of each study.

DR. TEMPLE: So, they weren't contributing equally. So, that must be dominated by study 63.

DR. GILMAN: And 65. Dr. Van Belle?

DR. VAN BELLE: Sorry, I really don't like to interrupt a presentation but since the issue has been raised, one of the issues that I have with the integration is that there is, in fact, no homogeneity of results. In other words, you can show statistically very easily throughout the studies that the frequencies are not comparable and, therefore, the pooling may not have been legitimate. I just wanted to raise that now and maybe the sponsor will want to get back to that later on.

DR. GILMAN: Well, I think it is best to handle it now. It has come up now, let's handle it now. So, if the sponsor wants to respond, please do.

DR. RUPPEL: That is true, there was a significant test for homogeneity in these studies but, as you will notice, all of the studies are going in the same direction so it is not a conflict in direction but, instead, a conflict in the degree of effect across studies.

DR. GILMAN: But you are showing a bar graph, which has implications with respect to the different effects here, and I posit that that is not a legitimate way to look at these data since you are comparing different groups,

different genders, different doses.

DR. RUPPEL: Mark, I think your next slide shows the actual odds ratios.

DR. CORRIGAN: Actually, that is one of the back-up slides.

DR. VAN BELLE: Well, I think the question is whether groups that are not homogeneous with respect to endpoint should be pooled, and I think the usual answer is that you don't do that. At least indicate that that is the case.

DR. GILMAN: Dr. Grotta?

DR. GROTTA: I guess on a related issue, what I am struggling with is really sort of a meta-analysis of subpopulations from different trials. What I would like to see if you are going to do a meta-analysis is to take the studies, all of them and all the patients in the studies, and meta-analyze them for the mortality endpoint and see whether by looking at the studies in a meta-analysis you see an effect on mortality -- the whole studies, not the subsets.

DR. RUPPEL: Let me pull the relevant slide. I think it is M-29.

[Slide]

We felt, since we wanted to look at the efficacious dose in each gender, that it would not be

reasonable to include the lower dose females, any of the lower dose females since they would not be at an efficacious dose. So, that is what we were aiming for, the efficacious dose for each of the individuals.

DR. GILMAN: So, you pooled data across different doses for all grade IV/V cases. Is that correct?

DR. RUPPEL: That is true; that is true. They were numerically different but, again, we were striving to get the efficacious dose as the pooling criteria.

DR. CORRIGAN: For the intended recommended dose for treatment.

DR. GILMAN: Yet, again, you are combining non-homogeneous populations. I think Dr. Van Belle's comment still pertains. Dr. Temple and then Dr. Katz?

DR. TEMPLE: There is always a question. The original people who used to do meta-analyses actually did a single axis and showed where all studies were along it. So, you could just look at all the studies and reach your judgment about whether there was consistency. Nowadays, more people are more likely to throw them altogether and get a combined endpoint. But, in fact, they are telling you the same thing. It is sort of obvious if you look at grade IV/V that it is sort of going to come out okay when you pool them. You can see that from the individual studies.

Whether it makes any more sense to have a combined

analysis than to look at them individually is sort of a matter of judgment. My own view is you don't learn more from that. You can, in fact, predict the outcome of that by looking at the individual studies. You can sort of tell that 65 is not going to overwhelm 63 so it is going to come out okay. But I think the argument, if it is credible, is that if you look at that group in several different environments it all leaning the same way, and pooling them altogether doesn't make that case any better than looking at them separately, I would think. So, I mean, I don't know if it is good to throw men and women together or not, but it some sense there is no real answer to that. But, you know, you do get to look at individual studies and make of it what you will.

DR. GILMAN: Would you like to respond to that question?

DR. KOCH: Yes --

DR. GILMAN: Please introduce yourself.

DR. KOCH: I am Gary Koch. I am a statistical consultant for Pharmacia and Upjohn. This display basically has two roles. As you have heard, study 63 had a confirmatory inference for the subgroup. This slide basically shows how study 63 fits with the other studies in that subgroup.

The other way in which it fits is the sponsor

basically is looking for an approval for a dose of 6 for men and a dose of 15 for women in IVs and Vs, and this is their combined data on those 2 doses for men and for women in IVs and Vs. It simply gives you a picture of what they are looking for an approval for.

Certainly, the other questions that have been raised merit attention, but all this slide is trying to do is give you an overall picture of how the two doses the sponsor is putting forward for IVs and Vs look.

DR. GILMAN: Dr. Katz?

DR. KATZ: I still don't know, even with that explanation, how the integrated analysis adds anything above and beyond just looking at the trials, as Dr. Temple said, individually.

The other question I had is how does the fact that this subgroup was a retrospective subgroup in every study, except 63, affect the appropriateness of this sort of analysis?

DR. CORRIGAN: Well, part of the integration questions are also going to be addressed by Dr. Marshall. Maybe we can return to that because --

DR. GILMAN: Well, I would kind of like to have the sponsor answer Dr. Katz's question now, if somebody will, or maybe Dr. Van Belle will respond. Would the sponsor respond please?

DR. RUPPEL: I am sorry, could you repeat the question?

DR. KATZ: Yes, in three out of those four studies the subgroup of interest was retrospectively designated, and I am wondering how that affects the appropriateness of performing an integrated analysis above and beyond any other questions that there might be about the appropriateness of the integrated analysis.

DR. RUPPEL: I think clearly everybody accepts that 63 was prespecified for the high neurograde patients. I think you have also heard earlier that 32 could be thought of as confirmatory because essentially the exploratory work had been done on 29 and 65, and we also went back and looked at 32.

DR. KATZ: I don't think it has been established that 32 -- that the effect on the neurograde IV and V has been found to be confirmatory.

DR. RUPPEL: It has been suggested.

DR. KATZ: Well, a lot of things have been suggested --

DR. RUPPEL: Right.

DR. KATZ: But I think that is actually an issue that needs to be discussed. I just don't think it has been shown yet.

DR. RUPPEL: I think in the integrated analysis

our attempt was simply to get a look at the overall picture, what is the overall signal for this drug across all the studies, the whole body of evidence that we have seen and what we would like to go to market with. True, in some of the studies this neurograde group was exploratory; in others, at least 63, it was adequately prespecified. I don't think that that will necessarily detract from kind of the overall summary nature of the integrated analysis, and I am not by any means suggesting that the integrated analysis should stand by itself as supportive evidence.

DR. GILMAN: Not "some" were prospective; one was prospective and the others were retrospective.

DR. RUPPEL: That is exactly correct. I apologize if I suggested otherwise.

DR. GILMAN: Dr. Van Belle first and then Dr. Grotta.

DR. VAN BELLE: Well, I guess questions since I was the one that introduced the confirmatory/exploratory rubric here. The question to the sponsor is in the setting of this standard, before you did the prespecified endpoint of the neurograde IV and V in study 63, did you actually look at the data from study 32 in order to establish that standard? In other words, when Dr. Oliva presented the data the impression was created, certainly in my mind, that study 32 had not been looked at with that particular endpoint and,

therefore, could be considered confirmatory. If, in fact, the sponsor used the data from study 32 to set the endpoints in study 63, then I would certainly not claim that that was confirmatory. So, I think it becomes quite important what the process was for setting the standards for study 63.

DR. RUPPEL: You are correct in that the process we went through when we opened up 65 and saw that 65 was not positive overall, we then looked at the whole body of data that we had and that led us to the prespecification for 63.

DR. VAN BELLE: So, I would say study 32 was not confirmatory in that sense.

DR. GILMAN: Dr. Grotta?

DR. GROTTA: Just to follow this up, as I see it philosophically you have done four experiments, each study being an experiment, each one having a prespecified outcome, and two of those were positive and two of them -- well, actually only one of them was positive, the last one. What I would like to see if you are going to do a meta-analysis is a meta-analysis of the four studies as to what their magnitude of effect was on their prespecified primary endpoint. Recognizing that there is noise in there and that not all the patients were dosed at the dose that you are proposing, but if you are going to do a meta-analysis of trials I think that is the way you do it.

DR. GILMAN: Does the sponsor want to respond?

DR. CORRIGAN: Your point is taken, Dr. Grotta, but the only issue, of course, is that in the first trial the endpoint was vasospasm.

DR. GROTTA: If you do that, you may find that your meta-analysis still might be positive.

DR. GILMAN: Well, except that prospectively only a single study showed a positive outcome with respect to the prespecified endpoint.

DR. GROTTA: But if the others were trending positive and you do a meta-analysis and you have one positive trial and the others are trending that way the meta-analysis may show, in fact, that there is an effect of your drug across all these trials.

DR. GILMAN: Would the sponsor want to respond to that?

DR. RUPPEL: Could you put up M-5, please?

[Slide]

This shows the results across all neurogrades for the four studies. However, I must admit that it does not include the 6 mg females from 32 or 29 or the lower dose treatment arms either. It is only the 6 mg males from 32 and 29, and then all of the females from 63 and 65. As you can see, there is clearly not a significant effect across studies.

DR. GILMAN: There is clearly not a significant

effect across all studies -- I would emphasize.

DR. CORRIGAN: For all neurogrades.

DR. GILMAN: Right, for all neurogrades, in mortality for all neurogrades.

DR. CORRIGAN: Which wasn't the primary endpoint. So it really doesn't speak exactly to Dr. Grotta's point.

DR. RUPPEL: I apologize, I have been informed that that did include the females as well from 29 and 32.

DR. GILMAN: That did include the women?

DR. CORRIGAN: At a possibly ineffective dose.

DR. GILMAN: Dr. Cui?

DR. CUI: I just want to point out that the p value for study 33 in the sponsor's slides is 0.010 --

DR. GILMAN: Study 33? You mean 32?

DR. CUI: Yes, 32. Sorry. The protocol prespecified to do the adjustment for the multiple doses, which is three doses. So, if you adjust that the p value is 0.03. Then you have to account for the post hoc nature of the whole thing. So that nominal p value is somehow misleading. You do not mention anything about how to reach that p value.

DR. GILMAN: In the end then, what sort of p value would we look at there?

DR. CUI: The first thing I want to say is that the primary endpoint, vasospasm, was tested at the 0.05

level. So, if you test at the 0.05 level generally the overall efficacy conclusion will have a type-1 error rate larger than 0.05. As Dr. Oliva already said, the question is how much inflation of the type-1 error we are going to take.

There are 14 prespecified primary and secondary endpoints and -- there are 14 possible endpoints. The number of the combinations for the neurograde is around 31. I don't want to do very conservative adjustment, just very limited adjustment. So, if you adjust for 3 doses, if you adjust for 2 genders, if you adjust just for neurograde II, low neurograde and high neurograde, then the p value should be times 8. So just at that nominal p value, if it is times 8 it already exceeds 0.05.

DR. GILMAN: So that p value should be 0.08? Is that what you said?

DR. TEMPLE: Well, I want to argue that. That is for both genders so you don't have to adjust for that. But if you do take 3 groups and 2 endpoints, which is not very aggressive one, you come out around 0.05 or something like that. So, it means it is, at best, a marginal finding if you just make the absolutely minimal corrections that you have to make for 2 endpoints, and there are many more, and 2 subgroups. But I wouldn't correct for 2 genders because that includes both.

DR. GILMAN: Dr. Van Belle?

DR. VAN BELLE: I think there is really one big issue here. The FDA really wants to guard against type-1 errors, namely, to approve a drug that doesn't work. And, we got into this p value game by these kinds of considerations. It seems to me that clearly some kind of adjustment has to be made, whether you multiply by 3 or 8 is open for discussion.

I think in most statistical texts the issue is what is the family of comparisons across which you make these comparisons, and that is not completely clear. Clearly, the sponsor would like to minimize the membership in the family; FDA may want to maximize the membership in the family. I think we could agree that the family consists of more than one comparison, whether it is 8 or 30 is not clear. So, I would certainly agree that some kind of adjustment has to be made, particularly since this was not the primary prespecified endpoint, and I would be very concerned about that. But, I would be very reluctant to sort of assign a multiplier of 8, or 5, or 3 to this kind of exercise.

DR. GILMAN: Then where would you leave us? How would we interpret that result then?

DR. VAN BELLE: Well, I would certainly agree that the p value of 0.1 is nominal in the terminology that has been used before. I would have no problem moving it up to

0.05 even. Whether you make it 0.2 or not, I think that that would create more of a discussion in my mind at least.

DR. GILMAN: So, to stop playing the p games, then would you say that this becomes a marginal finding? Would you use that term?

DR. VAN BELLE: If marginal means greater than 0.05, I would agree to that.

DR. GILMAN: Marginal means on the margin, neither clearly positive nor clearly negative as a study.

DR. VAN BELLE: Well, if I were a patient I wouldn't be very interested in that particular treatment.

DR. GILMAN: Dr. Temple?

DR. TEMPLE: I think that is what, at the time I signed the letter, I thought this study represented. We also noted the rather striking result in the male subgroup and, therefore, thought that confirmation of this finding would be good enough. So, we certainly didn't think it stood alone. That is why we didn't approve it. Whereas sometimes a mortality finding can be the basis for approval we certainly didn't think that because it was kind of odd, and marginal once you make minimal corrections. So.

DR. GILMAN: All right. Dr. Corrigan, back to you.

DR. CORRIGAN: Okay. I would point out that mortality -- as you point to the p value game, there is still a 2:1 difference sometimes, and with these numbers it

is better maybe to count up the individuals.

DR. GILMAN: Would you explain what you just said?

DR. CORRIGAN: Well, if you have, for instance, a finding, as Dr. Temple pointed out, of no one who dies in the placebo group and you can do enough statistics so that that begins to lose significance, I mean it bears the replication that the FDA suggested, but there was I think, in 32, 19 versus 0 deaths. That is all.

DR. GILMAN: Dr. Penn?

DR. PENN: Might I suggest that a patient might make a different choice than the FDA would in terms of p values. If I were a patient with subarachnoid hemorrhage and there was one chance in twenty that a study had shown a positive drug for this I might want to take it. So, I mean we are using different criteria, and I think we ought to know that.

DR. BROOKE: Can I comment?

DR. GILMAN: Yes.

DR. BROOKE: Dr. Penn, you are right and that is the sort of emotional decision versus the scientific decision, and I do agree with you. I have often said myself that if I had some dreadful disease I would take everything I could lay my hands on whether I knew it worked or not. But I am not sure that that is a good basis for making the decisions that we have to make today. And, we are wrestling

with the four studies. I thought the addition of another yellow and red bar is rather like the car advertisement where you have 16 different additional features and not just 10 and that makes it more saleable. And, I get suspicious when I see a bar that appears on a histogram that doesn't have any clear reason to have been there.

We are wrestling with the four studies. Now you have thrown the committee into disarray by introducing a fifth -- oh dear, I am using my words carefully; don't take offense -- a pseudo meta-analysis which can only weaken the evidence of the previous four. You know, it is the old argument that four reasons for an effect are much weaker than one reason for an effect. So, I would just caution you, if you throw that up you are going to get a lot of discussion.

DR. CORRIGAN: Your point is well taken and it is certainly not my intent to throw the committee into disarray. Perhaps at this point I could invite Dr. Marshall to continue the discussions.

DR. GILMAN: Dr. Marshall, are we in disarray? I thought we were doing pretty well.

**Risk Benefit Assessment:**

**SAH Response to Specific FDA Comments**

DR. MARSHALL: Thank you, Mr. Chairman.

[Slide]

I must say at the outset I am impressed with the sophistication and the quality of the FDA analysis, particularly in the areas of safety, and I hope that I can perhaps appropriately respond.

Being prudent, I think we will pass the slides about integrated analyses because I feel the floor shifting under my feet --

[Laughter]

[Slide]

These were the issues raised in the FDA evaluation and both Dr. Oliva and Dr. Racoosin pretty much stuck to their outlined slides and, as I said, I thought they did a very good job in making their presentation. As I also pointed out, a prudent man doesn't go back to where he should or, as they often say, only a fool is his own lawyer. So, I am going to try to skip through that and skin the integrated analysis discussion because I think we are not going to make any new ground there, and move on to the issue of efficacy.

[Slide]

As you have seen with regard to the data, and we have had discussion about the p values in 32 and the retrospective nature, and I don't want to spend any time on that but, rather, move on to some of the issues of prespecified and unspecified risk factors that were raised

in the FDA analysis, and also to point out how the neurosurgical community sees it.

I also feel obligated to respond to the discussion of PMR2 and, having been mentored by Prof. Genet, I see his boat turning over in the loch in Scotland following that discussion this morning because, quite frankly, it begs credulity to argue that a patient who is flaccid on one side and decorticate on the other is considered better using your analytical methodology than a patient who is bilaterally decerebrate. It is not true; it isn't held up by any other analysis. And, I think it shows really sort of cocked up kind of mechanism of looking at severity because if you look at all the experience in the Glasgow Coma Scale the worse unilateral side is a much better predictor. So a patient, for example, who is decerebrate on one side and purposeful on the other does worse than a patient who is decorticate bilaterally who would be thought to be worse.

So, I think the PMR2 issue really is a creation that, in my view, has no validity and, having dealt with the GCS and the problems with the imputed score which have come up here, I would caution you in using that to assume that you have identified a relatively good grade or flaccid on one side and decorticate on another who would not meet your requirement for PMR2.

[Slide]

Having said that, I would like to move on to a number of the other factors. The first is the fact that intraventricular blood on CT scan is increasingly recognized in the neurosurgical and critical care community as a major risk factor for an adverse outcome in patients with subarachnoid hemorrhage, even more important than that traditionally -- excuse me, let me go back a minute -- the one of clot fitness which, over time, seems to have become less important, perhaps in part because part of the clot is evacuated and nobody knows exactly the reason.

The second area is the site of the ruptured aneurysm, which is shown here. I would also point out that generally in all the subarachnoid hemorrhage studies that have been done age is viewed as a continuous variable, and the cut point of 65 most of us think is no longer appropriate. The cut point should be 70-75. And, if you reanalyze the data you see nothing with regard to that.

[Slide]

Moving on to the next slide, aneurysm location is a very important factor in predicting the outcome in patients with aneurism or subarachnoid hemorrhage. For the lower neurograde patients, for which the sponsor is applying, one sees a consistent, unfortunate pattern from the sponsor's perspective, of a bias very heavily for unfavorable location of the aneurysm --

[Slide]

-- shown here, varying in the smaller studies from 0-20 percent, and in the most recent studies a doubling in here, about a 60 percent excess, which carries with it the implications in this group about doubling of mortality in this specific subset. So, the factors that are really important here broke against the drug and, in spite of that, certainly in 63 as we have shown, there is a positive effect.

With regard to intraventricular hemorrhage, although not so strong, again a much larger population of patients had it, and in 63, 65 and 29 again the frequency of intraventricular hemorrhage was much greater in the tirilazad-treated group, which serves to emphasize I think the strength of the observation that this appears to be an efficacious agent in patients with poor grade aneurysm or subarachnoid hemorrhage.

[Slide]

While we have heard a lot, and appropriately so, about retrospective analyses, I think that it is appropriate to comment a little bit about time to dosing and time before and after surgery because the trend from the time that these studies were conceived, in the late 1980s until now, has been to earlier and earlier surgical intervention. The idea is that one can prevent a re-rupture of the aneurysm and

institute what is thought to be acceptable treatment that may influence favorably delayed ischemic neurologic deficit or vasospasm. Also as a surgeon, although I am reluctant to say it, the notion might be that you could protect the patient from us during surgical intervention with a drug that might protect against retraction and so on.

So, while again I recognize that this was a retrospective look, to me it is an appropriate look as a clinical for an area where there might be a substantial effect. Indeed, with that caveat, one notes that in the 65 study, which overall clearly did not demonstrate significance, one sees a rather substantial effect on mortality in those patients who were dosed within 24 hours, and the data looks very similar obviously if you look at an odds ratio but, more importantly, if you look at surgical intervention, that is, the time that the drug was given before surgery versus after surgery.

[Slide]

Now, in terms of looking at the data, Dr. Brooke raised a very important issue, which was the whole issue of death not being necessarily worse than catastrophic disability. That is something I have worried about since I began my work in catastrophic diseases of the nervous system for many years. What we have here, on the left-hand side of the slide, is the actual vehicle distribution in the 3

months GOS -- taking the caveat of Dr. Grotta also which I would like to speak to as to when you look at outcome. In the middle slide we have a vehicle group modeled with a predicted 30 percent upward movement of the population. On the right-hand side we have the actual distribution of tirilazad.

What you see is the fact that you have a reduction in mortality, as has been shown by Dr. Corrigan and myself; an increase in good outcomes; a slight bump up in severe disability which raises the issue that Dr. Brooke spoke to, and I would point out that there are now a number of studies in aneurysm and subarachnoid hemorrhage, as well as in head injury, which demonstrate a substantial movement upward of these patients -- yes, Dr. Katz?

DR. KATZ: Is this, and I hesitate to use the word, integrated data?

DR. MARSHALL: Yes. This is all of the bad neurograde patients together.

DR. GILMAN: Excuse me, these are data we have not previously seen so could you explain how you integrated them? Was it done in the same way as we have heard previously?

DR. RUPPEL: No, these were combined to see how they fit the model. The 30 percent was the maximum likelihood estimate from a model and we projected the

vehicle predicted model and then just looked to see how well the observed tirilazad GOS response matched the predicted model.

DR. TEMPLE: But just pooling all the patients, each one counting one?

DR. RUPPEL: Exactly.

DR. TEMPLE: So they just threw all the patients together as if they were in a single study.

DR. GILMAN: Well, then I have to say that this has the same objections that we referred to earlier in throwing together data from people treated with different doses, different genders, and various different studies, and prospective versus retrospective analyses, etc. These, by the way, are all retrospective, I assume. Right?

DR. MARSHALL: Well, taking those caveats, Dr. Gilman, I think the point that I wanted to make from the slide with regard to severe disability is the fact that if you look at two years they move disproportionately, 2.5-1, to better grades rather than stay the same or down. In other words, 56 percent of these patients will be located in the good or moderate distribution at 2 years in this population. This has been seen and replicated in a number of studies.

DR. BROOKE: Can I get clarification on that?

DR. GILMAN: Yes.

DR. BROOKE: A 30 percent shift upwards, how is it

calculated? I mean, if you shift from a neurograde IV to III, is that 15 percent?

DR. MARSHALL: No, no, it is the total N, the N of the population, not weighting each grade up. And, this is in keeping with our experience in the large head injury trials as well.

I would made one other point since I saw Dr. Leber here this morning and we had a discussion in June. One of the problems with the severe disability category is that it is profoundly broad. It goes from patients who are institutionalized, non-ambulatory, who can't feed themselves to people who can live alone but they can't get from Gaithersburg to the Capitol for example. I think there is a problem there, and if a new, more extended GOS is used it may make it easier in these studies to look at that over time.

The point to make simply was that severe disability, and I said I appreciate Dr. Brooke's comment, is a category with a large potential for improvement over time. In fact, that is what we have seen in these patients over time.

DR. GILMAN: Dr. Marshall, could I stop you for a moment?

DR. MARSHALL: Sure.

DR. GILMAN: You said, "and this has been

replicated in many other studies." Could you discuss which studies you are referring to?

DR. MARSHALL: Right, there is a Danish study; t there is a Scandinavian study, both published, looking at 3-month, 6-month and 2-year outcomes of aneurysms and subarachnoid hemorrhages and what you see in patients who fall using the GOS, which we all recognize was developed primarily for head injury as an outcome measure, is that those patients, indeed, move up rather than die, become vegetative, that is deteriorate, or remain within the severely disabled category.

DR. GILMAN: These are patients treated with tirilazad?

DR. MARSHALL: No, I am just talking about in general the issue is what happens to patients in the severe disability category, and I think there is some data -- Dr. Wilke is here, from pharmacoeconomics, if you wish to discuss this issue in more detail later we could perhaps talk about that.

[Slide]

Another very important issue, appropriately raised I thought and very well done by the FDA, was the issue of identifying the target population. I would like to make a couple of preliminary comments that were not dealt with on the slide because the implication I got from listening to

the talk, and perhaps it was in error, was the notion that somehow intubated patients are likely to be conversational with the physician caring for them, and if you calculate the percentages of patients who were intubated in the two groups, in terms of looking at comparability, they were 55 percent in the placebo and 50 percent in the tirilazad group. Intubation in patients with this disease is an indication of a significant subarachnoid hemorrhage and a patient who is in real jeopardy of death or severe disability. The notion that you couldn't get a score in 41 percent of the patients is included in large measure, and there is a large experience with imputed verbal and eye opening scores in these patients. I mean, literally more than 100,000 patients have been studied. In structural diseases of the nervous system you can calculate, with about a 98 percent accuracy, based on the motor score and knowing whether there is eye opening or not what the patient's verbal score would have been.

DR. KATZ: If you actually look -- and Dr. Oliva talked about this a little bit, the patients who had an mGCS of 7, not including a verbal score, who were intubated -- well, actually if you don't look at the ones who were intubated, if you look at the ones who actually had a verbal score but had a 7 just adding the 2 components, eye opening and motor, on average those patients had a verbal score of

2.2 or something like that. So, on average those patients were a neurograde III.

DR. MARSHALL: I think that one of the problems is that with regard to neurograde one of the difficulties is the march of time, as Dr. Corrigan referred to, and the difficulty with the neurograde III patient is that they really fall into two groups, those who have focal deficits, usually because of the site of the hemorrhage, and those with depressed levels of consciousness which the Japanese, in fact, routinely divide into IIIBs, and they perform exactly like the IVs, and if you go back through this data, which I hope to show you, that is exactly what you see.

I mean, I think you have raised a very good point with regard to looking through the data as it was presented by the sponsor. I had asked them to do a number of analyses, which I would like to show you, to make the point that I think you can identify an appropriate target population at high risk given the concerns that you have raised about a potential signal in the patients in better grade.

[Slide]

This is one way of looking at the data, and these are patients with a motor score of 5 or less, that is, people who do not obey command and have, by definition in the CRF, a depressed level of consciousness. Just in looking at where you wind up with the total N provided by the

sponsor in the IVs and Vs, what you see is that you wind up with 92 percent of the patients. So, simply by saying patient no longer obeys commands, purposeful or worse, and has a depressed level of consciousness in terms of eye opening you can pick up what I believe is an appropriate target population of patients who are at high risk of death or severe disability.

DR. GILMAN: Now, what are you showing us here? What is the integrated score from?

DR. MARSHALL: It is not an integrated score -- oh, that is just the total of these populations. In other words, in 32, 97 percent of the patients would have fallen; in 29, 87 percent; in 65, 93 percent; and in 63 approximately 90 percent of the patients would have met the criteria of depressed level of consciousness, not obeying commands. In total, 92 percent of the patients in the study would meet that definition of a bad patient in terms of a subarachnoid hemorrhage.

DR. GILMAN: I don't think that that was compatible with the analysis of those cases, unless I am mistaken. The IV by V subgroup represented only 19 percent.

DR. MARSHALL: That wasn't the point. The point was if you take 100 percent of the IVs and Vs and simply used the definition of a Glasgow Coma Scale of 5 or less and depressed level of consciousness you still wind up with 92

percent of the cohort. That is the point.

DR. OLIVA: Are you changing the definition --

DR. MARSHALL: I am just trying to validate -- the FDA raised in your report the issue as to whether or not one could easily or appropriately identify the patients at risk, the population at risk --

DR. OLIVA: No, no, no, that wasn't what I said. What I said was can we easily identify the IV/V grades.

DR. MARSHALL: Right, and I am showing you a way to identify them that eliminates the confusion with regard to the so-called neurograde or any other indicators, and to go back to what is easily done in every critical care unit worldwide, in millions of patients daily, which is to take the motor score and the depressed level of consciousness, which is very simple to do and can be extracted from the CRF in this study without difficulty. That is my point. We can argue about it but that was the point.

DR. TEMPLE: I think the contention is that if you eliminate the need for doing the verbal score which is confused by the fact that some people are intubated, what you are saying is you can get most of the relevant patients without using that.

DR. MARSHALL: You said it better than I did. Right. Correct, exactly. And, they can be identified within the IVs and Vs --

DR. TEMPLE: Pretty well, yes.

[Slide]

DR. MARSHALL: And, if you go to the more severe, that is patients who are withdrawing to pain, all of whom, we know, are not speaking and do not have eye opening, then you wind up with this and, again, if you analyze the data, and I recognize that this is retrospective, you see the same trends in relative efficacy doing it this way.

So, I think one part of the discussion to focus on, if the issue of labeling were to come up, is that a better mechanism potentially to identify these patients is to use the GCS as it was intended and in a patient with a depressed level of consciousness because, to me, that is the appropriate cohort that is easily recognizable.

[Slide]

Now, in looking at the classification schemes and their effects, what one has here, again combining the data in response to an anticipated question from Dr. Gilman -- this is the modified GCS, the World Federation Neurograde, and the Glasgow Coma Scale best motor score.

[Slide]

The next issue was the whole question of complementary increased risk in the neurogrades I through III.

[Slide]

If we look at the data that was presented by Dr. Racoosin with regard to mortality at 3 months, what you see is what you showed, a difference of about 1 percent here and approximately 2.5 percent there. Then, combining the data, perhaps with some objections, one sees no substantial difference. But the concern was clearly appropriate to raise.

[Slide]

These are the odds ratios for these. Let's forget about the integrated to save me from being attacked again about that issue, but to say that one really doesn't see much.

Then, I asked the sponsor to carry out a number of other analyses to look at --

DR. GILMAN: Excuse me, could you go back to the previous slide?

[Slide]

DR. MARSHALL: Sure, you want the odds ratios, Dr. Gilman?

DR. GILMAN: Yes. What are you showing us here?

DR. MARSHALL: This just shows the odds ratios in the four studies of the neurograde patients I through III.

DR. GILMAN: I see.

DR. MARSHALL: Where the issue of harm has been raised --

DR. GILMAN: I see.

DR. MARSHALL: -- by the agency.

DR. GILMAN: All right.

DR. OLIVA: I just have a comment. The Y axis is a logarithmic scale, is it not?

DR. MARSHALL: Yes.

DR. OLIVA: So, really that shrinks the differences in odds ratios across different studies because they are just slightly above 1.

DR. MARSHALL: Well, except 32 obviously.

DR. OLIVA: I am just saying if you change the scale --

DR. MARSHALL: Well, 32 doesn't go above the line no matter what you do. Right?

DR. OLIVA: Right. My point is if you change the scale of the Y axis you would see the difference much better.

DR. KATZ: Exactly. I mean, what appears to be a lack of any sort of a signal -- we can argue about whether or not there is a signal but the apparent lack of a signal here is entirely related to the mode of presentation. If you just look at the numbers, as we have done, they are what they are. We can argue about whether or not they are meaningful but at least they are plain.

DR. MARSHALL: Well, let me show you another way

that I asked the sponsor to look at the data in all of these four studies with regard to the issue of signal, and that is to use a hard endpoint which is the patient obeying commands, all of whom should pretty much, by definition, fall in the III, II, I category, in an attempt to see if there was a signal there. What you see -- let's ignore these two and let's look at 65 and 63 where this was raised by the agency, and what you see is in the vehicle group 11.5, 10.6, so about a 1-point difference but flipped around now from the I through III as they were defined. Then, looking at 63, about a 1 percent difference in the other direction.

So, I think using obeys commands as a hard measure of patient function, one sees something that looks a little bit different but, again, it is hard to see, from my perspective, that this is a signal.

DR. GILMAN: So, what is your conclusion from that? You are not seeing any change.

DR. MARSHALL: Right, but the issue is that the agency has raised the concern, Dr. Gilman, that there is a signal of increased mortality in the group I through III. In looking at them on the basis of a hard measure as the patient is inputted into the study -- obeys commands, motor score 6, we see no harm. I think that suggests that the signal, if there, is very weak and not detectable using an alternative mechanism to look at relative goodness of the

patients in that group.

DR. GILMAN: Let me ask the agency whose data did you use for classification into grade? Did you use the sponsor's data or did you have your own data?

DR. OLIVA: No, we used the sponsor's data.

DR. GILMAN: They used the sponsor's data. So, are you saying that the sponsor's data were flawed?

DR. MARSHALL: No, no, I was just saying that if you look at the Is through IIIs, and perhaps we can go back to that slide --

[Slide]

In Dr. Racoosin's presentation, she pointed out appropriately -- I mean, she showed it for varying time efforts but at the 3-month date what the agency raised concern about was the 7.8 and 10.5 difference, although clearly it is not significant, and this difference in 65 of about 1 percent.

I then said, from my perspective, could we look at a hard endpoint -- obeys commands. There is no scale involved; it is very straightforward. We know it is a very good predictor of relative outcome. What you see when you look at motor score 6 --

[Slide]

-- and this slide combines patients who are not obeying commands, perhaps those hanging on the edge of III

and IV, and what you see again -- forgetting 32 and 29 -- is that in 65 you see a very small difference, and here you see the same difference you saw when you include those patients who might be sort of a beginning bad grade III -- a little bit larger difference but clearly not significant.

So, the point was that I don't believe, and I think the sponsor doesn't believe as well obviously, that there is a signal in the I through IIIs of harm. That was the point.

DR. GILMAN: Dr. Brooke?

DR. BROOKE: I don't mean to interrupt you in midstream. I wanted to congratulate you. It is really nice to hear somebody who is obviously involved in clinical medicine talk. You are pulling out the groups, of course, that all of would have been fascinated by rather than the ones that are rather artificially pulled out.

Unfortunately, it is all retrospective, as you have commented, and there are two reasons for a retrospective analysis. One is that you can't believe that you have spent ten percent of your life on a study which is negative --

[Laughter]

And, the other one is in planning for the next study so that you can really answer the question which should have been answered in the first place. I wonder which

of the attitudes you prefer to adopt for this particular --

[Laughter]

DR. MARSHALL: Well, this is the reason why I became a CEO of managed care for a while in San Diego -- I am just kidding! My point only was that in reading through this and having been involved as an investigator in the American subarachnoid hemorrhage studies, and as the principal investigator of 36 which we have not yet come to, one of the head injury studies, one is always concerned about the whole notion of imputed scores, weighted scores, and so on. Dr. Wayne Alves is here, from the University of Virginia, and we wrestled with a lot of this in the 17 and 36 study, and we carried out a validation study to look at the issue of imputed scores with regard to the verbal and eye opening score. I think, from my perspective as a clinician, I would like hard numbers that I can have absolute confidence in, and the GCS has stood the test of time and has been validated now in a number of diseases. It falls down in stroke because in stroke, obviously, you are worried in most instances about focal deficit and, therefore, using the worst GCS initially when the concern was preventing cerebral vasospasm a focal deficit was appropriate.

What should have happened, in retrospect, if you look at the CRF, is that they should have switched it when

the endpoint became mortality and they did not, and I think that leads to a certain amount of noise which the agency appropriately in their analysis pointed out. I think that criticism was correct.

I then said let us look at a number of other variables to look at relative wellness or illness, percent intubated, to see if there is a difference. Is it true that the agency has identified an imbalance that says the placebo group is worse? That does not appear to be true. It is off by less than a percentage in literally thousands of patients.

When I saw the PMR2 calculation, I then went back and looked at it and said, now, is it logical based on what we know about patients behave that someone who might be flaccid on one side and is decorticate on the other falls out of a bad classification? And, the conclusion was of course not. It doesn't make sense and when, in fact, you go back and recalculate the data it blows the whole PMR2 thing away. I mean, it is not a valid construct. It was not prespecified by anybody. It was a mechanism used to look for potential harm or effect of the agent. I think somebody said it is counter-intuitive and it is just not right because we know that in that setting the worst side predicts the outcome even if the best side is substantially better. A decorticate patient, bilaterally decorticate, does less well

than a flaccid patient on one side who is withdrawn or focused on the other. They do worse.

DR. GILMAN: Do you want to respond to that, Dr. Cui?

DR. CUI: I just have a quick comment. Certainly I do not have insight or a medical reason for why PMR2 is significant, and Dr. Marshall suggests that there is some literature saying that some other factors are important. But I want to point out that basically before doing a trial we only have a guess; what will happen we don't know. We don't know which is a good predictor -- maybe in the literature but when you do the trial that may turn out to be not so important.

I just want to point out that in this trial I used the PMR2. I identified the PMR2 in the control group. The patients with PMR2 had a significantly higher mortality rate as compared to the patients without PMR2.

The other thing is about the rupture locations. I did try to check that. Actually, I tried to be fair. I checked all the baseline --

DR. MARSHALL: I am sorry, I couldn't hear you.

DR. CUI: I tried to be fair to check all the prognostic factors. Something like the baseline blood pressure seems, to me, not significant. Some things the sponsor checked, say, the age, therapy and that kind of

thing I regard as outcome instead of baseline measurement. So, I basically don't want to use that. If you want, I can explain.

But the rupture location, to me, it is very hard to use because for rupture location there are four or five categories, like middle cerebral and posterior cerebral, that kind of thing. So, if you say that is important the question is -- you have that table showing that it is a bias against the drug in terms of the distribution of rupture location. My question is which category you used to select?

DR. MARSHALL: That is very straightforward. I mean, posterior circulation aneurysms are those that appear on the vertebral or basal artery and its tributaries, including posterior cerebral, and that is a well-defined adverse risk factor forgetting about this trial.

DR. CUI: I want to know if the mortality rate was associated with the posterior cerebral in the control group.

DR. RUPPEL: In the 63 study the patients -- you want the vehicle group? The patients in the vehicle group that had the posterior cerebral rupture location had a 50 percent mortality as compared to 41 percent in the others. In the 65 study there was a little larger spread. It was 50 percent for the vehicle patients with the posterior compared to 36 percent in the other locations.

DR. CUI: Okay, but for PMR2 the mortality rate is

an 8-fold increase.

DR. MARSHALL: You know, I have to say that I attempted, having received the report only fairly recently with PMR2, to not only recreate it but it was very hard because it wasn't quite specified how you got the patients. I also tried to validate it on over 100,000 patients, about 8,000 with aneurysm or subarachnoid hemorrhage, 56,000 head injuries, and another catastrophic disease of the nervous system and, in fact, I found the data to go in the other direction. That is exactly what was found by Brian Genet and Graham Teasdale in the initial worldwide studies of the Glasgow Coma Scale. That is, to make it very brief, unilateral badness, as you have defined it, ordinarily in a very large cohort does not overcome one side that is worse in your definition. So, I couldn't replicate it. It does not go along with previous experience. I understand you made the observation. It was unspecified obviously and it is interesting, and perhaps it may be useful and unique somehow, but it has not been seen before and is not validated.

DR. CUI: That is my concern. If you find something unique, and you know everything is done in a post hoc way and subgroup analysis, there is potentially a lot of bias introduced. That is exactly the problem I have in interpreting this trial. Say, for rupture location -- not

this one --

DR. MARSHALL: Well, we can put that up. It broke basically against the drug 2:1, and that has an unfavorable bias --

DR. CUI: Right. That exactly shows how bad the subgroup analysis is. If you use the subgroup analysis everything is biased, no matter the direction of the bias but it is the bias. A lot of bias is introduced and we don't know what will happen.

DR. MARSHALL: Well, I think there is a little bit of a difference when one has risk factors for which it is well known, for example, that they have a adverse effect on outcome and you identify them, you prespecify them, and they are there or they are not. I mean, here the fact is that aneurysm location, unfortunately, if one would like to have truly balanced populations, broke against the drug in every study. It is just chance. But it has a significant potential effect on the overall outcome and in spite of that we saw a beneficial effect of tirilazad.

I will return to this issue in response to Dr. Racoosin's remarks about 17 where this was such an egregious imbalance in the frequency of patients with bilaterally unreactive pupils that completely explains your concern about herniation. But this is an example of what can happen. These are clinical trials. We are not studying people under

the kinds of conditions would like in a laboratory and, in fact, people are now saying we should have a specific scale for ventricular hemorrhage and subarachnoid hemorrhage, and in fact many patients with posterior circulation aneurysms weren't even operated on until microsurgery, and Dr. Drake, who was the father really of posterior circulation surgery, would never operate in those patients early because of the excessive risk. So it is a different ball game. They are much more difficult technically and they do worse.

DR. GILMAN: Let's pause here. There are a number of questions from around the table. Let's stay on this point. Dr. Grotta and then Dr. Katz, then Dr. Temple.

DR. GROTTA: Well, the fact of the matter is that we can criticize the people for looking at a post hoc analysis of these studies, and we have to do the same thing or at least the agency does. I mean, really your safety analysis and the risks to the good grade patients is purely based on post hoc analysis of the data.

What Dr. Marshall is simply showing is that if you identify the good patients in another way, that is by their motor scores as opposed to their grade I, II or III, that the so-called higher mortality that you think you are seeing in the good patients disappears or is certainly attenuated. So, you know, we have to be consistent here.

Personally, I am not bothered at all by this. I

mean, I don't really think that the safety issue really is a big one. I think that it is all post hoc generated, and I also feel that the identification of patients clinically is not a big problem. We use the Glasgow Coma Scale and the Hunt and Hess Scales take two seconds to carry out in an emergency room every day. We impute the verbal score all the time.

I think that rather than focusing on the real issue which is whether efficacy has been demonstrated, we are getting hung up on whether we can identify the patients or whether there are safety issues which I think are really based on post hoc, unreliable analyses.

DR. GILMAN: Maybe Dr. Racoosin wants to respond, but she did comment on how meager the data were concerning adverse events.

DR. RACOOSIN: Safety by nature is a post hoc review. Its efficacy -- we know that there are prespecified endpoints and there is a certain level of statistically significant that has been established that needs to be met in order to get a win. In safety we have to review the data in the way we think will best identified safety issues. We can argue about the strength or weakness of the signal, but I just want to clarify that this is the nature of the safety review.

DR. GROTTA: Well, that is fine but then you can't

criticize Dr. Marshall for using the same post hoc way to identify in another way the same patients you are trying to say are at risk, and he is saying -- at least I think he is saying that if you simply define the good patients another way this safety concern disappears.

DR. RACOOSIN: I understand that. My concern is that the sponsor has put forth the group that they are asking for approval in, and those were the groups that we used to review the data, the safety data and the efficacy data. So, I can understand the point of wanting to find other clinical measures of severity of these patients, but the application is asking for approval for an indication in a subgroup that has been established that is different from the groups that Dr. Marshall is describing.

DR. GROTTA: You are right, they asked for approval in IV and V, and so it is valid to look at safety issues in group IV and V. But then you were trying to say that there are safety issues in another group, the good patients, and that is not the group that they are looking for approval in. So, he is simply saying that when you define the good patients another way you don't see the safety concerns.

DR. RACOOSIN: Could I just respond to that? I understand that that is where they are asking for approval. So, you could also call into question, well, why did I

discuss the exposure in head injury patients, and why did I discuss the exposure in the stroke patients. Any patients who have been exposed to the medication are pertinent to the safety review.

DR. MARSHALL: I would like to point out though this is the difference we are talking about, forgetting my reconstruction of this using obeys commands, we are talking about a 1 percent difference in 65 in 500 patients and 2 percent here. I mean, we are not talking about a significantly large signal suggesting anything.

DR. GILMAN: Dr. Katz, then Dr. Temple.

DR. KATZ: A few things, first of all, the relative risk is whatever it is, and we can go back and look at what we calculated it to be but we thought it was reasonable to draw attention to it. It seemed relatively big.

The other thing, just to emphasize what Dr. Racoosin said, looking at the complementary subgroup in the subarachnoid hemorrhage studies is very relevant because there is a question, given the combined concern about whether you can identify these people but these are people who might get treated, and that is something you need to think about.

The other thing is you presented 3-month. Judith also looked at 20-day mortality, the reasons being the ones

she described, which are that when you are looking at a drug-related adverse event it is perfectly reasonable to think about events that occur in close temporal association with the treatment. As time goes on other events that cause death begin to accrue, and you might lose what appears to be a signal. The relative risk may go down, which we saw, in time. So, if you look closer, and given the half-life of the drug, 20-day mortality is a reasonable thing to look at. At least in one study where the relative risk goes down in time the risk difference stayed the same so that it is really sort of a power question. The difference is still there, you just don't pick it up significantly and it is all related to what happened early. So, you know, there are different ways to view this. I think there still is a signal there, and some of this obscures that.

I have a couple of other questions also. I won't say always but we are often faced with the fact of a sort of retrospective identification of covariates that tend to make the analysis look better, and it is always obvious after the fact that these were the covariates that should have been included in the analysis. If it is so obvious and these are all so well known, they could have been included in the protocol, and have it prospectively stated that the analysis was going to be adjusted according to these covariates --

DR. MARSHALL: But they were prespecified and

identified for posterior circulation aneurysms and --

DR. KATZ: In the primary analysis of the data? I don't think so. You might have even stratified the randomization if you thought that was the case. So, I think this is all coming after the fact.

The other question I have is about PMR2 because it was the most powerful predictor of outcome in this data set. You said that you didn't exactly understand how we determined it.

DR. MARSHALL: No, I said I was trying to identify the sample size from which it came and I couldn't get the numbers to match up according to the worst motor scores that were reported in the CRFs. The other problem I had with it, quite frankly, is the historical evidence --

DR. KATZ: Well, that is what I am asking --

DR. MARSHALL: -- at the worse side overcomes a better score bilaterally dramatically. So, if somebody is flaccid on one side and withdraws on the other, which would not fall within your PMR2 calculation, those patients traditionally do much worse. And, you said, no, they don't; they do better than the patients who are bilaterally decerebrate. That just isn't so.

DR. KATZ: Well, one of the questions I had was you said you took this definition and validated it against these other data. I am just wondering how you did that.

DR. MARSHALL: Not validated. I mean, what you have done is you basically identified an unspecified point, which is perfectly fine, and then you tested and trained on the same data, which you caution us all the time not to do. I then went back and said, okay, maybe this is real. I mean, this is an interesting observation. I tried to identify the cohort of patients from which you then tested the assumption and couldn't get the numbers to match in terms of the total number of patients. I then went back to our own very large database based on clinical trials, some sponsored by the government and some sponsored by the pharmaceutical industry, and looked in those databases in three major diseases and could not confirm this observation. In fact, I found the contrary, which is that the worse side dictates, if you use the worst motor score, the outcome in these patients with a substantial degree of power versus what you found.

So, it is an interesting observation. I can't confirm it. I couldn't identify your N properly within the database. So I don't know what to make of it. I mean, I did the best I could to try to get at it. I thought it was interesting. It was important to me personally because it goes against all of our experience in 1970.

DR. GILMAN: Dr. Temple next.

DR. TEMPLE: Well, I think you are probably

demonstrating why after the fact subset analyses are treacherous because this was a very powerful predictor. We see that all the time, the expected covariates don't work out and something clearly unexpected turns out to be the major predictor.

DR. MARSHALL: Right.

DR. TEMPLE: That is not unusual. I wanted to make one point though, which is that the definition of the risk groups I through III has particular weight here because it is by removing them that you get all the good news.

DR. MARSHALL: Absolutely.

DR. TEMPLE: So, that all seems very reasonable, but if the complementary set is the group that in some sense has to do badly, or a little badly, if the other group is going to do well --

DR. MARSHALL: I think what I am saying here is there is no difference in the well patients and you see a substantial effect, or the sponsor is proposing that there is a substantial effect throughout the trials whether the endpoints were the same, whether pseudo meta-analyses are justified -- that is the point.

DR. TEMPLE: I understand. My prediction, however, is if you divide patients according to this method you will find that the complement of that no longer shows the benefit.

DR. MARSHALL: Well, I showed you, in fact, that that is not the case and --

[Slide]

-- if you look at this slide, what you see is what the agency showed in 63, which is that it is about the same and the difference becomes much smaller in 65. I just think it is a very weak signal and I don't think it is real, and that is what I am saying. It is half a percentage point in 303 patients in the vehicle group and here it is 2 percent. Nobody would even look at that and say it is a trend if you had a favorable effect of an agent.

DR. TEMPLE: Right, it is very hard to say these things are real and that is what Judith said, and you sort of do your best. But, the as defined group that turned out to be the dividing line that lead to wow, we win here and this isn't so good has particular credibility in answering that question --

DR. MARSHALL: Sure.

DR. TEMPLE: -- because there is always the concern that you can slice data any way you want and eventually, if you keep doing it, you will find a group that looks better --

DR. MARSHALL: There is no question and, as I said at the outset, I was impressed with the care and sophistication of the analysis carried out by the agency

because I think they raised a number of very valid points. I think there are appropriate and meaningful responses to those points. I mean, you know, we will show you what we found. We felt that since there was concern raised in the report with regard to the whole application of the World Federation Neurograde and confusion as to whether patients were categorized correctly, we have a hard point, universally accepted, as Dr. Grotta said, easy to apply, and we didn't see anything. I mean, that is the best I can tell you.

DR. GILMAN: Can I interrupt at this point to summarize and see if we agree about where we are now?

Dr. Marshall, if you now recategorize these cases as the more severe grades we find, according to your --

DR. MARSHALL: Less severe. These are less severe. These are less severely affected patients. Less severe patients are shown in these two slides, and I am saying we do not see harm.

DR. GILMAN: All right, you are saying we do not see harm, and can you now tell us what happens when you look at the worst grade cases, having decanted some of those worst grade cases into the better grade cases --

DR. MARSHALL: We showed that.

DR. GILMAN: -- with respect to efficacy?

DR. MARSHALL: I have already shown you that

slide. We can go back to it. We showed the motor score. There is another slide, perhaps I omitted it but basically --

[Slide]

-- less than or equal to IV, which is the more severe patients, which the same thing. What you see is, in fact, a more robust effect in 63. The mortality difference is even greater, that is, in favor of the drug, and you get a little closure of the trend in 65.

So, what I think it does, it shows you it confirms the observation of a very strong trend in the studies -- not getting back into the arguments of the validity of 32 and 29 about retrospective identification of that group. This is all neurogrades, predominantly, obviously, heavily loaded by the IVs and Vs because, by definition, these patients are, at best, withdrawing to pain which means they all pretty much fall into the IV and V category and the effect is still robust and still maintained. So, this is an even-handed application throughout the population.

DR. GILMAN: How many cases have you moved?

DR. MARSHALL: Eight percent of the patients would have been in IV and V and they dropped out. So it would be 8 percent. The distribution of those patients is the same. So there was no real change.

DR. GILMAN: So, again, this is a retrospective

look at efficacy and these are data we have not had a chance to examine in detail.

DR. MARSHALL: Well, Dr. Gilman, in all fairness, I didn't suggest this be done to demonstrate efficacy. It was simply really designed primarily to address the issue which the agency appropriately raised about classification of patients, and I wanted to show that you could very easily, as Dr. Grotta said -- he said it in two seconds. He is quicker than I am; it takes a bit longer -- to get the patients classified appropriately and it holds up very well using a very hard, easily determined validated measure.

DR. CUI: May I ask a question? How do you combine the studies?

DR. MARSHALL: I am not sure because I didn't create the slide.

DR. GILMAN: Use the microphone, please. Identify yourself and give us your answer.

DR. CORRIGAN: Yes, I am Dr. Corrigan again. I am just going back to your point and the committee's. It would probably be more useful here to look at the individual studies.

DR. MARSHALL: But the point is also this was not added together properly, as somebody pointed out. It adds up to 169, not 270.

DR. CUI: May I make some comments on this slide?

Actually, one thing I want to say is that this finding doesn't contradict our finding based on PMR2. This finding still tells you that the patient's situation is really bad; then you will see the larger treatment effect. This is one thing.

DR. MARSHALL: That has nothing to do with PMR2 because --

DR. CUI: Right.

DR. MARSHALL: -- we excluded the top two categories that make up more than half the patients.

DR. GILMAN: He understands that.

DR. MARSHALL: It has nothing to do with it.

DR. CUI: But it seems that the patients here have best motor scores less than 4.

DR. MARSHALL: Equal to or less than 4.

DR. CUI: Equal to or less than 4. So, those are severe patients. Right?

DR. MARSHALL: Absolutely.

DR. CUI: For the severe patients you will see larger mortality trends. That is the same thing. The second thing is --

DR. MARSHALL: It is not larger mortality trends. The issue is also that, although it was done to demonstrate validating the identification of the patient, it still supports the fact 63 is a strongly positive study.

DR. CUI: Another thing I want to comment on is the post hoc analysis. I agree we are doing post hoc analyses but the motivation for that is because of the subgroup analysis in nature. There is a substantial difference between our post hoc analyses and the sponsor's analyses in terms of the purpose. Our post hoc analysis is in order to use our results to show there is a significant finding --

DR. MARSHALL: But in all fairness, this is a response to your criticisms of the issue of patient identification, and that is the only reason it was presented, and how else could I respond by showing you that I can validate the data, using the scales used within the study, except to show you something that is universally accepted? And, I have shown you that.

DR. GILMAN: Dr. Penn?

DR. PENN: Larry, you could help us a little bit. If you had designed the study you probably would have stratified for a number of things --

DR. MARSHALL: Right.

DR. PENN: -- differently than it was stratified for.

DR. MARSHALL: Yes.

DR. PENN: And, you would have chosen different endpoints undoubtedly with the knowledge you now have. The

issue we are going to have is, given what we have got, not being prospective, how are we going to make these types of judgments?

The second thing I want to say that I agree entirely with being able to identify the population groups that are at risk. I don't think that is going to be a problem, grading the patients and having two grading scales. It grates us to have two grading scales but even if you didn't follow the, exactly, as you have shown, and you used reasonable clinical judgment about a bad patient and a good patient, they would fall into the right category. So, I think that is different.

The problem that we have to focus on, I think, is the one that we don't have prospective, appropriately stratified data to prove the point, and it looks like we are always teasing ourselves with Phase II, almost Phase III studies, and how do we deal with that crucial dilemma from the committee's standpoint?

DR. MARSHALL: Can I give an answer, Dr. Gilman?

DR. GILMAN: Yes.

DR. MARSHALL: Thank you. Well, first, I think the point about patient identification is a real one and you have addressed it, and I think there is some agreement among the clinicians in the room that we can appropriately identify the patients to deal with what is a purported

suggestion by the agency, which I don't think is supported, that there is increased risk in the better grade patients.

I think that the argument about the integrated nature of the analysis, the sort of "this is apples and oranges" or "apples and airplanes" has some validity but it is not entirely valid because, after all, we do change the doses of drugs all the time. The meta-analyses that have been published repeatedly and have been support in the PDR, allowed by the FDA, have allowed very different dosing, for example, for the prevention of infection from implants where the doses range over wide ranges and even different antibiotics the studies have been put together to show that you can then conclude that implant infection is reduced by the use of antibiotic therapy in a very broad sense. So, I think that it is not entirely fair, in my view, to attack those.

I think the issue of retrospective nature, of going back, is obviously of concern. At the same time, looking at it from the sponsor's perspective, in reading the letter -- and, obviously one can interpret it differently -- it seemed to me that the point seemed to have been established, and the neurosurgical community had established, as well as with a number of regulatory agency worldwide, that the effect was robust and demonstrated in males, and that the issue was whether or not you could see

an appropriate effect -- and I agree absolutely with what Dr. Racoosin said that when you are looking at risk you have to look for very small signals because they are very important and we must not do harm. Study 63 is a very robust study demonstrating efficacy. The trend in 65 was favorable a bit to the drug but did not show significance. If 65 was adverse I would not be standing here because my view would be that there is no evidence that this drug is effective in poor grade subarachnoid hemorrhage patients. But the fact is the trend in 65 was favorable, although not significant. We recognize that that is the case. I recognize that is the case. But the effect in 32 was remarkably robust, recognizing the endpoint was vasospasm. The effect in 63 was also very robust and the endpoint was prespecified. That is the best answer I can give you.

DR. GILMAN: Thank you. Dr. Burkhart?

DR. BURKHART: Yes, I would like to make two comments, first regarding the way we were approaching safety. I think the way we looked at it is that the findings in the non-subarachnoid hemorrhage studies actually did influence the way we looked at the subarachnoid hemorrhage studies. So, we went into those a little bit concerned. You had some unusual observations coming from the stroke studies and from the head trauma studies.

So, within that context I think there was

significant concern about what the bump in the relative risk meant in the low neurograde patients. Now, if you had found a cause to explain that increase I think you might have had a much stronger signal than you had and, as you heard from Judith, she was unable to find anything. Perhaps that was due to the quality of the data.

DR. MARSHALL: Well, I agree --

DR. BURKHART: So, I don't think we were arguing at all that there was a strong signal in the subarachnoid hemorrhage studies. I think, if anything, it would be weak if I were to characterize it. It is really the other studies.

The second comment is really to have you back up on the slides, back to the best noted score of 4.

DR. MARSHALL: That is quite a way back. They will find it. Go ahead.

DR. BURKHART: Well, I remember the numbers. Anyway, you said that about 8 percent of the patients would be moving from IV/V, out of IV/V using the best motor score. If you will notice in 63, there are 154 patients that were in IV and V and I believe that represents 102. So, that is quite a bit more --

DR. MARSHALL: No, wait a second. That is not what I said. What I said was this is the best motor score. Remember that that is the best they are. Some of these

patients are decorticate and decerebrate on the other side. If you take that into account and if you look at the data what you will see is instead of 8 percent moving 10 percent moved, but the distribution with regard to outcome is the same. And, I asked that question very early because, of course, that is a concern that you suddenly wind up -- you know, the funnel is getting smaller and smaller so, remember, this was only to look at the issue, again, worrying about the issue raised, I think very elegantly by the agency, of classification of patients. Can we identify the patients, particularly since you raised the issue of harm?

DR. BURKHART: But wasn't your point that using this approach you had 92 percent of the patients in IV and V? Didn't you show us that?

DR. MARSHALL: No, I used the Vs and said "do not obey commands" and have a depressed level of consciousness. But this is just sort of trying to make the argument about the issue of harm.

Can I go on, Dr. Gilman?

DR. GILMAN: Let's see if there are any other questions for you at this point. No? Yes, please continue.

DR. MARSHALL: Yes, if I can find out where I was.

[Slide]

Two other issues were raised by the agency. One

was the so-called deleterious effect of nimodipine. And, I think we need to review, one, what the agency approved nimodipine for. It was approved in subarachnoid hemorrhage patients with good grade and to reduce the incidence and severity of ischemic deficits. So often, unfortunately perhaps, what is said in the approval and what happens in clinical practice don't seem to be coincidental. Many neurosurgeons in North America lack confidence in oral nimodipine. Thirty percent of the patients in grade I in the United States do not receive the drug in grade I because of that lack of confidence in the oral form but the manufacturer saw fit not to apply for approval of the intravenous form which is used throughout the rest of the world. It is the standard of care worldwide in Hunt and Hess IV and V patients. The Petruk study which Dr. Wiers who was here, Chairman of Neurosurgery, University of Chicago, was involved with and probably knows more about nimodipine than anybody in the world, was a study at a dose 50 percent higher than the dose used and approved by the agency. The complication that led to the increased mortality is hypotension, which is well-known to be associated with excessive dosing. There are a number of other studies looking at grade IV and V patients using the more traditional Hunt and Hess which show no evidence of harm and, in fact, suggest benefit.

So, I think that the notion that somehow there is masking and unmasking of an effect of tirilazad on nimodipine, to me, seems very unlikely. In other words, the Petruk study, we believe, is an overdose study. The agency did not approve the drug at 90 mg, and there is a meta-analysis, which you may or may not like, was published in 1996 in the Journal of Neurosurgery dealing with the issue of overall outcome with nimodipine. So, in my view, I had a great deal of difficulty accepting the notion that there might be harm in association with nimodipine.

The last issue --

DR. GILMAN: Dr. Marshall, can we stop you there?

DR. MARSHALL: Absolutely.

DR. KATZ: I have a couple of questions. The notion that the increased mortality in the sick patients in the Petruk study were related to hypotension -- maybe a word or two about how we actually know that. I mean, there might have been increased incidence, I don't know, but that is one thing.

The other thing is I am interested, because we have looked through the literature and maybe we missed it but we didn't find any other studies that really specifically addressed the question of the effects of nimodipine in the bad neurogrades.

DR. MARSHALL: Could I call on Dr. Wiers?

DR. WIERS: Well, in that study, in fact, the only differences in mortality were in Hunt and Hess grade III group. There were virtually no differences in the IV and V groups. The difference -- there were only 7 patients, an excess mortality of 6 over 1 -- 7 patients who died in the grade III nimodipine group, and 3 of them died from rebleeding, 1 from a radiological disaster and 1 from a surgical disaster. So, there was no common thread conceivable between nimodipine and those deaths. I think most of the neurosurgeons in the world ignore the FDA decision which was based, I think, on a misinterpretation of this paper because there were no differences in the IVs and Vs; it was in the IIIs.

DR. MARSHALL: I heard his presentation. He was concerned in the worst grade patients about hypotension in the IVs and Vs although it did not play out in mortality.

DR. WIERS: We didn't see a difference in mortality.

DR. GILMAN: So, if I understand what has just been said, the FDA's approval for grades I through III is justified -- maybe not for grade III, but for grade IV and V it is not currently approved but it should be. Is that what we were just hearing?

DR. MARSHALL: Well, that is the clinical impression of neurosurgeons but that may be the triumph of

hope over fact.

DR. GILMAN: But then we have the problem of where this is published and what are the data that this conclusion is based upon.

DR. MARSHALL: Go ahead, Dr. Wiers.

DR. WIERS: Well, this paper was published in the Journal of Neurosurgery in 1998.

DR. GILMAN: Plus, do you know whether this was the paper upon which the FDA made its decision?

DR. WIERS: Well, this was one of the bases upon which it made its decision.

DR. GILMAN: Were there others showing different effects?

DR. WIERS: Yes, there was a much bigger British trial of all grades.

DR. DORSH: I am another neurosurgical consultant, Nick Dorsh, from Sidney, in Australia. In fact, although the study didn't specifically address bad grade patients with nimodipine in subarachnoid hemorrhage, the British trial that I am sure you are familiar with showed a very significant overall improvement, and they do mention that among the factors individually related to outcome were Glasgow Coma Scale, etc., and adjusting simultaneously for that treatment remained highly significant. That is the only other study which really included any large number of grade

IV and V patients.

DR. MARSHALL: That is the so-called BRANT study, run by John Picard.

DR. DORSH: The other study of May and others, also in England -- Edward May, I mean, not me --

[Laughter]

-- did show again a strong trend towards overall improvement but it only had very few grade IV and V patients. But I agree that worldwide experience, particularly in our country, where we are able to use the intravenous nimodipine, in which it is much easier to avoid the problem of hypotension, is that we would be very wrong not to use it in grade IV and V patients, Hunt and Hess neurograde or anything.

DR. GILMAN: All right. I guess that is the triumph of experience over science.

[Laughter]

DR. MARSHALL: I would like to move on to the issue of safety, which Dr. Racoosin has addressed in some detail.

[Slide]

She did present the entry criteria, and so on, and Dr. Grotta made a comment about the study. I think, you know, what happens with the tincture of time if you look at this is that if 81 had gone first, where there was a strong

trend in favor of tirilazad, then one would have clearly continued. Given the fact that the studies were extremely similar in terms of dosing, endpoint, etc., if one looks at the combined mortality in these studies at 10 days and 3 months I think it is very difficult, from my perspective, to say there is any kind of significant signal.

As Dr. Grotta pointed out not only here but previously, having read some of this papers, the whole issue of placebo mortalities in some of these previous drug trials is an issue, and I would point out that in the lubeluzole trial, at least one of them, the mortality in the placebo group was in excess of what one sees here.

So, I think the point is that the idea that there is a signal in stroke, from my perspective, is very hard to see and, again, remembering that we had really divergent trends -- 81 very favorable, small sample but almost 2:1 in favor of the drug with regard to mortality, and not quite the converse but an unfavorable signal in 88 against the drug. If you look at them together, which I think is appropriate, you basically see nothing.

[Slide]

We also looked at the issue about causes of death that has been raised, hypotension, and I am going to come back to that.

DR. GILMAN: Sorry, can I stop you there? Would

you go back to the previous slide?

DR. MARSHALL: Sure.

[Slide]

DR. GILMAN: What are you telling us? That if you combine the two studies --

DR. MARSHALL: Well, they are exactly the same. The mortality in one was exactly the opposite in the other, and if you look at the two together you see no difference in mortality in these two major stroke studies. This study was stopped, 81 was stopped because of the futility of going on because 88 was unfavorable. So, 81, at the time that it was stopped, was quite favorable to tirilazad in terms of outcome.

DR. GILMAN: They are very different in total N, are they not?

DR. MARSHALL: Yes, but the difference was rather large. In fact, it was disproportionate with regard to the impact. The mortality difference in the vehicle group which was unfavorable in 88 is essentially wiped out completely by a much smaller study, showing that the trend was stronger, in fact, in 81 in favor of tirilazad.

DR. GILMAN: Dr. Burkhart?

DR. BURKHART: Just to point out real quickly, the difference between vehicle and drug in study 81 all occurred late. You know, if you look early there is not as much

difference, I don't think, in 81.

DR. MARSHALL: We can look at the data in detail.

DR. BURKHART: Well, I think you are raising a philosophical point. When you are worried about safety are you going to focus on 3 months or are you going to focus on early exposure? I mean, you are going to have to decide because you may see different results. In fact, if you look at 88 you will notice that the relative difference between drug and vehicle declines over time and that the absolute difference between the two remains the same.

DR. MARSHALL: I think, in all fairness, you are looking at vehicle mortalities that are extremely low. As Dr. Grotta pointed out, many of these run around 20 percent. If you look at the differences in these studies and you look at 3 months, and while I recognize that the argument can be made with regard to when to look, how to look, about time, when you have such a small mortality to make the inference that you are seeing harm from an early signal, to me, I think is troubling because I don't think the signal is amplifiable and, in fact, you see exactly the opposite in the other study, in 81.

DR. BURKHART: Well, I guess the way I would look at it -- and I would not have combined the data but looking at them separately as Dr. Racoosin showed them, if you look at study 88 you see quite a bit of relative difference

between drug and vehicle within the study that reaches some level of statistical significance. What you do about that is another question.

DR. MARSHALL: Why don't we look at 81?

DR. BURKHART: I was just going to say you are absolutely right, if you look at 81 it is completely different. Now, does that mean that 88 is not true? I mean, I don't know how to answer the question. I can tell you that if you have one or two positive studies going the wrong way out of seven, eight, nine or ten, that is an unusual feature to see in a drug. We still approve drugs when that happens for efficacy but for safety that is an unusual thing to see in NDAs.

DR. MARSHALL: Well, I think we ought to put up 81 just in the interest of clarity in terms of the size, as Dr. Gilman pointed out.

[Slide]

This is 81, day 10, 8 patients, 14 percent, 6, 12 and then 3 months, 33 percent, 19 patients, 10 and 19.

DR. BURKHART: So, the point I am making is that is a remarkable change in the vehicle over time. Most of the difference that is happening, that you are pointing out about vehicle is occurring after the infusion, whatever that means and that is an unusual feature in that data set.

DR. MARSHALL: Well, I think the key is whatever

that means. I think the interpretation could certainly be -- you know, I am not certain there is any effect in stroke, and this is not the indication being applied for, but I think it is very difficult to say that you can show harm. Are we going to assume there is harm in the placebo group? I think obviously not. In my view, therefore, it is appropriate to combine this --

DR. BURKHART: It wasn't my point to say that 81 was showing harm.

DR. MARSHALL: No --

DR. BURKHART: Okay.

DR. MARSHALL: -- but what I am suggesting is that given the fact that the studies were quite similar, they are divergent, but what happens is you have a very low mortality, as Dr. Grotta appropriately pointed out, in 88 in the vehicle group and here what you have is a somewhat higher mortality in the vehicle group, and these both go about the point that has been in the lubeluzole trial and with other trials of agents which are potentially efficacious in stroke. That is all.

DR. GILMAN: For this committee the question really revolves around whether it is justifiable to add those two studies together. I think that is for each of us to determine in our own minds having heard the arguments on each side.

Please identify yourself.

DR. ALVES: I am Wayne Alves. I was a non-voting member of the 88 safety committee. The decision to stop the trial, while there was concern over the apparent signal in mortality, particularly the correlation with primary cause of death being neurological factors, the committee did not find significant issue with respect to the SAEs or AEs, except to note that the SAs were probably correlated with the increased mortality. They then considered the futility analysis. At that point, based on the imbalance in mortality, the study was underpowered and it was futile to continue. That served as the basis for their decision to stop, not the mortality difference.

DR. RACOOSIN: Can I interrupt for one moment, please? This is quoting from the abstracts, Volume 136, page 6, this is the abstract for study 81: The study was terminated after Pharmacia and Upjohn decided not to pursue the evaluation of tirilazad for this indication following termination of a similar multinational study (as recommended by safety and monitoring committee) due to an increase in mortality rate in the tirilazad group and the outcome of the futility analysis. So, this is a response to that previous comment.

DR. HANLEY: Dan Hanley, I am a neurointensivist and another Upjohn consultant. I would like to speak to this

issue of sample size and mortality in stroke. Understanding and respecting the safety issue, 81 and 88 are small studies which we just heard about.

[Slide]

This is a listing of some of the larger stroke studies and the mortality in the placebo and treatment arms from the literature, including the NINDS study, ECAS 1 and 2. You see that in larger studies the mortality ranges from about 10-15 percent to about 20 percent. Death has not been a selected endpoint in stroke studies because it is a low frequency event and you have to get to large sample sizes in order to get an adequate estimate. The sample sizes that were looking at safety are the ones that we can only look at for tirilazad, and it is smaller. When you do the combination you begin to see that the combined data approaches this range of 15-20 percent mortality rate, and looks similar to all the other stroke mortality rates out there. Whether that helps you with a drug safety effect at 90 days I don't know.

The other way to look at this data if you want to look specifically at the drug infusion effect is to compare to the day 10 data, which we will show you later on, of mortality in the subarachnoid hemorrhage group of patients. Those are individuals who are at the end of their infusion time period. That is the only other comparison that I can

see that you can make to allow you to look at the safety issue for drug effect.

DR. KATZ: We see this a lot as well. When there is an outcome when folks aren't particularly happy with the explanation is often that the placebo rate was too high or was too low. It is what it is within the trial. That is why you have a placebo group. It might be discrepant with previous placebo rates. There may be explanations for that. It might be that the trial is too small at that point. It might be that there are different patients enrolled. We don't know. It doesn't change the entire validity of the analysis performed in that trial, and that is what it is.

DR. GILMAN: But then that again raises the question as to whether one should combine two trials.

DR. TEMPLE: But in some sense it doesn't really matter because you get to look at the two separately and then you get to look at the two combined, and as usual in these cases, you don't quite know for sure; you are guessing to a degree.

DR. GILMAN: Yes, we may not have to guess if there were ten more trials but there are not.

DR. TEMPLE: That is right. But I think from our point of view it isn't that we think that there is proof positive that the complementary group is harmed, or anything like that, but it raises a concern that is not easy to

dissipate, that could make one think that you want to be quite sure of the benefit. I think that is the context that this is raised in. So, it is important to see how strong it is and how much worry you should have.

[Slide]

DR. MARSHALL: I am not going to spend much time on this because I think Dr. Racoosin covered this, and I think you are left to your own conclusions as to whether these very small numbers represent any kind of a signal. Obviously, my conclusion is not.

This is 81 and 88, both separate and combined. So, you have the opportunity that you are left to your own devices to add this up, since I have gotten into trouble adding it up together myself.

[Slide]

I would like to move on now, if I can, to talk a little bit about the head injury studies. As I pointed out in 36, I was the principal investigator in the trial. Although involving Europe, Australia and Israel and it occurred abroad, our clinical trial group managed the trial. Study 17 was a domestic study in the United States and Canada. As Dr. Racoosin pointed out, this study was stopped after having essentially accrued all its patients, but accruing additional patients to replace one center whose performance and quality of data was very poor and 17, as she

indicated, was stopped because of concern that there was excess risk in the tirilazad group.

She also pointed out that the issue of herniation early on, between day 3 and day 7, was a major potential signal of excess risk from brain swelling or increased intracranial volume, as we chose to call it in analyzing the data.

[Slide]

That just summarizes what I said. As she also pointed out, 36 showed no evidence of harm but some suggestion that in traumatic subarachnoid hemorrhage there was a beneficial effect of the drug.

[Slide]

These were some of the issues when I got the data and suggested that an advisory committee to the manufacturer, at that time the Upjohn Company, be put together. Some analyses obviously pre-recognized prognostic variables. If you look at bilaterally unreactive pupils, most of the people in the audience who are in clinical practice will recognize that this is probably the most ominous predictive signal, and there was a 7 percent difference between the vehicle group and the tirilazad group. There was an excess frequency of subarachnoid hemorrhage and also pretreatment hypotension.

Certainly, I am not up here to tell you that these

variables are completely independent. They are not, but they are not entirely interdependent either. In terms of the number of patients in terms of pupillary reactivity --

[Slide]

-- what you see is that there were actually 27 with bilaterally unreactive pupils and in the group with 1 reactive pupil only there was an excess of 9 in the tirilazad group.

If you then look at the mortality data, what you see is exactly what you would expect. That is, overall the population has a death rate of about 13 percent. The patients with bilaterally unreactive pupils have a mortality rate of approximately 3-4-fold greater. If you then calculate it, you essentially wipe out the difference.

The other point to make here is that when you do an analysis of this you have to weight it for its impact on mortality, not just throw it in as one of several variables. In addition, there were differences in the frequency of extradural hematomas, again unfavorable to the drug. And, if you look at the cause of death, identified by Dr. Racoosin, which is increased intracranial pressure and, therefore, herniation, this is exactly what you would expect, and you would expect those deaths to be early, between days 3 and 7.

So, the safety monitoring committee which Dr. Andrew Maas, who is the Vice Chairman of the European Brain

Injury Consortium, and Dr. Franco Servadi, the Secretary, came to participate in, concluded that there was no evidence of harm in the 17 study and that the difference between the vehicle and tirilazad group could be completely explained by the difference in pupillary reactivity prior to randomization if one did the analysis in an appropriately weighted way, and there are algorithms for weighting patients which have been provided to the agency over the last several years by a number of sponsors in anticipation of other drug trials going forward.

So, as I said, I think that the evidence in that study was fairly compelling, in 17, that there is an appropriate, not post hoc, explanation for the difference in outcome in mortality, and that it fits very nicely with the observation of early deaths from herniation.

[Slide]

I would like to move on now finally to the issue of increased risk in subarachnoid hemorrhage. In the filing from the FDA the concern about the discontinuation rate of the drug in patients was raised as a potential signal of harm. I would point out that that certainly is appropriate but it is also appropriate to look at mortality, the frequency of adverse events and the frequency of serious adverse events in terms of what you are seeing in terms of the signal.

[Slide]

This is the SAE leading to discontinuation analysis. One point to make is that after the first studies it was decided that changes in liver enzymes of greater than 3-fold would lead to a discontinuation. So, you see a difference in the frequencies here in this study as opposed to here, that is, of the total number of patients discontinued.

If you look at the calculated frequencies of discontinuation from certain adverse events -- brain edema shown here, intracranial hypertension shown here, and lung edema which I am going to speak about in a moment --

DR. RACOOSIN: Could I ask one question, please? I am sorry to interrupt. Your patients dosed under tirilazad 6, 644, which studies are those from? Is that just men?

DR. MARSHALL: That is all studies.

DR. RACOOSIN: And this is the pivotal trials?

DR. MARSHALL: It is 32, 29, 65 and 63.

DR. RACOOSIN: Thank you.

DR. RUPPEL: Larry, that does include all placebo-controlled SHA studies that had a 6 mg arm.

DR. RACOOSIN: So that would be 19 and 7 as well?

DR. RUPPEL: Not 19 because it did not include nimodipine. It is all placebo-controlled nimodipine studies.

DR. RACOOSIN: Thank you.

[Slide]

DR. MARSHALL: In looking at these adverse events leading to discontinuation, as I said, we see very little. Then looking at 14-day adverse events with frequencies greater than 10 percent, remember, in a population of people with a bad disease, cerebral vasospasm, cerebral infarction, brain edema.

I want to make one comment with regard to something that Dr. Racoosin said, which is that it was not clear, and it may not have been clear in the CRF as to the diagnosis of brain edema. That is made on CT scan by the investigator. That is a sort of standard mechanism within all of these trials. For intracranial hypertension, which is a rubric under which many of these things fit, one saw nothing. Hydrocephalus, pneumonia and the issue of lung edema.

Now, the concern raised in, I believe it was 63 and you can correct me if I am incorrect, Dr. Racoosin, was the issue of the sort of acute respiratory failure issue, acute pulmonary failure and ARDS. But I think it is important to point out that when you have a significantly improved survival in a group of patients who are being resuscitated with large volumes of fluid you have to live to get the treatment, and if you die then, therefore, the frequency of the events in absolute terms would go up but

the percent frequency is not likely to be substantial. This is an observation that has been made in a number of trials, those sponsored by the NIH, both head injury and other catastrophic diseases, that is, if you have improved mortality it is inevitable that some of the other treatments such as hypervolemia therapy can lead to complications. The issue is, is the quality of the complication worse than the disease you are treating or does it lead to an adverse event such as death? That clearly is not the case.

So, I was struck by your report and was concerned about it, but in looking into the data I did not feel that there was anything, and in terms of looking at the actual percent frequencies one also sees nothing.

[Slide]

This is looking at serious adverse events in this population, the total patients, and again one does not really see anything in terms of a signal increasing, suggesting increased risk. Again, the overall impression here I think is that there is no substantial or identifiable increased risk of any kind of brain signal that is a disease or process suggesting harm.

[Slide]

Then looking at the others, and I have already covered the issue with regard to respiratory. In the cardiovascular area there was the issue of modest

hypokalemia. I think most of the clinicians in the room would recognize that in this population of patients who are cared for in critical care units a potassium of 3.2 is not a substantial risk and will not lead to patient harm.

[Slide]

Then there was the issue of skin disorders. There was an increased incidence of rash but, again, I think when talking about a disease which has a potentially fatal outcome that is a relatively small concern, although I think it clearly was identifiable in the patient population.

[Slide]

So, I think that one can conclude with regard to safety in terms of the subarachnoid hemorrhage patients, and I have also discussed the others, that the concern about serious cardiac and pulmonary adverse events seems to me to not exist, and that they occurred at the expected rate; that the safety profile is good for CNS events; and that there were no clinically relevant changes in EKG or in cardiac, liver or renal laboratory results, recognizing that modest hypokalemia did occur in this patient population but, again, from my view as someone who is concerned with critically ill neurosurgical patients, not of any substantive concern.

[Slide]

One can conclude that the mortality in patients with bad grade subarachnoid hemorrhage, appropriately

identified, can be reduced by almost 40 percent; that the outcome occurs despite baseline imbalances against Freedox, particularly in posterior circulation location and intraventricular hemorrhage; that the drug has a good safety profile and has a favorable risk/benefit ratio.

DR. GILMAN: Where did the figure of 40 percent come from?

DR. MARSHALL: From the 37.5 overall reduction in mortality in the trials when they are added together.

DR. GILMAN: You are adding all trials together?

DR. MARSHALL: Yes.

DR. GILMAN: I see.

DR. MARSHALL: Thank you very much.

DR. GILMAN: Any questions for Dr. Marshall? Again, we have expressed some concern about adding together trials from different studies. Dr. Corrigan has only two slides. I assume that you can be fairly succinct and then we can go to lunch perhaps.

**Concluding Remarks**

DR. CORRIGAN: I have one slide so I will shorten it even further.

[Slide]

Subarachnoid hemorrhage is a catastrophic disease for which there is no pharmacological agent indicated in the United States. We believe that the substance of the evidence

that we have presented demonstrated treatment effect for patients who are easily identified as being poor neurological candidates for it, and we urge the committee to recommend in favor of approval to the FDA. Thank you, sir.

DR. GILMAN: Thank you, Dr. Corrigan. Any questions for Dr. Corrigan?

We will break for lunch in a moment but first I would like to caution the committee not to discuss this agent or this morning's proceedings over lunch. This is all supposed to be discussed in public only. We will take a one-hour break and we can reconvene at 2:25.

[Whereupon, at 1:25 p.m. the committee recessed for lunch, to resume proceedings at 2:25 p.m.]

- - -

A F T E R N O O N P R O C E E D I N G S

DR. GILMAN: Before we ask for public opinion, the sponsor has asked for two minutes for wrap-up time. We want to ensure that the sponsor has every opportunity to present their case so, please.

DR. MARSHALL: Thank you, Dr. Gilman. I just wanted to point out that in the discussion, which got rather heated this morning about the integrated analysis, that it is important to recall that in study 32, first, the entire study was positive; second, the study was positive for its primary endpoint, cerebral vasospasm in men; and, third, it was positive in the group IV and V.

I think looking at the entire picture, as I said previously, with regard to outcomes we have a disease for which there is no treatment approved, that is, poor grade subarachnoid hemorrhage. We have a positive study in 63, prospectively designated, and we have a positive trend in all the studies, including 32 which I referred to initially.

So, I think while there certainly are very legitimate and appropriate criticisms of the construct of combining them, looking at the studies as a whole from a clinical perspective we have reduced mortality, the number of deaths is reduced in each instance when this drug was given in a controlled trial with positive and significant results in 32 and 63. Thank you.

DR. GILMAN: Thank you. Any other comments from the sponsor? Have you had a chance to present all the information you would like to present?

Then, from the agency, is there anything further the agency wishes to present to us? If not, then is there anybody who would like to present at the open public hearing?

**Discussion by Advisory Committee**

If not, we will turn to the committee. The agency would like us to address the issues that have been raised in Dr. Katz's narrative. So, if you open your red books you will find the narrative on page 23. There are 6 questions. We will just go through them one at a time. I will read out the first one:

We are very eager to hear your views on these issues which include but may not be limited to the following. One, a very critical question is whether or not there is any bona fide finding across these four studies that can be considered to have been independently replicated or corroborated. That is question one. Let's deal with question one first.

So that the agency will have the sense of this committee I will just lead off by attempting an answer to that question. In my view the answer to that question succinctly is no. I find a series of studies that have had

different Ns, different populations, that have been examined with different doses over time, and I find very little that I can carry over between studies with respect to result. Consequently, I believe the answer to question number one is no.

Now, let me go around the table and see what my colleagues think about that, anybody who wants to comment on that?

DR. BROOKE: I would agree with your comment.

DR. GILMAN: Dr. Drachman?

DR. DRACHMAN: I would agree with that. I believe the way these studies really have been used is more like pilot studies. Each one led to the next, and then a backward look. That is not replicating. That is moving along. We have heard more and more about how to pick out the most at risk subjects, and I think the ideas are good. Looking at those with poor grades on the various tests makes sense, simplifying really does make sense. But this is all retrospective. So I view it as pilot studies, not leading to a confirmation.

DR. GILMAN: Dr. Temple?

DR. TEMPLE: Just to be sure we understand, there was one study that was done, based on a lot of prior experience picked out, that I believe the sponsor would say was prospective, number 63. So, it is a very good question

whether that constitutes replication. It is not two independent studies, each meeting something, but they could argue that there is a kernel of replication in that they formed a hypothesis from a series of studies and then did a study to check it out. Now, they didn't do two studies to check it out, obviously.

DR. GILMAN: I view that as one study in which a subgroup, identified prospectively, did have a beneficial outcome on drug. But to answer the question that Dr. Katz has posed, I do not see replication across the four studies that we have examined.

Let me ask my other colleagues, other thoughts? Does anybody disagree? Dr. Kawas?

DR. KAWAS: I don't disagree but I am not completely sure what I think. I have four studies and it has been very hard for me to put the interpretation of these studies together. I see some positive effects in each study; I see some negative or non-effects.

I think though also that this is the only time among the questions that are posed to us though that we might have the opportunity to come back to the issue that was raised by Dr. Brooke. If we were to design a study for subarachnoid hemorrhage, it seems to me that the grade IV and V to prevent mortality would not be one of the first things that we would be trying to do clinically. And, the

ethical issues that Dr. Brooke raised and everything sort of factors into this.

If the question is are there two studies that convincingly tell me that this drug works in that group of patients, I think I agree the answer is probably no.

DR. GILMAN: Other comments?

DR. LACEY: Mr. Chair, I would say no. I did not feel that there was substantial evidence for the proposed indication in the things that I saw. I am left with too many conflicts regarding issues related to gender. The issue of women and the age of women as to whether or not the dosage is appropriate for the older woman that has been selected for the study so far -- I am not clear on that at all. Various other issues that have come up have left me with opposing points of view. I see a lot of promise and hope related to a drug that could come into play at a later time and, with that then, I would rest my point in that I say no for today.

DR. VAN BELLE: I would agree with that. My concerns are, first of all, the changing in the endpoints, what I call a lily pad analysis. You step on one lily pad and move to the next one. It just does not hang together in terms of a confirmatory approach.

In addition to that, if you look at the sample sizes for studies 32 and 29 in this neurograde IV group you

are really talking about very small sample sizes, and that really leads me also to be concerned about the effectiveness of the overall study.

Then, as was pointed out, the fact the results from study 32 were used to generate the hypothesis for the subsequent studies clearly does not make it a confirmatory study.

DR. GILMAN: Don't those creatures that move among lily pads hop rather than step?

[Laughter]

DR. GROTTA: I think I feel pretty much the same, except that I am left with a nagging notion, when all is said and done, that actually what has been shown on the positive side does appear to be a consistent signal in all of the studies that there is a biological activity of this compound. As was brought up earlier today, even in the desperate clinical situation that these patients are in you would like to be able to feel that we could move ahead. But having said that, you would like to see a study that is positive for the particular population and for the particular endpoint, and not one of these studies was primarily focused and designed to test whether the drug is effective in this subgroup of patients, that is, women or men with severe grade hemorrhage at the appropriate doses. Even the final study, 63, even though it was a prospective

analysis, it is only a subset of those patients and, consequently, the numbers, as has been pointed out, upon which the conclusion is to be drawn are very small.

So, I am just not confident enough that what we are seeing isn't type-2 error, and I feel like another study needs to be done focusing on that particular population, in order to make me convinced that the drug should be indicated.

DR. PENN: I would weigh in that I am afraid we are in a type-2 error situation also. As a clinician, I would love to see this available after a proper study and, unfortunately, the business of taking it each step led each subsequent study a little bit astray and that is the trouble we have gotten into here because it appears that there may well be a good biological effect, but the problem is we have to deal with the data as it comes to us not as we would wish to have it.

DR. GILMAN: Thank you, all. Question number two, the question is related to practicality of determining an unambiguous identification of the proposed subgroup of patients with subarachnoid hemorrhage who might be candidates for treatment, assuming a finding of substantial evidence is made. In other words, is it possible to identify a group of patients who would benefit from this medication?

I would say yes, that could be done. We have heard

a couple of different ways of doing this. The Glasgow Scale is one way of looking, as Dr. Marshall did; using another approach or a series of approaches could be taken. What I would urge is that a consistent approach be taken and used so that we don't have to go retrospectively and say, well, the original protocol was incorrect, we should be looking at another way of analyzed these data because we then get into the problem of a retrospective study. So, I think the answer to question number two is yes. I think it should be possible to identify a subgroup, if one exists, that will respond to this medication.

Maybe we should just go around the table, starting off with Michael Brooke.

DR. BROOKE: Yes, I think subgroups can be identified in a prospective trial. I rather like the way Dr. Marshall identified some of the potential subgroups.

DR. GILMAN: Dr. Grotta?

DR. GROTTA: Yes, I am not bothered by the methods that were used and I think it can be identified. I guess the problem I have is do we know the group? While it is tempting to think that we do know that it is the severe patients I am not sure. That is what bothers me in part. We may be jumping to a conclusion based on data that has small numbers and, thereby, depriving patients who might otherwise benefit from being included in another larger trial.

So, I don't know the answer to the question about which group should be included, but I suspect there is a group hidden in this population that will benefit and can be identified.

DR. GILMAN: Dr. Van Belle?

DR. VAN BELLE: I am not a clinician but on the basis of the evidence presented here it would seem to be reasonable, but I would not want to state this too strongly either way.

DR. GILMAN: Dr. Lacey?

DR. LACEY: I would agree more with the statement just made. I am intrigued by the idea that there seemed to be some promising groups but I am not clear at this moment about what those groups are. So, I can't say that they definitely can be identified. I think there is promise that they can be identified.

DR. GILMAN: Dr. Penn?

DR. PENN: I think there is no problem clinically about identifying these groups, and that is not a matter of major concern to anyone doing these studies in a practical way, and the FDA's concern with how one subgroup fits into the other with the two rating scales is not a concern that would in reality be of great importance. I think you can find a lot of ways to identify the sickest patients, and there may be a slightly better one than what was used in

this study but I would have no problem with the way they did it.

DR. GILMAN: Dr. Kawas?

DR. KAWAS: I agree. I have nothing to add.

DR. GILMAN: Dr. Drachman?

DR. DRACHMAN: I agree, yes.

DR. GILMAN: Good. Thank you, all. Yes, Dr. Temple?

DR. TEMPLE: Let me be sure I understand. The concern might be that if you found a convincing effect in a group, that if you defined that group some other way in practice the conclusion might not apply to that group. But what you are saying, if I understand it, is the definition, for example that Dr. Marshall suggested, is close enough to that group so that that would not be a major worry if you were convinced that it was effective in that group.

DR. PENN: I think we can find the sickest patients and that probably using a whole bunch of scales you will not be off significantly. Eight percent was the difference that Larry came up with for a different rating scale that worked just as well.

DR. GILMAN: Let's go to question three then. Any concomitant increased risk in the complementary subgroup of patients, neurogrades I through III, which has ramifications related not only to the potential inappropriate treatment of

these patients but also for determining the validity of the beneficial finding in neurograde IV/V patients.

This sounds a little bit cryptic. It refers to the question of whether the medication if approved for, say, grade IV/V would then be used appropriately if it were found to be damaging to grades I through III. That was the thrust of this point.

DR. KATZ: Yes, that was part of the thrust. The other part had to do with the fact that there was a worsening in III. What did that mean for the validity of the primary finding, which is benefit in IV and V? I am just clarifying what the last half of that question meant.

DR. GILMAN: Right. So, first, I don't think anybody can control how physicians will use a particular medication once it has been approved. The approval process can go through the FDA and there can be clear labeling on the drug but then, once it is available, physicians are free to use it as they see fit. If there is a clear message though that this drug could be beneficial to a certain group of patients and damaging to another group patients, and the evidence is convincing, then I think its up to committees such as this one to say, yes, this drug is effective, assuming it passes safety evaluations, then the committee presumably would say it is safe or has relative safety but it should strictly be used for certain classes of patients

and no others. I think that is a reasonable outcome. With respect to this particular drug it has not yet -- going back to question number one -- been shown with replicated studies that it is effective in a certain group of patients and that it is definitely harmful in another group of patients. There certainly is evidence suggesting each of those may be true but it is not yet clear.

DR. BROOKE: I will start off. My favorite position is the top of the fence, and I don't think it has either been proven or disproved that this drug may be harmful. I listened with interest to the interpretations but I think I am not convinced that it is harmful and I am not convinced that it is not harmful. I think an important question is would I be comfortable recruiting patients for a trial on this drug knowing what I know about the potential harm, and I think I have to say the risk to benefit ratio would be low enough that I wouldn't hesitate to recruit patients in a trial but I am going to sit on the fence.

DR. GILMAN: Dr. Grotta?

DR. GROTTA: Well, the data certainly don't prove one thing or the other. That is for sure. But I don't see that that causes me significant concern about the safety of the drug.

DR. GILMAN: Dr. Van Belle?

DR. VAN BELLE: The issue is I think if this is

going to be applied to the grade I through III population. If that is not the case, then I don't think there is an issue with respect to grades IV and V. So, I am not quite sure yet what the question is. I would think that the sponsor would not be asked to do a randomized clinical study on patients in grades I through III. I mean, that just wouldn't be done so I am not quite sure what the question is.

DR. GILMAN: Dr. Katz?

DR. KATZ: I guess what I meant by the first part of the question was do you think there is an increased risk, an unacceptably increased risk or an increased risk in the I, II, III patients? I mean, that is really in a straightforward way what I would like to hear people address.

DR. GILMAN: Maybe I didn't address it clearly. There has been evidence shown suggesting that there may be increased risk in the I through III group. I am not convinced that that is correct; I am not sure that is true as yet. I don't think there has been enough data presented for us to be certain about that issue. Do you want to comment further?

DR. VAN BELLE: Let me just make sure that I position myself on one side of the fence or the other. So, I would say that in my opinion the evidence has not been shown

conclusively of substantial side effects.

DR. GILMAN: Not side effects; damaging effects upon a certain group is the question.

DR. VAN BELLE: Okay, damaging effects on a group not intended for therapy.

DR. GILMAN: Yes. Dr. Lacey?

DR. LACEY: I will say that from the point of view of the patient, if I were a patient having been graded I through III of the neurogrades I would be concerned. I would opt for no treatment over this treatment at this point on the basis of the information I have seen. If I were an older women, which I am not --

[Laughter]

-- one to be considered for treatment with the drug I would be concerned that I may be getting a heavier dosage than I need. So, I am concerned about the idea that perhaps we have not been convinced of the safety of the drug in those for whom it is proposed.

DR. GILMAN: Dr. Penn?

DR. PENN: I don't have any major safety concerns. If there were a robust effect of the drug in the group that is intended to be treated I think the risk/benefit ratio would say to go ahead and treat that group. We are talking about really trying to avoid a misuse of a drug in a group of patients, but I think that that is a question of

physician education and should not concern us as a major point.

DR. GILMAN: Dr. Kawas?

DR. KAWAS: I have the same amount of confidence with regards to the deleterious effect on groups I through III as I have with regards to the beneficial effects in IV and V. Since everyone is talking about if they were a patient, if I were a patient, with the data I have seen, I would feel comfortable to be randomized if I were a subarachnoid I to III to a study involving this agent. If I were a IV or V I might have to think twice.

DR. GILMAN: Dr. Drachman?

DR. DRACHMAN: The risk would be modest if the benefits were real. Notice that I am using the subjunctive for the condition contrary to fact, and because of that, you know, it is so conditional that I really do not feel that this -- it is a moot point.

DR. GILMAN: Dr. Grotta?

DR. GROTTA: I would just like to come back to one point, actually going back to the efficacy issue. I don't feel that the efficacy is clear in either women or men at this point. I think if further studies are done I would like to see the benefit corroborated in both populations and at the doses that are proposed, that is in men at the lower dose and in women at the higher dose.

I don't want to open another discussion but we didn't really answer the question as to whether the 6 mg dose is correct in men. That was the highest dose that was tested. I guess in other disease entities is when the dose has gone above 6 mg/kg and I guess that is where side effects were possibly seen. But, in any case, I think if further work is to be done it needs to be done in both populations.

DR. KAWAS: I did want to add that of the safety data, far and away the most compelling evidence and disturbing to me was what appeared to be a possible dose response in the sense that the higher doses resulted in the excess mortality. Although I doubt that the trend was significant, each successive dose had an increasing point estimate and I think that is the most worrisome of the safety data rather than the other findings.

DR. GILMAN: Dr. Temple?

DR. TEMPLE: You have touched on this a little bit and maybe you will want to defer it, but one of the things we are going to need to ask you about is what further studies are needed and appropriate. It is just worth mentioning that in study 32 there was no particular distinction between severity. There was an effect in men in that study seen in both groups. So, it is not obvious that that group needs to be abandoned either. But we will ask you

whether you think it would be sufficient to study just the sicker ones, but you can do that later.

DR. GILMAN: Dr. Temple, I would like to do that, if you don't mind, when we get to questions one and two that have been posed for us.

Any other comments on question three? If not, let's go on to question four. Question four concerns the potential deleterious effects of nimodipine in neurograde IV and V patients and the effects, if any, on the analysis of the trials.

Now, this brings up a point of considerable question in my mind because we heard that nimodipine has been approved for grade I through III with the HH type evaluation of patients but it is deleterious on group IV/V cases. Yet, we have heard from the neurosurgeons present that it is state-of-the-art to use nimodipine in essentially all patients, grades I through V. I find myself puzzled by that situation, and I was attracted initially to the thought that perhaps should another trial be conducted that trial ought to be on patients not treated with nimodipine. Now the question is whether it would be appropriate or inappropriate to withhold nimodipine from a group and that is a significant problem.

So in the light of the information we have heard today I am puzzled by this issue. I am not sure I can come

up with a formulation for you at this point. Let me ask my colleagues. Let's go in the other direction this time. Dr. Drachman?

DR. DRACHMAN: That is a very difficult question. The recruitment of candidates might be badly interfered with when someone would approach the family members -- the patients clearly wouldn't respond themselves, saying "here's what I usually do but because this trial must be done without nimodipine I am recommending that we forgo the customary, although non-labeled, use of the drug and use this non-approved drug or placebo in its place." So, my own view is, first of all, that the evidence showing that nimodipine was an additional risk was rather thin, and the ability to get around and deal with another study would be seriously flawed or obstructed.

DR. GILMAN: But would you comment on your level of conviction that there is a deleterious effect of nimodipine on grade IV/V or can you say?

DR. DRACHMAN: Low level.

DR. GILMAN: Low level?

DR. DRACHMAN: Low level of worry that nimodipine may have been a major contributor and should not be used in future studies; that it is risky and should not be used with tirilazad in a future study -- low level of concern.

DR. GILMAN: Thank you. Dr. Kawas?