]

The FDA review of study MP-US-MO1 will focus on three issues. The first one is whether survival difference between the two arms in the intent-to-treat population is a persuasive finding.

[Slide.]

Two cutoff dates were used for survival analyses. None of these dates were prespecified in the protocol. The update survival data using September 8th cutoff date was submitted about three weeks ago before this meeting.

Forty patients were alive on September the 8th. Thirty-three of these 40 patients had last follow-up date 30 days or more beyond the cutoff date. Four deaths occurred after the cutoff date, and three were on the histamine/IL-2 arm.

[Slide.]

This slide shows median survival at different times. The first two days were used by the applicant, and the last one was by FDA only. At March 8, 2000, cutoff

date, 243 patients have died. That is, 80 percent of the events have occurred. By September 8th, 2000, an additional 22 patients have died, an increase of 7 percent events. Four more patients died if the last follow-up date of all patients were used. Regardless of what dates were used, the difference in median survival between the two arms is about one month.

[Slide.]

This slide shows the hazard ratio and log-rank p-value of the three survival analyses. Hazard ratio is the risk of death on the treatment arm in this trial is the histamine/IL-2 arm divided by the risk of death on the control arm, a measure of relative treatment efficacy of the histamine/IL-2 combination.

A lower hazard ratio indicated a lower risk of death on histamine/IL-2 arm. As you can see, the hazard ratio from all three studies are around 0.8. That is, patients on histamine/IL-2 arm had about 20 percent less chance of dying when compared to patients on IL-2 alone arm.

It is interesting to note that the log-rank p-value is the smallest at the September 8th cutoff date. Overall, these hazard ratios were very similar indicating that there was no significant change in the survival difference between the two arms with longer follow-up.

[Slide.]

This is the Kaplan-Meier curve of survival analyses using the March 8th cutoff date.

[Slide.]

This is the survival curves using the September 8th cutoff date.

[Slide.]

This is the survival curve using the most recent follow-up data.

There is not much difference between these three analyses in terms of Kaplan-Meier curves.

[Slide.]

All these p-values I just showed you are not adjusted for multiple comparisons.

[Slide.]

We performed subgroup analysis to look for consistency, which we did not find.

[Slide.]

We looked at three subgroups which are listed on this slide. 129 patients had liver metastases at study entry. These patients represent 42 percent of the study population. Median survival is 5 months on IL-2 arm and 9.2 months on histamine/IL-2 arm, strongly favoring the histamine/IL-2 arm with p equals 0.0033.

176 patients had no liver metastases at study entry. These patients represent 58 percent of the study

population. Median survival is 10.3 months on the IL-2 arm and 8.7 months on the histamine/IL-2 arm, favoring the IL-2 arm, but this does not reach statistical significance.

We then look at the subgroup of patients with skin/lymph node/lung only disease. These patients are considered to have good prognoses. The median survival of patients with skin/lymph node only disease is about 7 months although there is wide variability. Median survival for patients with lung involvement in the literature is about 11 months.

82 patients, that is, about 27 percent of the study population have skin/lymph node/lung only disease. Median survival is 12 months on IL-2 arm, and 10.4 months on histamine/IL-2 arm favoring the IL-2 arm, but this does not reach statistical significance.

[Slide.]

This is the survival curves in the liver metastases subgroup, 129 patients.

[Slide.]

This slide shows the survival curves in 176 patients with no liver metastases.

[Slide.]

This is the survival curves in 82 patients with skin/lymph node/lung only disease.

[Slide.]

We looked for supporting evidence from tumor response rate and time to progression, but did not find any.

[Slide.]

The overall tumor response is very low, 3 percent on each arm. By FDA assessment, here was only one CR, which is on IL-2 arm. The other two CR's reported by the applicants were not confirmed by appropriate imaging studies.

[Slide.]

FDA assessment of time to tumor progression differed from the applicant's in 223 patients out of 305. Sixty-two patients had no follow-up scans after baseline studies, 34 were on IL-2 and 20, histamine/IL-2. These patients were not included in the FDA time to tumor progression analyses.

Other differences in the methods used to determine time to progression are listed here. FDA censored time to progression on the date of the last imaging studies submitted in the dataset. The applicant censored time to progression on the last day when the patients were known to be alive, which in some cases was the cutoff date for survival analyses.

FDA did not count death as progression unless progression was documented. All deaths were counted as progression by the applicant.

For patients who progressed, FDA based the question date on the date of the imaging studies documenting the progression. The applicant based progression on the date when the imaging study was read.

[Slide.]

This slide shows the FDA analysis of time to progression in 243 patients. There is no difference between the two arms.

[Slide.]

In summary, we have the following comments on the efficacy results in the intent-to-treat population. Survival difference in the ITT population did not reach statistical significance.

P-value dependent upon cutoff dates, which were not prespecified in the protocol.

Lack of internal consistency across subgroups.

No supporting evidence from tumor response or time to tumor progression.

[Slide.]

The applicant is seeking marketing approval of the histamine/IL-2 combination in patients with melanoma that has metastasized to the liver. The efficacy claimed is based on results in a subgroup of 129 patients with liver metastases at study entry.

The second issue is whether survival difference in

this subgroup of patients represent persuasive evidence of efficacy.

[Slide.]

MP-US-MO1 is not stratified by the presence of liver metastases at study entry. Therefore, no stratified randomization was performed in the liver subgroup. There were many imbalances in prognostic factors consistently favor the histamine/IL-2 arm.

[Slide.]

This slide listed the major imbalances between the two arms in this subgroup. Except elevation of LDH, the imbalances consistently favors the histamine/IL-2 arm. There is a striking imbalance in a number of metastasis sites between the two arms. The histamine/IL-2 arm has almost twice as many patients with one site of disease as compared to the IL-2 alone arm.

Having one metastasis at a site is considered a favorable prognostic sign. In the series from University of Alabama reported by Dr. Balch, et al., median survival of patients with one single distant metastasis site was 7 months, with two sites, 4 months, and with three sites, 2 months.

Recent publication from M.D. Anderson by Ethan, et al., emphasizes that single visceral organ involvement does not have the same poor prognosis as multiple visceral organs

were involved even when the most favorable group of patients, that is, those with lung metastases alone, are excluded.

[Slide.]

We don't have complete information on prognostic factors. For example, preceding stage of tumor is a prognostic factor, but the study did not collect this information. Information was collected on metastases evaluable at the study entry, but not on those at first diagnosis except the diagnosis date.

When we compared the baseline performance status, which was obtained prior to randomization to that on the first day of treatment, that is, cycle 1, day 1, we found that 58 patients had a decrease in performance status, 33 on IL-2 arm and 25 on histamine arm including 18 patients who had a deterioration to performance status 2 or worse.

Within the liver subgroup, 28 patients had a decrease in performance status including WHO-4, who had a PS 2 or worse on the first day of treatment.

[Slide.]

Using what we had, we performed adjusted analysis to account for the detected imbalances. The performance status using a model is based on performance status because 47 patients did not have performance status on cycle 1, day 1.

The FDA statistician, Dr. Sridhara, will go into more details on this later, but briefly, the covariates included in FDA models are slightly different than those included as sponsor's model, and the details will be addressed later.

After adjusting for these imbalances, both hazard ratios and the p-values shifted indicating that these imbalances contributed to the observed survival difference in the liver subgroup.

[Slide.]

This slide showed the hazard ratio and p-value in the FDA adjusted analyses. It is our view that because of the large shifts in hazard ratio and p-value, one cannot reliably estimate the magnitude of treatment effect by the histamine/IL-2 combination in this subgroup of patients.

[Slide.]

A closer look at the liver subgroup revealed that the overall survival difference is the most apparent in a small group of 20 patients with only one site of metastases, that is, liver only disease.

[Slide.]

As shown on this slide, there was a very large difference in median survival favoring the histamine/IL-2 arm in 20 patients with liver only disease.

However, we must note that there were twice as

many patients on histamine/IL-2 arm as compared to IL-2 arm. As I mentioned earlier, metastasis is a prognostic sign and there were more patients with this prognostic sign on the histamine/IL-2 arm making it impossible to assess. With patients with disease both inside and outside the liver, the treatment effect of histamine/IL-2 combination is less apparent.

[Slide.]

Again, we looked for supportive evidence of tumor response and time to tumor progression, but did not find any. There were two partial responses occurred on the histamine/IL-2 arm with a response rate about 4 percent.

[Slide.]

Follow-up available in 96 patients, about 74 percent of the liver subgroup, by FDA analysis there is no statistically significant difference in time to progression between the two arms in the liver subgroup.

[Slide.]

Our comments on efficacy in the liver group are listed on this slide. Imbalances in prognostic factors consistently favor the histamine/IL-2 arm. These imbalances contributed to the observed survival difference between the two arms by FDA adjusted analysis.

Treatment effect of the histamine/IL-2 combination is most apparent in the small group of patients with liver

only disease.

No supporting evidence from tumor response and time to tumor progression.

[Slide.]

The third issue that we are going to address, is the histamine/IL-2 combination a well tolerated treatment regimen?

[Slide.]

As shown on this slide, the majority of patients completed no more than three cycles of therapy. Most of the patients dropped out because of progression of disease.

[Slide.]

The histamine/IL-2 regimens are administered in the outpatient setting except the first two doses of IL-2, which was given in the clinic. Patients on IL-2 arm need to self-administer subcutaneous injections twice a day for two days during week 1 and week 3, and twice a day for five days during week 2 and week 4.

[Slide.]

Patients on histamine arms must self-administer subcutaneous injection of histamine twice a day for five days for four weeks in addition to the IL-2 injections.

[Slide.]

According to the applicant, vials or pre-filled syringes were not returned to study sites to check for

compliance. Patient compliance was assessed by diary. But a lot of patients didn't complete diary because the diary was implemented during the middle of the trial. Dosing information was available in some patients by other source such as home care records.

Overall, we do not believe that patient compliance with medication was adequately assessed in this trial.

[Slide.]

This slide listed safety information on MP-US-MO1. Based on the applicant's dataset, about 14 percent of patients required dose reduction for adverse events. We could not tell from the database whether the dose reduction was for IL-2 or for histamine or for both.

Fifty-four to 60 percent of patients experienced Grade 3 to 4 toxicity, and 11 percent died within 30 days of the last dose of study medication. The numbers are slightly higher in patients with liver metastases.

[Slide.]

Thirty-three patients died within 30 days of last dose of study medication. Three deaths were attributed to study medication by the applicant and all three were on the histamine/IL-2 arm. We could not exclude toxicity-related death in 12 patients based on review of the patient narratives.

[Slide.]

Specific types of Grade 3 toxicities are listed on this slide. There toxicities have been described in IL-2 trials. More patients on histamine arm had Grade 3 headaches. Otherwise, there was no excessive Grade 3 toxicities on the histamine/IL-2 arm.

[Slide.]

About 5 to 7 percent of patients suffered Grade 4 toxicity. The incidence of any specific type of Grade 4 toxicity is rare. Again, these toxicities have been described in IL-2 trials.

[Slide.]

MA-0103 is an ongoing single arm study using the same histamine/IL-2 regimen as in the major trial MP-US-MO1. The applicant submitted updated adverse events data on 90 patients. Of these 90 patients, 16 died within 30 days of the last dose of study medication, and 8 of these 16 deaths were in patients with liver metastases at study entry with an incidence rate of 23 percent.

The incidence of death is higher than what was reported in MP-US-MO1.

[Slide.]

Grade 3 toxicity occurred in 54 percent of patients, very similar to what was observed in MP-US-MO1. No Grade 4 toxicity was reported. Incidence of specific type of Grade 3 toxicity was also very similar to what was

reported in the MP-US-MO1 except chest pain, which was higher in the single arm trial.

[Slide.]

The next few slides listed FDA overall summary. Only one randomized study was submitted in this NDA to support the marketing approval. The efficacy of the IL-2 regimen used in this trial is not known.

[Slide.]

The survival difference in the intent-to-treat population did not reach statistical significance.

Survival difference in the subgroup of patients with liver metastases should be interpreted with caution. Imbalances in prognostic factors favor the histamine/IL-2 arm.

It is our view that effect of these imbalances precludes the reliable assessment of the efficacy of the histamine/IL-2 combination, and there is no supportive evidence from tumor response rate or time to tumor progression.

[Slide.]

Safety review showed that 58 percent of the patients suffered Grade 3-4 toxicities and 11 percent died within 30 days of the last dose of study medications.

I included here the reported safety data in the most recent DTIC trial to show that incidence of Grade 3 and

4 toxicities could be as high as 36 percent in patients who were treated witn this drug which is considered to be well tolerated.

I mean this is to make a point that it is not possible to adequately assess the incidence of drug-related toxicity in the absence of a non-IL-2 arm.

This concludes my presentation, and I think Dr. Sridhara will come up here and to address some important statistical issues in a major study MP-US-MO1.

DR. SRIDHARA: Good afternoon. I am Rajeshwari Sridhara, statistical reviewer of the study being presented here.

[Slide.]

I am going to talk about the major statistical concerns in the randomized M01 study focusing on the efficacy aspect of the study. The major concerns can be categorized into two broad areas, namely, the overall finding in the ITT population, and the second one, liver metastasis subgroup finding.

[Slide.]

There are three specific problems in the overall finding. Problem 1 is regarding evidence of efficacy. Problem 2 is regarding multiple survival analyses. Problem 3 is internal consistency. I will be focusing on problems 1 and 2 as problem 3 has already been addressed by Dr. Chiao.

[Slide.]

Let us look at the problem of evidence of efficacy. First, the original protocol did not specify time of final analysis date done even on a specific data cutoff date. However, in the statistical plan submitted by the sponsor in November of '99, an arbitrary cutoff date of March 8, 2000 was chosen for final analysis. As for this date and NDA submission, there is no difference between the two treatment arms with a p of 0.1255 and a difference in median survival of 0.9 months.

Note that the data by this March date was reasonably mature with approximately 80 percent occurrence of even. Even in subsequent analysis with updated data, unadjusted p-value has remained above 0.05 level with 88 percent of even already occurred.

[Slide.]

The second problem refers to multiplicity issues. This graph illustrates that repeated analysis result in different p-values. This emphasizes the importance of prior specification of time of final analysis based on number of even.

Per sponsor's statistical plan, March 8th should be technically considered as the final analysis date. It appears that the nadir p-value was reached with September 8th data cutoff date.

The red line represents the 0.05 alpha level and the green line represents 0.042 alpha level, which was the alpha allocated for final analysis in ITT population per protocol. Note that all p-values, that is, the blue diamonds, are above 0.05. The x axis here is not to be scaled by time, and the main point of this illustration is that these p-values are not reliable.

[Slide.]

Regarding covariate adjusted analysis in the ITT population, the International Conference on Harmonization Guidelines states that, "When the potential value of an adjustment is in doubt, it is often advisable to nominate the unadjusted analysis as the one for primary attention, the adjusted analysis being supportive."

It further continues to state that, "In most cases, however, subgroup and interaction analyses are exploratory and should be clearly identified as such; they should explore the uniformity of any treatment effects found overall."

We believe that the ITT population was randomized and therefore we base our conclusion on unadjusted analysis, and the adjusted analysis are used only to check for internal uniformity and consistency of results in the ITT population.

[Slide.]

In summary, the one randomized MO1 study conducted by the sponsor in metastatic melanoma patients does not demonstrate convincing evidence of survival benefit with histamine plus IL-2 treatment in the overall ITT population.

[Slide.]

Moving to liver metastasis subgroup finding, there are two major problems in the liver metastasis subgroup finding. Problem 1 is the absence of overall survival benefit, and problem 2 is imbalances and multiplicity issues.

[Slide.]

In the absence of overall survival benefit, any subgroup advantage is questionable. Again, the ICH Guideline states that, "These analyses are not intended to salvage an otherwise non-supportive study but may suggest hypotheses worth examining in other studies or be helpful in refining labelling information, patient selection, dose selection, et cetera."

The common standard for strength of efficacy used is p-value of less than 0.05 in two, well-controlled, randomized studies. If the result of the single study considered here is acceptable as sufficient evidence, then, we should evaluate an interaction effect between treatment and covariate. If the interaction is found to be statistically significant, then, the overall results are not

interpretable and subgroups have to be studied individually.

In such a case, one can consider subgroup analysis. Since the sponsor presented Table 11b of my review, I would like to clarify that this model was considered to test for interaction and find a rationale for further subgroup testing.

This model was not intended to evaluate the treatment effect in either of the subgroups, that is, either the liver met subgroup or the non-liver met subgroup. A subgroup analysis was further performed in my review, as well.

[Slide.]

However, in the study, stratified randomization within liver subgroup was not done. Imbalances favoring histamine plus IL-2 arm in the distribution of patients are observed. Furthermore, liver metastasis subgroup hypothesis testing was added on to the original protocol after the study had completed enrollment.

No allocation of alpha for testing liver metastasis subgroup hypothesis was planned prior to the start of the study. A statistical plan was submitted by the sponsor prior to NDA submission with a plan for post-hoc adjustment of type 1 error for testing liver metastasis subgroup hypothesis.

[Slide.]

In this graph, red bars represent histamine arm and blue bars represent IL-2 arm. The message to be taken from this bar graph is that the red bars, which represent the histamine arm, have higher percentage of patients in all the better prognostic subgroups. For example, patients with less than 65 years of age, female, performance status of zero, no prior chemotherapy, et cetera.

Thus, this bar graph illustrates that the imbalances observed favors the histamine arm. It should also be noted that of the 14 characteristics presented in this graph, 13 of them favored histamine arm.

[Slide.]

Because of these imbalances in the liver subgroup, it is appropriate to further evaluate treatment, in fact, using covariate adjusted analysis. Sponsor has submitted models which are different models with different data cutoff dates. They are also different from protocol specified covariate and covariate specified in the statistical plan.

FDA has used consistent model at all times and no selection was considered, that is, all characteristics identifiable with imbalances were included in the model.

[Slide.]

In the original model, only two factors were prespecified, presence or absence of liver metastasis, and secondly, whether the patients received prior chemotherapy

or not. However, subsequently, an amended statistical plan was submitted in November of '99 with the covariates listed in column 1, namely, treatment, age as continuous variable, sex, race, number of prior anti-cancer therapy, metastatic disease sites categorized as skin, lymph, non-visceral versus lung, GI, kidneys, adrenals versus liver, bone disease, number of metastatic sites, LDH as a continuous variable, prior chemotherapy as a continuous variable, baseline performance status, and centers grouped into Midwest, North, West, and South, and each region being compared to South, not sure of reason behind this particular comparison.

In the model submitted by the sponsor in the NDA application, with the follow-up data up to March 8th, only 6--this is the second column I am talking now--only 6 of the 11 specified covariates were considered.

The covariates with asterisks were not prespecified in the statistical plan or were not as prespecified in the statistical plan. For example, the number of prior anti-cancer therapy and LDH were used as categorical variables instead of continuous variables as specified in the plan.

Also, only bone metastatic disease site was used in the model. The model submitted by the sponsor subsequently in November with updated data considered all

factors although some of the continuous factors were categorized.

The last column lists the factors included in the FDA exploratory model. All factors with imbalances are included in this model. Notice that three more additional factors - baseline albumin, disease-free survival from initial diagnosis to primary metastasis, and time from primary metastasis to randomization were considered because of the imbalances identified in these characteristics.

[Slide.]

Survival analysis adjusted for all covariates with imbalances illustrated in the earlier bar graph were included in these analysis. In this table, I am not listing the covariates which went into the model because I have included all the covariates that had imbalances, and only giving the treatment effect here.

With the follow-up data up to March 8th, the p-value for treatment effect for adjusting for factors for imbalances was 0.1070 with updated follow-up data up to September 8th using the same model, the treatment effect after adjustment for imbalances was 0.0573.

Note that the p-value has dramatically increased compared to the corresponding unadjusted p-value of 0.0033. It should also be noted that this model does not include all known prognostic factors in malignant melanoma.

Furthermore, the p-values presented in the table are not adjusted for multiplicity, meaning multiple hypotheses or multiple analysis.

For example, adjusting only for two hypotheses, one in the ITT population and one in liver metastasis subgroup, the p-value is 0.1146 with the updated data.

[Slide.]

Therefore, the take-home message is that the adjusted model results are sensitive to inclusion and exclusion of a covariate. They are also sensitive to whether a covariate is used as a continuous variable or a categorical variable.

More importantly, it should be kept in mind that these p-values are not adjusted for multiplicity either for multiple subgroups or multiple analyses.

Thus, there is no robustness in the liver metastasis subgroup finding, and it is not possible to assess the true treatment effect in this subgroup given the imbalances from one single open label study.

Dr. Griebel will continue the presentation addressing further the single study shift.

DR. GRIEBEL: Good afternoon. My name is Donna Griebel and I am the medical team leader. We are here today considering an application that has as its foundation a single randomized controlled trial, and my job on behalf of

the FDA is to briefly introduce the perspective on how important it is to keep in mind the limitations a single trial has for allowing us to draw efficacy conclusions.

[Slide.]

It was not until the Federal Food, Drug, and Cosmetic Act of 1962 that applicants were even required to demonstrate that their product was effective as claimed. In this Act, Congress required that applicants provide substantial evidence of effectiveness, and they defined that substantial evidence in the Act as data that came from adequate and well-controlled investigations, and importantly, the word "investigations" is plural, and the Agency required from that time on that sponsors submit more than one randomized controlled trial in support of the effectiveness claim.

However, in 1997, the Modernization Act, Congress gave the FDA some leeway to in some circumstances consider data from an adequate and well-controlled single trial as substantial evidence.

[Slide.]

The reasons why a single study has generally not been considered adequate are listed in the Agency's guidance to industry for establishing clinical effectiveness, and these include any trial may be subject to unanticipated, undetected, systematic biases that could lead to flawed

conclusions.

There is inherent variability within biological systems that may produce a positive result by chance alone. So, a positive result that may be thought due to treatment effect may have occurred because of variability in the disease across the study population or variability in baseline characteristics of the participants in the study.

Applying the traditional p-value of 0.05 with a two-tailed test, 1 in 40 studies of ineffective drugs will be positive. One in 40 does not sound like a very alarming proportion, but if you evaluate this proportion in view of just looking at positive trials and looking at the proportion of a group of positive trials that could actually be false positive, factoring in the background probability that you are actually testing an effective drug, you can come up with a more hypothetical situation where you end up with a more alarming proportion, where you could have as much as one-third of those positive trials actually being false positive trials.

[Slide.]

This concept has been described both by Dr. Simon in his chapter in DeVita's Textbook of Oncology and recently in a paper by Dr. Ian Tannock in a supplement to JCO just last month.

In this paper, Dr. Tannock listed the causes of

false positive trials. They include mere chance alone, prognostic factor imbalances that can happen just by mere chance alone, multiple analyses of multiple endpoints, and the concept that I just mentioned, low prior probability that a new treatment will be a therapeutic advance.

[Slide.]

This situation is where you are testing many drugs, very few of which are actually effective, and this seems like a situation which is applicable to the field of oncology today.

The example to clarify this in the paper that Dr. Tannock gave is very similar to the one that Dr. Simon gave in the chapter in DeVita. If you start with 200 trials and you assume that you are testing in 10 percent of those trials a truly effective drug, 20 of those trials will be true positive trials.

In the remaining 90 percent of the trials that you are conducting, you are testing ineffective drugs, so you have 180 of the 200 trials that are true negative trials.

If you go back to your column of the 20 true positive trials, and you have powered your trials 90 percent power to detect the effect, you are going to miss the true effect in two of those trials, 10 percent.

So, you are left with what is down in the lefthand corner, 18 true positive trials where you actually detected

the effect. Going back up to your 180 true negative trials, if you apply the 5 percent p-value of 0.05 in a single tail this time, you come up with 9 trials that are false positive.

So, in the end you have 18 plus 9, or 27 positive trials, but 9 of those 27 trials are false positive. Nine into 27 is one-third. One-third of your trials are false positive, and that is a more alarming proportion than 1 in 40 that we just mentioned.

You can see from this that if you bump up the testing of the truly effective side to, say, 50 percent, you will bring down that proportion. So, if you use 50 percent true positive trials, you will end up with approximately 5 percent of your positive trials being false positive.

[Slide.]

Well, unreproducible positive trials do happen in the area of oncology and have occurred in the setting of metastatic melanoma. Dr. Atkins brought up the tamoxifen history in his presentation for the applicant, and in 1992, an Italian oncologist group presented in a New England Journal article the data from a trial that examined the introduction of addition of tamoxifen to dacarbazine, and they found a survival benefit that was statistically significant favoring the tamoxifen arm.

There was a more striking survival benefit in the

subgroup analysis of that study in the female population of the study. Subsequent trials examining the addition of tamoxifen to chemotherapy or biochemotherapy have not reproduced that statistically significant survival benefit.

[Slide.]

So, what does the FDA's guidance on establishing clinical effectiveness say about relying on a single trial? Well, it says that if you do this, it leaves little room for study imperfections or non-supportive information. It should be limited to situations where it is either impractical or ethically impossible to repeat the study to confirm the result.

[Slide.]

The guidance says that the characteristics of a single trial that can stand a substantial evidence is a large multicenter study, a trial whose design is appropriate, the conduct is flawless, and there is minimal possibility of bias due to baseline imbalances, unblinding, and post hoc changes in analysis.

[Slide.]

The results should reflect a hypothesis that has been documented in the protocol and the results should be statistically persuasive.

[Slide.]

So, I return to the single randomized, controlled

trial that is the foundation of this application and ask has this trial met the bars that have been set in the guidance.

[Slide.]

Well, this study was a randomized, controlled trial, it was a multicenter trial, but it is a single trial. The intent-to-treat analysis of the primary endpoint was not statistically significant.

[Slide.]

There appeared to be a survival benefit in a subset analysis of the patients with liver metastases, but there were imbalances in prognostic factors in this subgroup that favored the histamine/IL-2 arm, and the FDA's adjusted analyses showed that imbalances influenced the observed treatment effect.

[Slide.]

So, I would like to briefly review our questions. We have five questions. The last two are approvability questions. We have the safety question and we have two on whether the study can stand a substantial evidence, and when we refer to substantial evidence, we are talking about whether the single trial can stand alone as substantial evidence or whether it should be, if it is not substantial evidence, that infers that it needs support of a second trial.

So, the first question is a question about the

intent-to-treat analysis of the entire study population.

Does the survival difference in the planned primary analysis of the study, the intent-to-treat analysis of survival, represent substantial evidence of the efficacy of histamine as an adjunctive treatment for IL-2 for patients with metastatic melanoma?

[Slide.]

The second question again asks the question about substantial evidence, but this time in reference to the subgroup of patients with liver metastases in the single study and whether it represents substantial evidence of efficacy as an adjunctive treatment to IL-2 for patients with melanoma that has metastasized to the liver.

[Slide.]

Finally, I will end with our safety question. We ask whether the safety profile is acceptable.

We would be happy to answer questions, but I would like to rest of the review team to come up and join me.

Questions from the Committee

DR. NERENSTONE: Does the ODAC Committee have any questions specifically for the FDA reviewers?

[No response.]

DR. NERENSTONE: Thank you.

I would like to then open the discussion up to ODAC. Dr. Pelusi has a question.

Committee Discussion and Vote

DR. PELUSI: I realize we are going to be talking about some of our significant concerns regarding the design, the implementation, and the evaluation of this study. One aspect that I don't want to get lost, when we look at trials especially for metastatic melanoma, is this whole issue of quality of life.

I think that it is very difficult to pick an appropriate scale, but the problem is, is if we pick scales that we know probably will never be completed or that we only have maybe some baseline, that doesn't tell us a lot.

We also already know that one of the biggest impacts of this disease is on the family members, and we expect the family members to provide a lot of that care whether it is with this drug or with others.

So, the question is, as from the very beginning, do we look at quality of life of the family and in relationship to what happens when we put them on these particular studies, because as providers, when we have to look at offering treatment or not, the question is, is the safety, is it safe in that home and what goes into making that judgment, and do we have that kind of information, and obviously we didn't see that here, but I think that that becomes very important.

Also, the issue of many of these patients seem to

have home health support, as well as wonderful education from nurses, and I think that this is a very important part is again when we look at safety, we have to think about what is the real world use of medications that are approved, and the question again becomes do people have the ability to always have a home health person go in, how do we look at the compliance of how often the drug is actually taken or documented in terms of side effects at home versus what just happens in the office.

Again, not in relationship to this study, but these are the issues that I think we really need to address on every particular drug because this is a disease, it is a disease that not only affects the patient, but the entire family, and sometimes when we say let's offer something just to offer it, again, what is it that we are offering, and then I think that again the drug design and stuff will take us in that direction.

DR. NERENSTONE: Dr. Dutcher and Dr. Redman were asked to review and I wanted to perhaps ask if they wanted to start with some comments.

DR. DUTCHER: Before we start that, I wonder if there is any more information on those patients in the first three months of treatment that fell off the curve before the 12-week evaluation.

DR. NERENSTONE: We are asking the sponsor to go

back to those patients, Dr. Redman's question from earlier this afternoon.

DR. GEHLSON: Yes. Before cycle 2, there were 37 deaths, 27 were due to progressive disease, 17 on IL-2 and 10 on the histamine and IL-2 group. Ten of those deaths were due to other causes, 7 on histamine/IL-2 and 3 on IL-2, and we have discussed those.

That leaves the remainder of those patients. There were a number of patients who withdrew consent or withdrew for any other reason and for progressive disease, and that is where we are today on that.

DR. NERENSTONE: Dr. Albain also had a question for the sponsor.

DR. ALBAIN: I wondered if you could comment on the reversal, so to speak, of the significant p-value between your assessment of time to progression and the FDA assessment of time to progression.

DR. GEHLSON: Yes. We can never anticipate what the FDA has done. Because we didn't have a first evaluation until 12 weeks of patients, did not have an evaluation or a scan within that period of time, but they died due to their melanoma, we considered death due to melanoma as progressive disease, because their previous scan may have actually been all the way back to baseline at that point, so there could be 15 some-off weeks before they had a scan, so that would

make a big difference in the data because ours was actually the date of death or the data of a confirmed progressive disease.

Those are really the only differences that I could identify. We haven't had a lot of time to discuss all of the patients that they have included or excluded or all of the fine details, but we certainly will do that with them when this is done, but that is one I know for sure.

DR. NERENSTONE: I am going to open it up for discussion then for anyone with any other comments or questions.

DR. DUTCHER: And then to comment on the statistical issues that were raised by FDA.

DR. SIMON: Well, I believe very strongly in the importance usually of two clinical trials, and the basic reason being I mean most of the clinical trials in cancer that are done turn out to be negative even the large ones, and so consequently, of the ones that are statistically significant, a high proportion of them are false positives.

If we are not going to be making available to the public a lot of ineffective and toxic treatments, then, the only way really to guard against that is to have the two trial rule. I think the situation is even more extreme when you start talking about subset analyses.

Regardless of whether the subset was defined in

advance or not defined in advance or whether defining it after accrual is completed is far enough in advance, it is still a subset, and it wasn't really the primary focus of this study. Had it been the primary focus of the study, the study would have been on patients just with liver disease and it would have been sized appropriately.

Consequently, I think if you approve drugs based on one trial and significance in a subset, even with adjustment for the multiple comparisons, you would be approving a tremendous number of false positives.

I think for this particular trial, I personally find it very hard to have any credibility in my mind in the results. There is just too many imbalances that I just don't understand.

I would like to believe the results and I hope the results are true, and I really hope that a second trial can be done and we can find out really whether it is true, but given all of the imbalances and ineligible patients and inevaluable patients and large numbers of patients dying in the first two months, I certainly don't find this even approaches convincing evidence of effectiveness overall or for the subset.

DR. NERENSTONE: Dr. Redman.

DR. REDMAN: I agree with your comments especially regarding the subset analysis. I know of no other drug--and

I am open for comments on it--cytotoxic, biologic, hormonal, that is approved for a specific disease site, and not a disease itself.

I know there have been multiple trials and subset analysis that have found benefit for certain patient populations for treatment, but overall, patient population did not derive a benefit from that intervention, and I don't know of any drug again that is approved for that.

I think the study that was done overall and the primary analysis, the intent-to-treat showed no benefit or no difference between the two treatment arms. It is hard to say that there is any benefit even from the IL-2 arm itself.

If we want to address this, I think regarding approving something for a single disease site, I think we have to have an excellently controlled trial to answer that question before we start going down that slope of approving drugs for specific disease sites.

I don't think this trial met that criteria.

DR. NERENSTONE: There being no further comments, maybe we can address the question.

Just to briefly review, we have talked about the study MP-US-MO1 with the summary table, summarizes the FDA's analysis of efficacy.

The first column, the ITT IL-2 survival of 8 months median versus the histamine/IL-2. 8.9 months with a

p-value of 0.0526. The subgroup liver metastasis IL-2, 5 month median survival.

The next box is the subgroup with the histamine/IL-2 with the median survival of 9.2 with a p-value of 0.0033. The non-liver metastasis subgroup, IL-2 alone of 10.3 versus histamine/IL-2 of 8.7 median with a p-value of 0.7808. With the FDA response rate between zero and 5 percent across the six categories.

The time to progression, the FDA, a p-value of 0.4108 in the entire group. The subgroup analysis, the liver metastasis 0.1315, and the non-liver metastasis 0.9075.

The FDA's analysis of safety have been reviewed with the IL-2, 10 percent death within 30 days of the study medication; the histamine/IL-2, 11 percent death within 30 days of study medication with a 5 versus 7 percent Grade 4 and 59 versus 52 percent Grade 3 toxicities.

In addition, there are subgroup analysis in the liver metastasis, IL-2 versus histamine/IL-2.

Discussion. Metastatic melanoma can have a variable clinical course influenced by prognostic factors. As they outlined, there were many imbalances in the known prognostic factors and other patient characteristics between the two treatment arms in a subgroup of patients with liver metastasis because there is no stratified randomization.

substantial evidence of the efficacy of histamine dihydrochloride as an adjunctive treatment with IL-2 for patients with melanoma that has metastasized to the liver? In other words, the subgroup with liver metastasis.

Discussion?

[No response.]

DR. NERENSTONE: All those people who think it does represent substantial evidence of efficacy in that subgroup?

[No response.]

DR. NERENSTONE: Opposed?

[Show of hands.]

DR. NERENSTONE: No abstentions.

We haven't concluded that efficacy has been shown, so we don't need that one.

In view of the efficacy and safety data presented, should histamine dihydrochloride be approved for adjunctive use with IL-2 in the treatment of patients with metastatic melanoma?

Discussion?

[No response.]

DR. NERENSTONE: All those who think it should be approved, please raise your hand.

[No response.]

DR. NERENSTONE: Opposed?

[Show of hands.]

DR. NERENSTONE: And the last question. In view of the efficacy and safety data presented, should histamine dihydrochloride be approved for adjunctive use with IL-2 in the treatment of patients with melanoma that has metastasized to the liver?

All those in favor?

[No response.]

DR. NERENSTONE: All those opposed?

[Show of hands.]

DR. NERENSTONE: Does the FDA have any further questions from the group?

[No response.]

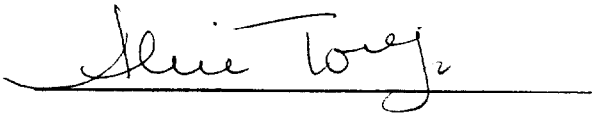
DR. NERENSTONE: Then, we stand adjourned, and we will reconvene tomorrow at 8:00 a.m. no matter what the weather.

[Whereupon, at 5:00 p.m., the proceedings were recessed to be resumed at 8:00 a.m., Thursday, December 14, 2000.]

- - -

CERTIFICATE

I, **ALICE TOIGO,** the Official Court Reporter for Miller Reporting Company, Inc., hereby certify that I recorded the foregoing proceedings; that the proceedings have been reduced to typewriting by me, or under my direction and that the foregoing transcript is a correct and accurate record of the proceedings to the best of my knowledge, ability and belief.



ALICE TOIGO

6

7

8

9

A

B

C

D

E

F

G

H

I

J

K

L

M

N

O

P

Q

R

S

U

V