The question is can one be informative to the other in some way because if we conclude that one can't be in formative to the other, we are left with a situation where to get a drug approved for either indication you need two trials in both indications when you know that the populations in both indications are going to have a lot of overlapping therapies. It would be an enormous task, and I think the standard, so far, in excess of anything else that any other sponsor has to do with the FDA for any other condition. I see the similarities more than the differences.

DR. THADANI: Do you want to amplify further on the ACS because all the clean, large trials with the oral agents are going in the wrong direction? You have OPUS with 16,000; you have SYMPHONY; and I know Eric Topol presented at the European Congress the totality of the data on IIb/IIIa orals is not positive.

DR. CALIFF: I am glad to hear you and Cindy believe Eric. That is probably the first time --

[Laughter]

DR. THADANI: -- and it is very relative to the generic structure, and if you include the ACS I am more or less convinced that there might be a difference in agents. I will buy that. But I don't think you can lump that. I think he should have shown the oral data here. I know it is not published but I am sure you have slides somewhere on the

oral data.

DR. CALIFF: I think I have to respond to this. First of all, David is not in a position to show that data. He is still a fellow. He is teaching me a lot. It is hard to believe he is still a fellow but he is.

We do have that data. It is preliminary. And, here again it is a matter of taste. Do you believe that a chronically oral way to administer drug, well after the event has occurred, is the same as an intravenously administered drug? If you do believe that, you know, we felt that for this meeting it is an issue well beyond this meeting. There are data about at least one of those trials that have not been discussed publicly and won't be for some time, that may not be consistent with the message of the other trials. So, I think that to bring that into this meeting would have not been appropriate.

I do think it is appropriate to at least raise the concern that the oral agent data in aggregate is trending towards no effect, and --

DR. THADANI: EXCITE is relevant here. Excite data was on ACS and the platelet inhibition is as good --

DR. CALIFF: Except that -- just one specific thing about EXCITE briefly, and the data haven't been published so I am uncomfortable saying much about it, but having it sent for publication by the first author, I will

just say that the effects in the first 48 hours are positive, which is the period of time in which these intravenous drugs would have been given. So, I think it is a very complicated issue, not quite as simple as a negative trial that would relate directly to what we are discussing today which is acute intravenous high-level therapy.

DR. PACKER: Why don't we ask Tom for his views and, Ray, you will get the last word.

DR. FLEMING: Actually, there are quite a few issues that have been brought up over the last several hours that I would like to comment on. What I will try to do is highlight comments on two or three of these main issues and defer the remainder of the comments until other points during the meeting when they will naturally arise.

Ray had asked me to be prepared to at least briefly comment on the statistical issues in conduct and interpretation of meta-analyses, and I will start with that but I would at least like to touch on some of the discussion here as it relates to the ACS versus PCI settings and the persistence of effect discussions that we have just had over the last half hour or so.

Beginning with meta-analyses, I think I can be brief. I think David has given, in my view, a very balanced assessment of what you can learn and where the cautions have to be in that analysis, but I will just briefly reiterate.

From my perspective, there are really two principal objectives that I see as motivating meta-analyses. One is to get further insights regarding more precise estimates regarding an overall treatment effect. The second is to get -- maybe I would call it descriptive insights regarding how generalizable the results are and how consistent they are across trials, across populations and across interventions.

I think many people will argue that you are most comfortable doing a meta-analysis in those settings where you are pooling similars. You are pooling similar patients, you are pooling similar protocols, you are pooling similar interventions. I think there is a tradeoff. When you are pooling similars, then I believe there is much more credence that you can give and interpretability that you can give to an estimate of the global treatment effect. If our goal is to come away estimating what is the magnitude of effect of this class of agents on death/MI at 30 days, I am more comfortable with that inference if I am pooling similar populations and similar interventions.

Of course, the downside to that is I get less generalizability insight if I am only looking at results across similar settings. So, the reverse setting is where you are looking at diverse populations and diverse interventions. In those settings, we are able to get more clues about generalizability and consistency but in that

setting we have to be much more cautious about using the global point estimate as something that we would believe applies to all of the populations and all of the interventions that were studied.

I think Freeman, Furberg and DeMetz give a nice summary of what are the concerns that can arise in terms of where dissimilarities are. They refer to dissimilarities in interventions, dissimilarities in the study populations, in the length of follow up, in the measures of response and the quality of data.

So, how similar or dissimilar are we in this setting? Are we looking at settings where everything is similar and we have the opportunity to get a point estimate that we can attribute to all populations and to all interventions, or do we believe that there are important dissimilarities? I might point out, and Rob has already made this point, that to do a meta-analysis such as this where we have A compared to placebo and B compared to placebo, and then try to conclude that this tells us something about A versus B is something that we have to be very concerned about or very cautious about.

So, specifically, I have great concerns in meta-analyses such as these in being able to come out with conclusions that different agents are, in fact, equally effective or unequally effective when all of them have been

compared to placebos and none of them have been compared head-to-head.

Dose issues also are important when I talk about different agents. So, it is not just different specific agents, but if we are looking at different doses of suboptimal doses those issues have to be taken into account.

Study populations -- this is really a key issue, particularly as it relates to ACS and PCI, and I am going to come back to this one. I want to comment a bit more on that.

Length of follow up and endpoints -- these are also significant concerns here because we have talked about several different endpoints, whether it is death, death/MI, plus/minus revascularization at 2-3 days or at 30 days or at 6 months. But issues arise as well here, using a clinical evaluation assessments or investigator assessments, and those issues as well have major influence on the nature of estimates of effect that we get even within given trials.

Quality of data certainly also has to be taken into account. Those studies that are done with higher quality being pooled with those that are not is an issue that would be a general concern.

An issue that I want to defer is one that relates to the temptation to do data dredging from a meta-analysis. I will come back to that issue later.

What I would like to do is turn to the specific

issue of acute coronary syndrome versus PCI. Marv has made comments about his sense in looking at these data, that the overall estimates, either with individual trials or with David's meta-analysis, certainly indicate to me as well that the magnitude of benefit could well differ in these settings. In fact, there is biological to expect that that could be the case.

I would like to go beyond that to discuss briefly the temptation to use data from one setting to conclude efficacy for another. Specifically, can we use data from trials that were specifically acute coronary syndrome trials to determine efficacy in the PCI setting? We have had some discussion already here about if you look within the acute coronary syndrome setting, can you break it down into those that received PCI versus those that didn't and get insights as to how the treatment had its effect and, specifically, was it mediated through an effect in those people who had PCI. That seems to me to be an interesting activity or attempt but it is extremely difficult to interpret the results. In fact, I consider it treacherous -- treacherous to do analyses in the acute coronary syndrome trials and to conclude that the treatment is effective in PCI or in non-PCI.

Let me be very specific. This is how these analyses typically might be done, take a population where at

time zero -- time zero, where you randomize, where you initiate the intervention is an acute coronary syndrome population for which subsequently there may be PCI. So, we would like to say, well, what would happen if there had not been PCI in this population? Can we use these data? Well, what we will do is we will censor people at the time of PCI. Well, that is a marvelous statistical technique, although the fundamental assumptions statistically to make it valid are probably completely invalid because what you have to assume when you censor somebody is that their subsequent likelihood of having the event, which would be death or MI, is exactly the same as people for whom you wouldn't have considered PCI. You have two people moving along here after, let's say, two days. One person undergoes PCI, the other one doesn't. You are looking at trying to estimate time to death/MI. You are going to use the experience in the person for whom PCI was not judged to be necessary to impute what the actual death/MI experience would have been in the person for whom you offered PCI had you decided not to do so. And, what we call that is informative censoring, and that informative censoring assumption is clearly violated.

The other approach that you might use is you are going to use these data, where time zero was acute coronary syndrome randomization, and try to infer or try to conclude what the benefit would have been if time zero had been

initiation of time of PCI. Well, we have at least three issues that come up here. First of all, treatment didn't begin in this setting at PCI; it began at some point before that. The second issue is how do you deal with an MI that occurs between randomization and the time of PCI? The third issue is you are having to assume that the treatment itself had no influence on the timing and whether you did PCI.

So, let me just summarize this to say again it is treacherous to use an acute coronary syndrome population and try to infer from that what the results would have been in a PCI or non-PCI population. These are different settings, and if you want to know efficacy in PCI you randomize at time zero PCI. If you want to know efficacy in acute coronary syndrome you randomize at that time zero. And, as Marv has observed and I concur with him, the results seem to be different across those settings.

The third issue, and I have taken probably more time so I am going to defer this one -- the third issue is the issue that has come up about persistence of benefit. I am completely supportive of the concept that if an intervention's effect is thought to be short term, let's say 2-3 days, it is very tempting to see the magnitude of the benefit at 2-3 days and look to see whether that magnitude is sustained out to a longer period of time. In fact, I would strongly concur that such an assessment is critical.

If death is the endpoint, for example, yes, I agree we are all going to die and so we know the difference is going to disappear but that is irrelevant. The issue is the difference sustained for a long enough period of time that we consider it clinically relevant. A difference in death at 2 days that is gone at 7 days, I would argue, is not clinically relevant. If it is gone at 30 days, I would argue it is not clinically relevant. But if it is out there to 2 years and is gone at 5, it probably is clinically relevant. So, it is important to look at whether or not the differences completely disappear so that you get a sense of what the overall magnitude of the effect is.

The problem that is going to come up, and this part I will defer because I think it will come up later in the day, is that it is easy to conceptually describe the motivation and to describe what we are going to do as we look to see if there is a significant and meaningful difference at 2 days, and look to see whether the magnitude of that difference is sustained at 30 days or 6 months. The problem is when you start then to operationalize this with statistical methods the properties of those statistical methods are, in my view, undesirable, and we need to come back to that.

DR. PACKER: Ray?

DR. LIPICKY: I don't want to prolong this very

much but the issue of how long does the effect last has been sort of drifting the way it usually is, and the idea is you give drug for 12 hours and, therefore, you are supposed to live a long time. I think that that really needs to be reexamined. If you believe what is going on here, something acute happens and something acute is modified. If you believe anything that is pertinent to extrapolations from plaque burden, there are 50 years of plaque that are there. You know, the circumstance is dealing with what happened in that millisecond a little bit ago.

So, in all probability by dealing with that, a few milliseconds or a few hours, one can influence short-term outcome. It isn't clear to me, if one thinks that way, that one has to mandate that there also be a long-term outcome. One certainly has to mandate that it not reverse or go in the opposite direction shortly. That would be adverse. But it isn't clear to me, and I just want to be sure that this gets discussed, that you have to have the effect size persist for 30 days or 6 months or 2 years, or whatever it would be, if the biological model that is going on is that you have an acute event and this is something that you are dealing with.

DR. BORER: I think that Ray has just sort of put his finger on what is a sort of a subtext here, and we really didn't discuss earlier what Tom just got into. The

issue is not so much whether the curves converge but the slope of that convergence, because ultimately one has to make a determination that if you have prevented something bad, did you prevent it long enough to be clinically meaningful -- the term that Tom used? And, we really haven't that but I just put a marker in there, we have to discuss it if we are going to talk about the way these drugs have to be compared, particularly if we are going to talk about indirect comparisons and putative placebos. It is true, Ray, it doesn't have to last forever but it has to last long enough for currently available methods to make that intervention clinically meaningful, and I think we have to keep that in mind.

DR. PACKER: Dave, thank you. We are going to go on to presentations by the individual sponsors. In doing so, let me first of all make a few general statements because it is really important to try to get from A towards B today, and I guess it needs to be emphasized that none of the sponsors should view themselves as being at risk. All of the drugs are approved for the indications for which they are approved and nothing that we say or do today will change what they are approved for. Although we are certain that each of the sponsors has an enormous pride in its database and views its database with some degree of affection compared with other databases --

[Laughter]

-- that is really not the point of today's discussion. We also are aware of what the trials are and what the overall results have been since this committee has considered many of them in the past. So, what we are really interested in, and we hope to achieve, is the individual sponsors' view on the questions before us. It would be important to hear what you can say that will contribute to the overall objectives of today's meeting. To tell you the truth, we would really prefer if you focused on those points as opposed to the relative benefits of your database compared to any other databases. With that in mind, we will ask Michael Kitt to lead off the discussion. Let me just say that the briefest presentations will tend to be the most effective today.

DR. LIPICKY: But you have just told them all of the things that I said the opposite about, Milton.

Trial Results: Eptifibatide

[Slide]

DR. KITT: With the instructions that Ray gave us, which are pretty close to what Milton said except for timing, I have been asked to discuss some of the design issues involved in conducting controlled studies of IV inhibitors of IIb/IIIa in the treatment of patients with acute coronary syndromes.

[Slide]

I will be presenting the basic hypothesis of how these drugs work, and then why the PURSUIT design was chosen, and how well we achieved the results of the PURSUIT design. I will then be using data from the PURSUIT study to illustrate some design issues in active controlled studies. Finally, I will be looking at the previous studies in acute coronary syndromes in order to further understand how design can affect the control event rate, an issue that was sort of touched on in the previous discussion.

[Slide]

It is important to start with the basic hypothesis regarding the effect of these agents, namely, that the drugs are only effective while being administered at therapeutic concentrations. For the sake of this presentation, this refers to levels of receptor occupancy in the range of 80 percent, an important issue when comparing clinical studies. When administered in therapeutic concentrations, these work by preventing thrombus formation in patients undergoing PCI, and preventing thrombus proliferation and accelerating thrombus resolution in patients with preexisting thrombus.

[Slide]

As simple as this hypothesis is in regard to the drug's mechanism of action, the treatment of acute coronary syndromes is remarkably complex. In contrast to ST-segment

elevation MI with its remarkable homogeneity and diagnosis and treatment of patients, these same issues of diagnosis and treatment are extremely heterogeneous, as we have been discussing. Variation in the extent and duration of thrombus characterizes the ACS population. There are some characteristics which are markers for the presence of thrombus, such as CKMBs and troponins, as was previously mentioned, and there are other baseline characteristics and comorbidity which are markers for outcome but, importantly, again as we have been discussing, it is different patient management strategies which can also affect the event rate.

[Slide]

The result of this heterogeneity and, in particular, differences in patient management, will have a marked effect on the control group event rate and effective of the drug treatment. I will be describing how clinical study designs, which deals with selecting patient population and how the patient population is treated, will affect each of these key features -- in other words, control group event rate and effective of drug therapy.

[Slide]

I will now use examples to illustrate some of the points I just made. We believe that at the time of starting the PURSUIT time there was good evidence for a treatment effect in patients undergoing PCI. Thus, the goal was not to

do another PCI study but, rather, to study acute coronary syndrome patients at the time of presentation of symptoms, the so-called time zero that Tom has just pointed at.

The basic premise in the design of the PURSUIT study was to demonstrate the efficacy of a potent platelet inhibitor in patients with acute coronary syndromes. Specifically, we were looking to gain insight into the treatment effect in a heterogeneous population with different management strategies, different regional cultural differences in medical practice. We believed that the large simple trial model would be the best tool to achieve this goal.

[Slide]

Thus, in order to accomplish this goal it was important to enroll a broad patient population with a broad representation of centers, global representation, real-life inclusion and just limited exclusion criteria into the study; to have no treatment mandate outside of randomization specifically in regard to catheterization and revascularization, and almost equally importantly, in regard to the timing of the revascularization.

[Slide]

The study alternatives that were available to us were either the traditional approval-directed trial or the large simple trial, and this is a diagram that I have

adapted from a publication that Rod Taylor and Eric Topol have done.

In the traditional approval-directed study one would choose a fairly homogeneous population with more restricted inclusion and exclusion criteria. In order to optimize the possibility for drug effect, as we have been discussing, one would choose to have a predefined treatment strategy which, in this case, would have patients undergoing PCI while on drug therapy.

Importantly, the results of a traditional approval-directed study lead to one having to extrapolate the results to clinical practice. In contrast, in a large simple study design we would enroll a heterogeneous population with broad inclusion and exclusion criteria, not dictating clinical care. We would in this case not necessarily optimize the timing of treatment to the disease process. But, importantly, the large simple trial leads to results that can be extrapolated to most clinical settings.

FDA has asked the committee to consider the results of the completed studies in acute coronary syndromes, and in order to assist the committee in this process I would like to briefly describe the basic design and results of the PURSUIT study.

[Slide]

Patients with chest pain or ECG changes were

randomized either to eptifibatide or placebo. Treatment was to be continued for a maximum of 72 hours but there was no minimum duration specified. All other decisions were left to the discretion of the treating physician, including when a procedures would be done and the timing of those procedures. The primary endpoint in the study was death and myocardial infarction at 30 days.

[Slide]

This slide describes the patient management -- how patient management actually occurred during the PURSUIT study. I have chosen the first 72 hours of the study because, as I mentioned, that is the time when patients were administered drug therapy. In the PURSUIT study a total of 64 percent of patients never went to the cath lab and a total of 87 percent of patients were managed medically. Only 13 percent of the patient population actually underwent PCI while drug was being administered. As you will see later in my presentation, an approximately equal number of patients underwent PCI when they were not on drug therapy, in other words, after drug therapy was discontinued. In other words, the PURSUIT design achieved its goal of studying patients with acute coronary syndromes who were managed according to normal clinical practice.

[Slide]

This slide presents the incidence of death and MI.

Just a brief reminder of the primary results -- as can be seen, the treatment with Integrilin significantly reduced the incidence of death and myocardial infarction while drug therapy was being administered, from 7.6 to 5.9 percent at 72 hours, and this treatment benefit was maintained at the 30 days, the primary endpoint, 15.7 and 14.2 percent, despite the fact that there was more than a doubling of events from the first 72 hours and that these events occurred while patients were not being treated.

The questions to the committee refer to patients with ACS who undergo PCI, among others. A lot of my disclaimer for what I am about to present actually was given by Tom Fleming a few minutes ago. So, I will at least make my disclaimer brief. However, I do want to mention that patients did not undergo PCI in PURSUIT by design. The investigators were told to practice medicine normally. Consequently, patients may have had PCI because of their disease state or PCI may have been due to the effect of the randomized treatment, or the PCI may have been planned all along. There is no way to get an unbiased estimate of the treatment effect in this population.

[Slide]

So, having given you these caveats, I am going to present a fairly complicated slide which, hopefully, you will be able to follow with me. Just to divide this slide up

very simply, on the left-hand side of the slide are all the placebo-treated patients; on the right-hand side all the eptifibatide-treated patients. On the left-hand side of the placebo group are patients who underwent PCI within the first 72 hours and the patients who were treated medically in the first 72 hours and, likewise, on the eptifibatide side.

What you can see in this slide is the actual number of events that occurred during the procedures. So, 1250 patients underwent PCI within 72 hours, 631 in the placebo group and 619 in the eptifibatide-treated patients. Important to note, there were 35 events even before the procedure was performed in the placebo group, 11 in the eptifibatide group. After the event, 71 events compared to 62 events.

Looking at patients who underwent medical management, there were 4108 in the placebo group, 4103 in the eptifibatide group. There was a reduction in events, 268 in the placebo group and 223 in the eptifibatide group, which occurred during drug treatment. What is very interesting and very important to note is that, in fact, many more events occurred after 72 hours, in this case 371 in the placebo group, more than occurred during the first 72 hours, and 376 in the eptifibatide group. As one would expect, when drug therapy was not being given there was

really no reason to believe why the effect should be any different.

[Slide]

I am now going to go ahead and, in spite of my caveats, present some of these results graphically. These are the results of the patients who underwent PCI within the first 72 hours. As you can see, there was a 5 percent reduction, 16.8 percent to 11.8 percent, again with all the caveats that I just mentioned.

[Slide]

I also want to point out the medical management. This was discussed earlier, whether the drugs have an effect independent of PCI. As tough as this particular piece of information is to draw out, I have pulled out the information based on the schematic that I showed a few minutes ago to point out here that patients who were only managed medically, who did not have a procedure in the first 72 hours, there was a 1.1 percent absolute reduction, and patients who were destined to go to PCI but had an event prior to that PCI -- in other words, they were managed medically -- there was a 3.7 percent absolute reduction in the event rate.

[Slide]

Having shown you the complexity of the PURSUIT study regarding how and when the drug effect occurs when no

management strategy is prespecified, let me show you some insights that we have for the design of active controlled studies in the future.

I am going to describe five examples of factors that need to be accounted for in the design of these studies. These are, again among others, regional differences in the population, differences in regional practice, implications of revascularization on outcome, the appropriate timing event point and the effect of new therapies on the control event rate. Most of these have already been brought up in general. I would like to bring up some very specific examples.

[Slide]

I won't spend a lot of time on this, other than discussing some very interesting findings from the PURSUIT study regarding some baseline characteristics. Not surprisingly, more obese patients in North America compared to the rest of the world, as there are more blacks; more procedures and more prior aspirin prior to coming into the study in North American compared to the rest of the world. Interestingly, there was more hypertension claimed in the North American population but on admission to the study slightly more patients were actually hypertensive.

[Slide]

Looking at probably a more dramatic difference in

medical practice, it is the prevalence of intervention on a regional basis. This slide describes the percentage of patients who underwent PCI and whether the procedure occurred during drug therapy or not.

First, the striped bars are the 72 hours and the solid bars are at 30 days. The first point to note is that at any time point there were more procedures done in North America, as we have already discussed this morning, compared to anywhere else in the world. But there is another even more important point to draw from this slide, 70 percent of the procedures performed in North America were performed early, while on study drug, whereas if you take Western Europe as an example, only 30 percent, 70.2 of the 24,8 percent of interventions that were performed were performed on drug, and roughly 50 percent of all procedures performed in the study were performed after drug was discontinued. Again, if you believe that the drugs work in this setting, this factor is very important in looking at control event rates.

[Slide]

Patient management strategy, namely whether a patient has a procedure or not, can greatly alter outcomes in studies of patients with acute coronary syndromes. I would like to contrast several key features of studies of patients undergoing PCI compared to trials of patients with

acute coronary syndromes.

In studies of PCI, the thrombus that is to be treated is induced by deployment of the device at the time of the procedure in contrast, with ACS patients who present with their plaque which has ruptured spontaneously. Likewise, the timing of plaque rupture is at the time of randomization in PCI studies, whereas in acute coronary syndrome patients it is almost always prior to randomization. The timing of treatment in regard to the plaque rupture in studies in PCI is simultaneous with plaque rupture and, obviously, this optimizes the opportunity for efficacy, and in acute coronary syndrome studies this is not at all timed with plaque rupture. In addition, as you have already seen, patients with acute coronary syndromes continue to have events over time.

[Slide]

As an example of a study in patients undergoing PCI, I would like to briefly describe some of the results of the IMPACT II study. As the committee recalls, this was a study conducted in patients undergoing PCI and doses of Integrilin that achieved approximately 50 percent receptor occupancy. To illustrate the point that I just made with regard to PCI, I would like to describe when the events actually occurred.

This slide describes the probability of an event

following PCI at any given time point of randomization. As opposed to the slide that Rob Califf showed earlier today where this time course actually was out to 30 days, this is a time course out to 24 hours, and whereas in Rob's earlier slide you were looking at most events happening in the first 3-5 days, this is from IMPACT II in 1994 when stents were barely used, and you can see the probability of having an event after about 9 hours is very, very remote.

[Slide]

Contrast this with some of the data from the PURSUIT study. Where time zero was device deployment in the IMPACT II study, the PURSUIT study shows a very different type of event pattern. This slide describes the timing events in the placebo patients of the PURSUIT study. The upper curve describes patients who were managed with early PCI compared to the patients who were managed with medical management early, in the first 72 hours.

What is apparent from this slide, as we have been discussing most of the morning, is that patients who undergo PCI early have their events early, with few new events occurring after the procedure. By contrast, patients who do not undergo intervention within the first 72 hours have more than a doubling of events from the first 3 days out to the 30-day time point and, therefore, controlling for this factor in a clinical study would be critical.

[Slide]

Based on the hypothesis of how these drugs exert their effect, it would make sense that treatment effect would occur while patients were on therapy. In addition, events that occur after drug therapy is discontinued should occur in roughly similar proportions.

[Slide]

Again, these are the results of the IMPACT II study at both the early time point, 24 hours, and at 30 days. There was a reduction from 9.6 to 6.6 percent at the 24-hour time point at the dose used in this study. As expected, the effect of drug therapy can be seen at this early time point when the intervention occurred and when the drug was being administered. After the procedure there was a small accumulation of additional events in both treatment groups, with little effect on the absolute reduction.

Just for a point of information, this study, which I mentioned earlier, was done at 50 percent receptor occupancy to show these type of results, and we are currently conducting a study looking at greater than 80 percent receptor occupancy.

[Slide]

This slide shows the time course of events in the PURSUIT study. There was a significant reduction at 72 hours, from 7.6 to 5.9 percent. Had the primary endpoint

been at this particular time point, we would have a very pretty p value, but the absolute reduction in death and MI was significantly maintained, nevertheless, out to 30 days. However, as I mentioned, after discontinuation of drug therapy there was roughly an equal number of events, as you saw in that schematic, and there was a doubling of the total event rate from 72 hours out to 30 days.

[Slide]

In designing clinical studies in the future one must also consider new advances in therapy that can affect the control event rate. What is shown here is the use of intracoronary stents in some previous studies. In the IMPACT II study only 3.5 percent of patients received an intracoronary stent, and that was performed in 1994. Three years later, in the PURSUIT study, of the 1250 patients who underwent a PCI, 50 percent or patients received an intracoronary stent and, as you are all aware, the usage of stent this year is well above 60 percent.

[Slide]

As a result of this increased stent usage, we have seen a dramatic fall in some of the complications of PCI. Shown here is angiographically observed abrupt closure in the IMPACT II study. In the placebo control group there was an incidence of 5.1 percent of abrupt closure, whereas in a more recently conducted study of stents you see an event

rate of only 10 percent of what was seen in the previous IMPACT II study.

[Slide]

Finally, in an attempt to illustrate some of the factors I described on the studies already conducted, I would like to show a comparison of some of the key features in the three previous studies of ACS. I will specifically discuss diagnosis, patient management strategies, endpoint definition and definition of new myocardial infarction. Again, I think this will help in some of the understanding of the meta-analyses that have already been presented.

[Slide]

Looking at differences in inclusion criteria, both PURSUIT and PRISM-PLUS enrolled patients both with the old definition of unstable angina non-Q-wave MI, non-ST-segment elevation syndromes. The CAPTURE study, on the other hand, only enrolled patients that were refractory to standard therapy and had been confirmed eligible for PCI by coronary angiography. Other than that, ECG criteria were remarkably consistent for the three studies, with only the PURSUIT and PRISM-PLUS study allowing CKMB positivity for entry criteria into the study.

[Slide]

I would like to contrast the relatively similar entry criteria with the marked differences in patient

management strategies among the studies. Looking at angiography was not required in the PURSUIT study, strongly suggested in PRISM-PLUS and was a requirement in CAPTURE. Waiting period, there was none specified in PURSUIT, 48-96 hours in PRISM-PLUS and in CAPTURE the patient needn't be refractory to standard therapy. PCI was not required in the PRISM-PLUS and PURSUIT studies, was required in CAPTURE. There were differences in heparin. Heparin was not specified as required in PURSUIT although 90 percent of patients did receive heparin, and heparin was used in the other two studies. Importantly, there was no minimum infusion duration in the PURSUIT study, 48 hours was specified in PRISM-PLUS and 24 hours in the CAPTURE study.

[Slide]

Looking at some important differences in the primary endpoint, death and MI were prespecified in PURSUIT, death, MI and refractory ischemia in PRISM-PLUS, death and MI and urgent revascularization in CAPTURE, which is certainly more similar to the PCI study. Timing event points, 30 days, 7 days and 30 days. Probably something that is extremely important to point out, and I think has been mentioned previously, is the role of the endpoint committees. In PURSUIT all suspected events were reviewed, where essentially in the PRISM-PLUS study only investigator identified events were reviewed.

[Slide]

Then, really getting down to some critical issues here in regard to differences in definition of MI, specifically looking at post randomization definitions in patients who were basically managed medically, in the PURSUIT study any elevation of the CKMB or, if there was no CKMB, CK was considered an MI. The PRISM-PLUS study was really looking at 2-fold elevation of the CKMB. In the post-PCI population it was pretty uniform, 3-fold elevation of the CKMB, although in PRISM-PLUS and PURSUIT it was a single elevation whereas in CAPTURE it required two samples, and some differences in the post-CABG definitions.

But what is really important to note is what effect these definitions had on the control group event rate, 13.5 percent in the PURSUIT study, 9.2 percent in the PRISM-PLUS study, 8.2 percent in the CAPTURE study.

[Slide]

Just one last slide looking at what the effect in the PURSUIT study would have been if we had used somewhat standard criteria, a CKMB elevation of 2-fold and it being at least 3 percent of the total CK, and this is death and MI not just MI, it brings the total placebo event rate down from 15.7 to 10.3 percent; the drug-treated group 8.8 percent, but what is interesting to note is the same 1.5 percent which shows that really the drug is working in the

larger MIs and the effect of drug therapy in the smaller CKMB elevations was not apparent, as shown here, without much difference.

[Slide]

I would like to conclude in pointing out that study design obviously has a very important impact on clinical outcomes. We have been speaking all morning about the heterogeneity of this population, but also major difference in management strategy among the studies really needs to be taken into consideration. The whole question of real world versus a standard study, approval-directed study, not trying to make the case of one better than the other but the fact that they are different all need to be taken into account because these all affect the control event rates and the magnitude of drug effect. Thank you.

DR. PACKER: Thank you very much. We will see if the committee has any specific questions. I think you have outlined many of the issues not only of heterogeneity amongst trials but heterogeneity amongst the point estimates which pertain to the meta-analyses that we heard about earlier. Jeff?

DR. BORER: I have two questions just to raise for discussion because I really don't think there is an answer, but in your slide on entry criteria you noted that unstable angina and non-Q-wave MI was the constant among the various

studies that you compared. But my reading of the studies themselves suggested that there is a marked heterogeneity in the definition of unstable angina and non-Q-wave MI among the various studies that you put up there. The PURSUIT definition was relatively tight, whereas, for example, some of the criteria in the tirofiban studies -- there were multiple criteria and some of them were very different. In some of the studies one of the criteria was the angiographic appearance of lesions as opposed to the clinical syndrome.

So, the first concern I have to raise is that presumably we are talking about one clinical entity or, more importantly, one pathophysiological entity and I am not sure that we really are and, you know, I would like some comment on that.

Another issue, just to get them all out and then you can deal with whichever you like, you suggested, and we heard before something that I would find fairly compelling, that is, that the absolute risk reduction was maintained over time even though the relative risk reduction wasn't as new events occurred due to other pathophysiologic processes that were sort of randomly distributed. The numbers you showed for overall event rate were small numbers so that differences among small numbers, which are even smaller, really are not appropriate bases to draw conclusions.

But, as I looked at the data from 24 hours to 30

days for the endpoints you showed, actually remarkably the difference between the placebo and the control group did decrease, and almost always by about 15-20 percent. Now, I don't know if that is meaningful, and I wouldn't suggest it is meaningful but that gets back to the point that Tom mentioned and that I raised earlier about the rate at which these rates are approaching one another, the slope of the curves of the convergence rate. So, I raise those issues for you to comment on.

DR. KITT: Let me answer the first one, the entry criteria, and I am sure Rick Sax can comment better on the different entry criteria in the tirofiban studies, but I think you are referring to the PRISM study as opposed to the PRISM-PLUS study. There were certainly some differences in the entry criteria between PRISM-PLUS and PURSUIT, principally the size or the magnitude of the CD changes and also the duration -- how long you were allowed to have chest pain before coming into the study. Those were probably the two major ones.

The PRISM study, on the other hand, which I did not address in this presentation and I think Rick will address, really was looking strictly at medical management over 48 hours. I didn't touch on that one. In that study, as I recall, patients were allowed to have angiographic evidence of coronary disease without ECG evidence of

coronary disease.

DR. BORER: You are right, and that is great, but my point is are we really looking at a common pathophysiologic process. You know, it has to be at least reasonably similar pathophysiologic process for us to accept all these results as being lumpable. It may well be. I just raise the question because it seems to me that the clinical syndromes are different enough so that maybe we are not really talking about the same processes here.

DR. KITT: I don't think I can disagree with this, and a lot of it has to do with the type of patients enrolled into the study. Even if you had actually very similar entry criteria into the two studies and, nevertheless, had enrolled patients in different places, let alone in this country -- for example, hospitals that are quick to go to angiography as opposed to hospitals that don't but, you know, specifically looking at the PURSUIT study we had patients enrolled with different types of disease, so to speak, in places like Eastern Europe and Latin America, even though they met similar enrollment criteria you see differences like the incidence of prior history of congestive heart failure and smoking histories and incidence of hypertension. So one could argue that even with sort of a crude measurement or crude entry criteria of what the disease is, there is still some dramatic heterogeneity in

that patient population.

Getting to the second question, which is how much of an effect acutely does one need to preserve at some later time point, I think is a very good question and I believe Doug Throckmorton is going to be addressing that later on. I mean, I don't know whether there is a good answer statistically or whether there is going to be a reasonable response to that but, again, since a large number of events occur over time whereas the absolute reductions remain similar but not identical, the question is how much difference in that absolute reduction is considered maintaining that benefit.

DR. BORER: Again, my question is are they really remaining similar? You know, we are talking about small numbers here --

DR. KITT: Right.

DR. BORER: -- small numbers of events, low event rates and what I may be asking may be totally beyond the capacity for resolution with the data that we have, but as I look at your data from 24 hours to 30 days for the several endpoints that you showed, you know, it seemed if I did quick calculations, which I did in front of me, the event rates converged; the absolute delta changed. It dropped by about 15-20 percent.

DR. KITT: You are probably referring to the

IMPACT II study. Again, the issues that we are really discussing here have to do with one disease state versus the other, but in the PURSUIT study the absolute reductions are very, very similar from the time of discontinuation of drug therapy until 30 days. It almost stays at 1.5 percent the entire time period. I agree with you, with the IMPACT II study, maybe because of the dosage which is only 50 percent of receptor occupancy, that might have changed.

DR. PACKER: Jeff, let me take one of the points that you just asked, and ask Michael and David to address. Since we are trying to discuss whether active-controlled trials are feasible and how they should be designed and interpreted, it would appear from the data that you described, and specifically some of the discussions that occurred when the PURSUIT study was discussed on the committee, that if one wanted to maximize one's chances of success, and maximize the magnitude of the benefit, one would, in studying patients with acute coronary syndrome study patients that were largely in the United States as they have the greatest incidence of percutaneous coronary intervention, define the endpoint early as opposed to late, that is, define it at 24, 96 hours as opposed to at 30 days, and all of these would be ways of maximizing the treatment effect. If that would be in the sponsor's interest to do this, does a historical group exist to allow a comparison?

In other words, just suppose the sponsor decided they wanted to do this but instead of going against placebo they wanted to go against active therapy, do the data exist to allow an estimate of the treatment effect for purposes of comparison in an active-controlled trial? Because, clearly, the sponsors will want to maximize their chances of success and the ability to show that they are equivalent or better than placebo would be in a population study that had the greatest chance of greatest magnitude effect, studied at a time point that would show the greatest magnitude effect. Maybe, Dave, you can begin with this.

DR. KONG: I think one of the challenges is not even like the magnitude of the effect but, rather, how confident we are about the magnitude of the effect and oftentimes you will see, just because the overall number of events may be good, the confidence intervals at early time points tend to be wider than if you measure the events later on.

So, I think there are several concerns that you have when you start talking about planning active controlled studies. One is which agent do you use as an active control? Now, we have been fortunate in the aspirin experience. We have a large number of trials using exactly the same agent, aspirin, versus placebo upon which we can base some approximate estimate of what aspirin's effect is against

placebo.

Unfortunately, it seems that in many other compound classes we have a great deal of difficulty having that same monotonous evidence base. So, one of the challenges that you have when you try to plan an active controlled study is, well, which agent, if you do choose to pick a single agent, is going to be your active control and does that skew your estimate one way or another?

One potential method would be to say, well, we would broaden the inclusion criteria for what you might use as an active control, although usually that is not necessarily good from the standpoint of being interpretable in the regulatory arena.

Another would be to say, well, we will pick some active control for which we have known experience with placebo. What that does, of course, is start limiting your previous experience. That is, you would have to ignore all the other work that has been done with agents that you haven't decided to be your active control, and limit your decision as to how good your active control is versus placebo only to those patients where it has been directly compared.

I think that is one of the complexities of the evidence based for glycoprotein IIb/IIIa inhibitors, and that we have certainly got enough evidence for each of these

compounds to be comfortable that each individual compound is reasonably effective against placebo. But one entertaining question is, well, is there enough evidence to use any one of these compounds as an active control?

Now, there are certain situations, in fact, where there are some people in the world who feel that it is even difficult to use aspirin as an active control for that reason, and that there are patient populations, say, if you want very specific patient populations, say, patients exclusively with acute coronary syndromes without ST-elevation for which, if you narrowed the definition for aspirin that closely, there are some differences in the trial results, namely AMIS, which is the largest trial to look at these types of patients, goes in the wrong direction.

So, it is concerning that even with the best evidence that we have in a particular drug class, say aspirin, we have sometimes difficulty selecting aspirin as an acceptable active control in certain situations. And, as a result, with the IIb/IIIa we also have a situation where there may be some uncertainty as to which agent we might select.

Now, experimentally and statistically there may be some way of creating adjustments for which active control you get and comparing that to placebo on a broad scale, but

that is probably not yet firm enough to be used in actual control designs without further exploration statistically.

DR. PACKER: Michael, do you want to comment?

DR. KITT: Yes. The question I believe you are asking has to do with how basically the time zero point interferes with some of the interpretation of these results. There are two members of the PURSUIT steering committee sitting here, so I could take advantage of that opportunity to ask them to comment on the design and the advantages of the design.

But an important point for us to make, and I don't think we were all that successful when we came here to present our data the first time, is that PURSUIT really was not designed to do the things that are trying to be pulled out of this discussion, and a very similar design to what these other studies were.

I go back to Tom Fleming's point or the whole issue of time zero. This study was designed to look at patients from the time they were presenting and, by definition, because of the way the study was designed, we were not, for better or for worse, maximizing the opportunity for drug efficacy. Therefore, you have all these other events occurring later.

Now, one can get into a discussion as to whether in the future this type of patient population should be used

for designing clinical trials for drug approval as opposed to the more directed type of design, but I can tell you that the information that one gains out of this type study is very applicable to the patient population that we will treat.

DR. PACKER: It is very applicable, and I think that that is one of the strengths of the study design, but it does allow one to, rightfully or wrongfully, take apart the data and allow a company that wants to do an active-controlled trial to try to maximize their opportunities. The question is do the data exist that allows them to do that in a way that can be interpreted. That is the question. Tom and Ray?

DR. FLEMING: Could I have you, Michael, put yours back up, and could I borrow your pointer?

DR. KITT: Yes.

DR. FLEMING: While this is going up, let me just reiterate what has been said a few moments ago. The PURSUIT trial was ideally designed to address the question of relative efficacy in the setting of acute coronary syndrome at time zero, in a setting in which about 12-15 percent of patients will subsequently undergo PCI in the first 72 hours. It provides us a directly interpretable answer in that manner.

The issue though is can we glean from this

insights into what the efficacy would be if we initiated the time of PCI or in a medical management setting. Do we have that slide?

[Slide]

There it is. The events here are broken down into those in whom there was PCI within 72 hours or not, and the events that occurred before PCI or after PCI. As we look at overall efficacy relative to the PCI analysis -- go ahead one slide, Michael.

[Slide]

-- this analysis is addressing the issue is there efficacy in PCI? It shows a fairly impressive 16.8 versus 11.8. I will remind you though that this 16.8 is made up of two groups and the 11.8 is made up of two groups. Let's go back to that previous slide.

[Slide]

What are those two groups? The 16.8 is made up by adding the 35 and 71 and the 11.8 is made up by adding the 11 and 62. The major difference here is the events that occur before PCI, the 35 versus 11, rather than the events that occur after PCI, which is 71 versus 61. So, if we are trying to glean from these data whether or not the intervention is effective initiated at PCI, most of this difference are the events that occur even before PCI is initiated.

What is interesting is we get to use those nice events not only for the PCI analysis, we get to use them for the medical management analysis because they occurred before PCI. So, the medical management analysis is going to pool the 35 and the 11 with the 268 and the 223.

[Slide]

There are those same 35 versus 11, and that is where most of the signal is. The bottom line here is that I am not trying to conclude from this that the intervention is not effective when initiated at PCI or is not effective in medical management, and maybe there are some clues here to suggest that it might be but most of these signals here are in the people who had the events before the PCI, who ultimately had PCI in 72 hours. The study is not designed to address initiation at PCI. The study is designed wonderfully to address a very important question, time zero is acute coronary syndrome in a population for which about 10, to 12, 15 percent will undergo PCI in the next 72 hours. At that time zero does initiation of this intervention provide benefit?

Now, if you can tell me who at time zero are the people that are going to undergo PCI in 72 hours -- it is all the males, and all the females aren't -- now I have an interpretable subgroup. But I don't know who those people are in this time zero. It is what Yusef and colleagues would

call an improper subgroup.

The only interpretable analysis in this study is the one the study was elegantly design to address, which is relative efficacy in acute coronary syndrome.

DR. LIPICKY: But there, I think, studies in people with acute coronary syndrome where the basis for randomization was that they were going to have percutaneous intervention. So, there are point estimates for what the effects of percutaneous intervention randomization are in a patient population with acute coronary syndrome. So, the answer to your question, Milton, is that there is data available that will allow one to decide whether you could do a positive control if you were using randomization on the basis of predetermining that you were going to have PCI in a patient population with acute coronary syndrome, but you could not dissect the PURSUIT data and pull your magnitudes out from there.

DR. PACKER: I think the difficulty is that in an attempt to maximize their chances of success sponsors are likely to design inclusion and exclusion criteria, primary endpoints and the timing of primary endpoints in a way that at least fulfills their bias or pretest hypothesis as to where the effect is likely to occur. The question then arises are there placebo control data in that subgroup of patients that one can use as a comparator because one

couldn't use the whole meta-analysis as a comparator. One would likely use as a comparator the point estimate at the confidence --

DR. LIPICKY: Oh no, I don't think so. I mean, you make it sort of sound evil. The signal here is pretty small. That is, all told the treatment benefit is small. So, basically it requires large populations. Then, if you are going to do a positive control trial that introduces other problems. So, it may be appropriate to have your primary endpoint be the place where the signal is the largest. So, you know, it is not evil to think of doing that; it may be the right thing to do. So, that is point number one.

Two, one obviously has to be sensitive to how the inclusions and exclusions are written but, in fact, one has point estimates for something in each of these circumstances, that is, for elective percutaneous intervention, for sort of urgent percutaneous intervention, and for acute coronary syndrome. And, I guess that that is where the questions sort of get addressed, and then Rob sort of led the committee astray a little bit ago, without anybody challenging him, by saying that studies are supposed to be for the practice of medicine, not to figure out whether drugs work and that is going to influence that discussion also because the question is, is that really what you want to do if you are a drug developer? I wouldn't want

to do that if I was a drug developer.

DR. KONG: One comment would be that certainly the easiest active-controlled trial to interpret is where the new compound actually beats -- the superiority over the active control. Although if you select a very specific population that may have a large portion of treatment effect, that subgroup, based on your previous evidence, may be so small as to give you rather large confidence intervals around that point estimate of larger treatment benefit which would actually make that candidate as an active control harder to beat if you are trying to measure against placebo.

DR. SEIGEL: Milt, I think your comment several minutes ago was something to the effect that there are incentives to study in drugs where the treatment effect is expected to be larger. It is important to note in this context that, from the regulatory perspective, if the standard of approval is to show a difference that incentive exists and that, of course, is the case for placebo-controlled trials or the types of trials David just mentioned.

If we are heading an area where we will be discussing standard of approval to show that there isn't a difference between your product and an effective drug, the incentive exists to study in a population where the effect of drug has minimal effect so that even if your drug is not

effective, it will not appear different.

I guess a lot of people here were at a meeting about a month ago where I spoke about lessons from lytic therapy at some great length. To suffice to say, there, where we have many trials where by meta-analysis you get many standards of deviations of effect, you can look and see that for people treated after 12 hours, for people with inferior rather anterior MIs or people with less ST-elevation you can get good, precise statistical estimates of the impact on effect size. You could, if a regulatory agency would allow you to do that, design a trial such that you are not likely to find a difference even if you use an effective drug.

Of note from that experience -- those were all entry criteria -- of not from that experience there is, for the reasons that we have discussed, it is harder to address, you can compare U.S. to European trials where procedure rates were different but it is not a randomization factor.

Interestingly, when you get to the issues of can you pool effects from multiple different therapies to draw appropriate bounds for the use of any one of those therapies, although there was some significant consensus on that point that you could when the total database was maybe only 100,000 patients -- so, as the databases grow, as more people are questioning whether accelerated TPA, for example,

might be different from other therapies, there is more question as to whether, in fact, those pooled estimates -- in fact, a lot of people sitting in this room have expressed a strong opinion that applying the same pooled estimate to a different therapies is quite inappropriate.

DR. CALIFF: I would like to address three or four points that have been brought up. First is Jeff's point about the effects converging. In fact, at least in our analysis, that is not what the data show, and I think it is important to point that out. I think the panel has a systematic overview and if you look at each of the endpoints, death, death plus MI, or death plus MI revascularization, the absolute difference actually gets larger between the early time point and 30 days and then comes back down again at 6 months to about the same place that it was at the early time point. It is only the relative effect that diminishes. So, the absolute difference actually does not converge in the pooled data, and for the specifics of what we are discussing today I think that is an important point to make.

Secondly, Ray's point that the treatment effect is small, unfortunately, if we actually force ourselves to measure clinical outcomes, that is almost always the case. So, as we are thinking about the future, if we want to measure blood pressure differences we can see big treatment

effects. If we want to measure effects on mortality or infarction of recurrent hospitalization for heart failure, we are going to see small effects. So, we are going to have to think of methodology to deal with small differences because that is what treatments do.

Third, I do want to comment on the driving force and the point of view of the steering committee behind the trial design in PURSUIT, which I think is very important to consider and gets back to Ray's comment about my misleading the committee. I think Ray is misleading the committee, and we argue about this all the time, but we can do very targeted, small trials in situations that magnify or allow us to see that a drug has an effect. Then we can all smugly approve the drug and go home assuming that when it is used out in the American public that it is going to have some benefit for the whole population. But I would argue that we have had a lot of experiences over the last five years where that smugness has been discounted by what happens in the complex environment of the real world, where we can't control exactly the populations that get the treatment, where practitioners give multiple other therapies at the same time.

So, the driving force on the COR Therapeutics at the time PURSUIT was designed was that we have to make this decision in the emergency department before we know who is

going to get an intervention, and where we don't know what the intravascular ultrasound looks like, and where practitioners are going to treat different patients different ways, and we would like to at least be confident that if we unleash this treatment on the public that it is not going to do harm in the general context of the way it might be used in practice. So, that was the driving force.

We could have designed a study that would have maximized the potential drug benefit and then felt good about the p value, but I would argue we should feel bad about that as a committee for what we might do to the public with the decisions that we make. I know it creates a difficult time for drug developers but maybe in good conscience we should be attacking the difficult problems.

DR. LIPICKY: But I guess it obviously isn't helping here now. If, in fact, one goes with one's bias that it looks as though the intervention stuff is the place where you get the most effect, and I make the argument that that is where it looks like to me, then, in fact, if you are going to develop another drug you and want a positive control trial, I would argue that that is the population you ought to study, people who are going to have percutaneous intervention, because you will have a smaller positive control trial to do. You may even be able to study two doses -- wow, that would be a tragedy, wouldn't it! -- as opposed

to doing acute coronary syndrome, and you might learn something in addition to developing a new agent. And, the results of PURSUIT would be thrown away for purposes of doing that because it wouldn't be useful.

So, indeed, there are competing things here from the vantage point of what is it you want to do when you are developing a drug. One of them is to figure out whether the drug works in some specific circumstance, and most economically, I would say, because it is certainly not crazy to try to develop a drug in an economical fashion. The other is to figure out what doctors should do in the emergency room. Sometimes you can put the two together. I don't know whether you think that would really work with PURSUIT. It obviously got it approved; so it did. But whether that was really the best thing to do, I don't know. As a practicing physician and someone responsible for patient care, you may think so and I wouldn't disagree with you. So, I don't disagree with what you are saying. I just think there are these competing things and they shouldn't be ignored as being competing.

DR. PACKER: Ray, you make the point that it would be in a sponsor's interest to pursue an enriched population -- let me use that term, but Jay makes the point that that would be true only if you were going against placebo.

DR. LIPICKY: Well, but I think that is wrong.

DR. PACKER: If you were going against active therapy --

DR. LIPICKY: Because the stipulation is that in a positive control trial you have to be able to make the argument "I would have beat placebo had placebo been present." So, Jay's example would not allow you to do that, and that would be a lousy positive control trial.

DR. SEIGEL: The positive control trial gets to one of the questions we discussed in pooling, if the estimate of how close you have to be, based on the question Ray just asked, is based on a broad variety of populations, high effect populations and low effect populations, if you determine based on a pooled meta-analysis that you have a 4 percent total effect and you want to be within 2 percent, then your incentive is going to be to choose that low effect population where the true total effect size may only be 1 or 2 percent because it will be easy to be within 2 percent of that even if you are inactive. Of course, the appropriate regulatory approach is to limit the estimation of the effect size of the active control to studies that use the same entry criteria, the same drug, and the same concomitant therapy as the one that is being used in the active control comparison.

Remember, all of these active control comparisons are, in fact, indirect comparisons -- the same ones that Rob

a few minutes ago said are never valid. He was saying that if A is compared to placebo and B is compared to placebo you can't compare A to B. When we are talking about active controls we are saying A is compared B; B was compared to placebo and we are drawing inferences about A versus placebo. The same problems exist if the populations are different, if the concomitant therapies are different, if the active control therapy, if its dose is different, if any of those are different, since you don't have randomization to account for those, you just have to use your brain power to guess which of those things matter and which don't. That was my point before.

At least we have some data in the case of acute MI and which ones matter, and we can make some corrections. Here we have a lot of reason to speculate but not nearly so much data.

DR. LIPICKY: I don't disagree with anything you said except the first two sentences. That is, the standard for approval, I think, should not be that you have not lost more than X. That is how you started out. You have to start out saying I would have beat placebo had placebo been present. So, if you select a patient population and/or a set of controls where you couldn't make that argument, then you are 100 percent right and, obviously, the regulatory agencies have to be careful how those decisions are made.

That is what we are trying to figure out so that one doesn't get into that business because the principal thing is I would have beat placebo had placebo been present. Then you can say, okay, well, then you probably are right but you have to have an effect size that is somewhere close --

DR. SEIGEL: To put an end to this, we are saying the same thing --

DR. LIPICKY: Yes.

DR. SEIGEL: -- when I am saying you have to be within a certain amount, I am suggesting that you have to be within that amount to assure that you have activity.

DR. PACKER: Let the record show that Jay actually favors the comparison to a putative placebo but concomitantly says that sponsors don't calculate it that way and just want to show that their treatment effect is within a certain range, and that is not the way to do it. Is that right?

DR. SEIGEL: I wouldn't say it is not the way to do it. I think if you are going to be within a certain range, that range is usually calculated to ensure that you are either better than placebo or so much better than placebo, it may actually technically work better to do an indirect comparison to placebo, which is what Ray is saying. I don't disagree with that.

DR. PACKER: We will have Tom and Rob before we

break for lunch.

DR. FLEMING: Rob has made a very important argument on behalf of designing trials in ways that allow us to reliably learn the effort of an intervention in the real-world setting. The acute coronary syndrome and PCI settings are related but they are different and it is important to understand the efficacy in each of those settings.

If the question is what is the effect of intervention, in this case Integrilin in the PURSUIT trial where time zero is as defined in the acute coronary syndrome setting, the conduct of this trial, and Michael Kitt has made the point in terms of the real-world setting and large simple trial aspect, is in essence a large part of why I consider this an elegantly designed trial. It is a trial that was designed to give us a reliable answer about the impact of this intervention in the setting of acute coronary syndrome.

We need that answer as well as the answer in the PCI setting. If we only had the answer in the PCI setting and we see, as we have seen from David's meta-analysis, about a 35 percent reduction, the reality is, David's meta-analysis is telling us, in the acute coronary syndrome it is about a 12-15 percent reduction. If we only had the 35 percent reduction and now we enter into an active-controlled trial of a few intervention against one of the previous

IIb/IIIa's, and we are using the 35 percent reduction estimate but we are actually doing this trial in the acute coronary syndrome, we are going to be greatly misled. So, it was extremely important that studies were done in both settings. As Rob pointed out, it is true that it is going to be harder to show that you have a significant effect in the acute coronary syndrome setting, but it is a different setting that requires an accurate understanding of what the relative efficacy is, and for all the arguments that Rob made, studies like the PURSUIT study are properly designed to address that setting, and are necessary if we intend in the future to do an active controlled study against Integrilin, for example, in acute coronary syndrome.

DR. PACKER: Rob, last word?

DR. CALIFF: I thought Ray had a great comeback there, actually. I was impressed, and I agree with his contention that if you were going to start in this setting with an active control comparison that percutaneous intervention is where you got the clearest effect and would be the right place to start. But what I would worry about would be saying, okay, that is enough. Do that with your new drug that has some kind of antiplatelet activity and then just give it to the whole world in whatever setting you want to. That would be very pertinent to the PCI setting.

But beyond that, if we want to know if it should

be used in ACS, my concern is that it probably really needs to be studied there because in the real world patients with ACS are getting exposed to not just percutaneous intervention but a variety of other drug therapies, and we need to have some understanding of how it all fits together to recommend it. So, I am afraid we don't get out of the box, we don't get out of the argument by saying, you know, here is one setting and we can do an active-controlled trial and it really gets us through all the issues that we need to get through.

But the final thought I have, which is probably the scariest one, is that there is a tendency to want to cling to this standard of showing you are better than placebo. I understand the reason why we want to think that. That is what the law says. But I have to wonder, going into the future, whether it is a standard that we are sort of hiding behind, super-mathematical sort of mumbo-jumbo, to hide the fact that in essence we are doing what Dr. Seigel said, which is that we are using a historical control that we can't be sure about. So, no matter what kind of confidence intervals or p values or whatever you end up with, you still have a matter of judgment as to whether your so-called putative placebo event rate is really what you think it is.

DR. PACKER: Rob, that is truly in many ways

throwing down the gauntlet because I think Ray would probably remind us that the law doesn't say that it has to be against placebo. The law says that the recommendations of the committee or decisions of the agency need to be based on trials which are convincing to experts.

DR. LIPICKY: It really doesn't matter what the law says. Have you ever heard me cite the law?

[Laughter]

DR. PACKER: Anyway, I think that the arguments can be made both ways. I think the greatest degree of comfort and confidence has accompanies the concept of comparison against a putative placebo. I think one widens the degree of uncertainty considerably if one pursues alternative models, but that is for discussion after lunch.

DR. CALIFF: I just want to say that the comfort level that you feel about a putative placebo may simply be an illusion.

DR. PACKER: It is something to think about over lunch. We will reconvene at 1:30.

[Whereupon, the proceedings were recessed at 1:30 p.m., to be resumed at 1:48 p.m.]

AFTERNOON PROCEEDINGS

DR. PACKER: We will resume with discussions for this afternoon, and ask Rick Sax, from Merck, to talk about the principles raised with trials of tirofiban.

Trial Results: Tirofiban

DR. SAX: Thank you.

[Slide]

I am going to spend the next few minutes talking about the results from the tirofiban trials, in particular the PRISM-PLUS trial and the RESTORE trial, mostly to use these results to illustrate a number of points looking forward towards the active-controlled trials and, in particular, to talk a little bit about the inclusion criteria that were used in PRISM-PLUS, the trial design and some aspects that we have touched on already in the morning session of medical management and percutaneous coronary intervention, issues related to composite endpoint, some issues that have come up a number of times already today on durability, and talk just a little bit about subgroup and cohort analyses.

Overall, I think that these issues have to be factored into the consideration for active-controlled trials and may raise the question as to if one can show, for example, that one is not inferior to a certain time point and certain endpoint what other factors might need to be

considered as one looks at non-inferiority. I also want to talk about the PCI trial, the RESTORE trial, in particular on some issues of selection of endpoints and how this may relate to meta-analyses, and come back at the end to the issue of durability.

[Slide]

Just to remind you, the PRISM-PLUS program consisted of three trials, the PRISM trial which focused on the period of medical stabilization; the RESTORE trial in which the drug was initiated and the catheterization laboratory in the setting of angioplasty; and the PRISM-PLUS trial which was an ACS trial, focusing on all aspects of the management of patients from medical stabilization through angiography and through angioplasty.

Since I don't think that active-controlled trials are likely to go against tirofiban on a background of heparin, I am not going to discuss the PRISM trial here. I will be glad to answer any questions about it. Nor am I going to discuss the dropped arm which was a tirofiban alone arm in PRISM-PLUS, but just focus on PRISM-PLUS as the ACS trial and RESTORE as the PCI trial.

[Slide]

So, let me begin by addressing one of the questions that the committee has to face this afternoon, namely, is there a population that can be identified that

would serve as a standard for future active-controlled trials? These were the inclusion criteria for the PRISM-PLUS trial. As Dr. Kitt has already indicated, they were very similar to those for the PURSUIT trial, the only major difference being the duration of therapy. For PRISM-PLUS they had to be randomization within 12 hours. In the PURSUIT it was 24 hours.

But, precisely, patients had to have symptoms, anginal symptoms and, in addition, had to have objective evidence of electrocardiographic changes, ST depression, transient ST-elevation, less than 20 minutes, or deep T-wave inversions, or had to have enzymatic evidence suggestive of an infarction, namely, elevated creatinine kinase or elevated CK. I think nowadays we probably would include troponins as a marker but at the time troponins were not widely available. I think these are objective findings that can be recognized in patients that would serve as a basis for defining inclusion criteria that could be used in subsequent active-controlled trials.

[Slide]

The study design had something very particular in mind when we set this up, and this was done with certain clinical and regulatory considerations and, as has been indicated, did differ from the PURSUIT trial. The patients presenting with acute coronary syndrome at time zero, as Dr.

Fleming has talked about, were randomized but then there was a period of treatment where patients were not to have procedures, a medical management period that lasted 48 hours unless the patient developed an endpoint, in which case they could proceed to procedures.

So, here we have a defined medical management period, and this was put in here specifically to address the question as to whether this drug was active, independent of the setting of angioplasty, because we do recognize that the treatment patterns do tend to favor patients going on to angiography and angioplasty in particular in North American, as has been pointed out. We did allow patients then to continue on therapy through angiography and most patients did undergo angiography. It was not mandated but they did undergo angiography and angioplasty if the physician felt that was clinically warranted. This was up to 108 hours.

Our focus, however, was on the overall management of patients with acute coronary syndromes and, therefore, we chose an endpoint at 7 days, thinking that that was reflective of the drug effect. We wanted to look at the overall drug effect and look at the effect for the in-hospitalization period. We were also cognizant of the fact that we should show durability looking at 30 days and 180 days, and I will come back and talk a little bit about the implications of that.

[Slide]

Let me talk for a minute about the primary endpoint. In the PRISM-PLUS trial the primary endpoint at 7 days was a composite of something we called refractory ischemic conditions, new myocardial infarction, and we used all-cause mortality. Again, it was at 7 days in patients with acute coronary syndrome, non-ST-segment elevation, syndrome of unstable angina, non-Q-wave infarction.

The refractory ischemic condition endpoint really was designed to represent a failure of medical management, and the patients continued to be objectively symptomatic with objective evidence of electrocardiographic changes -- and this is what makes it somewhat soft -- through optimal medical therapy, namely beta blockers, nitrates and maybe calcium channel blockers titrated to heart rate and blood pressure, and they were to continue to be symptomatic.

There was considerable debate, and continues to be considerable debate -- I know this is one of the questions the panel has to address this afternoon -- as to whether this mandated proceeding to a procedure, and that gets to the question as to whether procedures are good or bad or whether they are part of a practice pattern, and we decided that because of issues related to practice patterns around the world that we did not want to mandate use of a procedure as part of this condition but really focus on the symptoms

and evidence of ischemia. New myocardial infarction in the study was defined as a 2-fold elevation of CK or an elevation of CKMB, but there was no preordained screening of CKs in this trial. This was based on clinical symptoms and/or electrocardiographic changes. So, it was really driven by symptoms. Death, as I mentioned, was all-cause mortality. So, those are the components of the endpoints that were involved in the PRISM-PLUS trial.

[Slide]

The result has been discussed here before. There was a reduction of the primary endpoint from 74.9 percent to 12.9 percent, a 5 percent absolute reduction. The confidence bounds have been described. I will show them on the next slide. So, this could serve as a basis if one wanted to try to do this. There is a good treatment effect. The control group rate is I think understood. It could serve as a basis for an active-controlled trial.

[Slide]

However, it is important to look not just at the composite endpoint in the treatment effect. This was a one-third reduction in the overall event rate. But when you come to some understanding whenever one has composite endpoints one needs to look at what drives the composite endpoints. Here, as has been talked about before, there is really no effect on overall mortality, recognizing that the numbers

are very small here, but this is consistent across all of the trials. We do not see an effect directly on mortality here, and it is really driven by both refractory ischemia and myocardial infarction. In this case, both were of good magnitude and both were significant even though the trial was not powered to pick up any differences between the two, and the composite endpoint, of course, reflects primarily the fact that this is driven by myocardial infarction.

But as one thinks about active-controlled trials, one needs to not focus just on the composite endpoint but, again, one can envision trials and, in fact, we know of at least one large trial where the composite endpoint was positive by pooling of variety of arms but, in fact, mortality went in the wrong direction. So, if one is, in fact, going to go after a composite endpoint one would clearly need to at least look at the components of that to make sure that they are going in the correct direction, and we think this is guidance that the agency has certainly given, but then it becomes an issue of, well, is this all driven by refractory ischemia or refractory that leads to procedures, in which case one gets into issues of procedures, or myocardial infarction and then one gets into the definitions of myocardial infarction. So, it is very important to define those quite accurately as one looks at active controls.

[Slide]

I want to turn now to the question that has been discussed in the morning about medical management because this trial, again, was not powered to look at the medical management period but we were very interested to know whether there was an effect in medical management prior to any procedures and, in fact, again, the trial was not powered to look at 48 hours; it was a secondary endpoint, but you can see a magnitude of treatment effect that was not significant at 48 hours but certainly was in the same ballpark as the overall effect at 7 days.

Again, here prior to procedures there was a reduction in clinical myocardial infarctions and this was not driven, as I said, by procedures. Procedures were quite rare during this time period. So, there was an important reduction in myocardial infarctions and deaths were too few to really count them in any meaningful way, and refractory ischemia went in the right direction but the primary driver here was a reduction in clinical myocardial infarctions.

So, I think that one can begin to address the question as to whether these drugs really have benefit in the medical managed population independent of angioplasty, getting to some of the issues we talked about earlier.

[Slide]

Let me turn now to the question of durability.

These curves I think are illustrative of the same point Dr. Kitt has already raised, but the primary effect of the drug is quite early on. This also came up in Dr. Califf's talk. But after the initial time period for the study one sees an accrual of events that essentially is in parallel between the two treatment groups.

This is for the composite endpoint, which in this figure also includes readmissions for myocardial infarctions and readmissions for unstable angina. What you can see is that readmissions plus the initial events accrue at the same rate out to 180 days. Again, the study is not designed or powered to do this, but it turns out that at 180 days we did achieve a level of statistical significance, but for all intents and purposes, the treatment effect that was seen at 7 days and certainly at 30 days is really just maintained. The absolute delta is really not much different between these two time points. That is true also of death and myocardial infarction. Again, if you look at the delta at 7 days, which was an absolute delta of 3.4 percent, 3.2 percent at 30 days, 3 percent at 180 days, essentially all one is doing here is preserving the initial treatment effect, and I think this is just a demographic illustration of the point that has been raised throughout the morning.

[Slide]

I want to venture into these very treacherous

waters that Tom Fleming has talked about already and talk about some very improper subgroup analyses -- I think Ray is probably going to cringe here but the FDA asked us to do it, because I do think it is important as one thinks about unstable angina claims of non-inferiority to recognize that one could design a trial where the outcome is really based on outcomes related to angioplasty, a trial such as Dr. Packer has described already.

This was an analysis that -- again, I will admit that these are improper subgroups; they are post-randomization subgroups, very confounded by the fact that they are after time zero and they are predicated to some extent on the fact that they may have occurred during medical management, but they can be used not for inference but for an understanding of what drives the outcomes in the trial.

Again, I will not focus on the statistical aspects of these because they are improper subgroups, but it is important to note that as one looks at the all-patient cohort in the trial that there was a good effect in patients whether they subsequently through randomization underwent angioplasty, underwent bypass surgery or were in medical management. As Dr. Fleming has already pointed out, patients undergoing angioplasty or bypass actually get counted for their medical management period as well, but the main thing

is that the overall results are consistent across these various post-randomization subgroups, indicating that there was not one subgroup that completely drove the effect, and since the management of acute coronary syndromes really represents a heterogeneity of practice patterns, including medical management, bypass surgery or percutaneous coronary intervention, I think that as one thinks about non-inferiority of an endpoint at some time point after these events occur one needs to have some understanding as to what is driving the endpoint.

[Slide]

Just to point out again that there is a period of time, as Dr. Fleming pointed out with the PURSUIT data that happens before angioplasty, it also turns out here that there was a benefit after angioplasty. As one might expect for an agent that has complete inhibition, there is injury at the time of presentation and then re-injury at the time of angioplasty. However, again as has been discussed, this is not to say that one should ever make inferences from this subgroup that this should lead to an angioplasty claim. I think that that would require another trial. So, in fact, that is what was done and let me turn to that.

[Slide]

The trial that we used to study angioplasty was the RESTORE trial, and in this trial tirofiban was

randomized versus placebo on a background of standard angioplasty heparin, at least standard for the time, and there was a very low incidence of stent use in this trial. Time zero was at the time of angioplasty, and there were prespecified analyses in this trial at day 2 and day 7, again, related to the belief that the drug effect was going to occur early.

However, as has been the case with all the angioplasty trials in the field, the primary endpoint was specified to be at 30 days and a secondary endpoint at 180 days. I just wanted to touch on a couple of issues here, again, as they relate to active-controlled trials and meta-analyses, related to the selection of endpoints and selection of timing, which are also part of the questions.

In this trial, the composite endpoint -- and I talked a little bit about the risks of composite endpoints, but here the composite endpoint was all repeat revascularization due to ischemia, stent placement for abrupt closure. So, here stent placement was actually an outcome rather than part of the procedure, and that is something that may be very different than future trials going forward from here. New myocardial infarction -- I am going to spend a little bit of time on this because it is quite important and relates to something that Dr. Thadani has been emphasizing over a number of years now. And, again,

we can all agree on all-cause mortality, I believe.

The primary endpoint here was at 30 days, but I do want to point out that unlike any of the other trials in this field, here the population studied was a population of acute infarction, that is, ST-elevation or non-ST-elevation infarction or patients with unstable angina. So, this, in fact, was the sickest going into the trial of any of the populations that were studied in the setting of angioplasty.

[Slide]

Let me turn to the issue of endpoints because the selection of endpoints does make a big difference, both in the interpretation of meta-analyses in the field and also in a looking forward way at active-controlled trials. Again, this is an endpoint that I think everyone can agree on. But there has been a lot of discussion about the definition of myocardial infarction and the meaning of the levels of CK elevation. I am not going to get into that discussion, except to point out that in RESTORE, like PRISM-PLUS, the definition of infarction was driven by clinical symptoms. The protocol did not specify routine screening of CK during the course of the post-angioplasty period, and this has major consequences if one uses the concept of clinical infarction versus enzymatic-based infarction. One can get under-reporting of events by investigators looking for clinical symptoms, or one can get under-reporting of events,

as you will see, if one doesn't draw CKs. That is something that needs to be considered and very carefully specified in an analysis of a trial, again, on a looking forward basis.

The other major difference in RESTORE with regard to the endpoints was that here the steering committee, and there was considerable debate about this, thought that what was relevant for the patient was all angioplasties for symptoms, not just urgent revascularization. There was considerable concern among the steering committee members about how to interpret what did or did not constitute an urgent revascularization. So we, instead, went ahead and favored counting all revascularizations.

I just want to point out the consequences of that approach in the urgent acute setting, and what implications that would have for an active-controlled trial. The other trials in the field have all used the emergent urgent definition. Here, stent for procedure failure was something that was important at the time, but also would have to be reevaluated because of the high use of stents up front and not waiting for procedure failures. So, that is an issue that really needs to be addressed for any active comparator.

[Slide]

This just gets to the issue of CK analysis and what happens if you don't screen for CKs and rely on clinical symptoms. You can see that actually in the PCI

setting, as one looks at death and myocardial infarction, there is a very good balance irrespective of the entry criteria -- I shouldn't say balance but very good concordance, almost irrespective of the entry criteria if one uses death plus MI driven by serial CK screening. The event rates, irrespective of the trial, irrespective of the type of patient population actually are fairly close to each other, except for RESTORE which did not screen serially. That is true whether one looks at 30 days or 6 months. Again, I talked about death and MI here but this is almost entirely driven by myocardial infarction.

So, I think as one thinks about the PCI trials one can probably say that PCI may not make a difference for your population entering the trial, but it does make a difference how one collects CKs, and if one does not do this in a systematic way one is going to have trouble with the active comparison.

[Slide]

Again, these curves are probably similar to other curves that you may or may not see, but the effect of these drugs is quite early and then persists to some extent. In RESTORE, however, there was a narrowing of the curves between 7 and 15 days which, as I will show you, was related again to the endpoint definition. The early benefit was seen right after the procedure -- that has been described a

number of times -- and clearly persisted to 7 days but was lost at 30 days, in large part, due to the endpoint definition.

[Slide]

And, that relates to the fact that when you count all procedures rather than urgent revascularizations. If we had counted urgent revascularizations we may have seen a curve like this where essentially all the urgent revascularizations occurred in great proximity to the drug effect so, almost by definition, one has urgent revascularizations during initial hospitalization and then the curves remain essentially flat. There are almost no urgent revascularizations that take place after he patient's discharge. This, in essence, forces the endpoint to the time where the drug is having its greatest effect and, therefore, when one talks about durability one is actually talking about durability in terms of an event that happens here, not something that is occurring out here.

If one looks at the other criteria of non-urgent revascularizations, those begin to accrue after hospital discharge, say, at about 7 days and then continue to accrue over the time course and then continue to accrue over the time course and, in fact, as I will show, they accrue out to 6 months.

[Slide]

Here, the trend went in the wrong direction for tirofiban compared to the placebo group.

[Slide]

But lest one think that this is unique to tirofiban, the same effect was seen with the abciximab trials, the EPIC trial and the EPILOG trial where the non-urgent angioplasties, the non-urgent bypass surgery within 30 days for this particular trial went in the wrong direction. The 30-day data, to my knowledge, haven't been published, but the same thing is for true at the 6-month period with non-urgent revascularization going in the wrong direction. So, again, from a looking forward point of view, when one thinks about counting urgent revascularizations, then one is really looking at the drug effect quite early on. If one counts all procedures, then this leads to a potentially different conclusion.

[Slide]

Now, all procedures have a common endpoint at 6 months, and when one counts all procedures at 6 months, again, one sees the same pattern that one sees with all of these agents over and over again, that the effect is right here, right when the drug is available and at that point events accrue in parallel going out to the 6-month period. In fact, the delta at 7 days is virtually the same as it is at 180 days, this being again an illustration of the point

that has been already made.

[Slide]

Just in conclusion, I think that the lesson we can learn relative to active-controlled trials is that the choice of endpoints is critical. One needs to understand the meaning of the endpoints as they relate to the drug effect and as they relate to the issues of durability. I think in this field with the IIb/IIIa's we need to clearly identify that the drug effect is early so over time it may be acceptable, but I think, as has been iterated by the committee a number of times, durability of effect may be important.

But it is probably not sufficient to look at just the primary endpoint or specified endpoint. One really needs to focus on not just non-inferiority or comparability, equivalency -- whatever term one wants to use, but really look for consistency and interpreting the results across the components of the composite, across subgroups, and I think, in response to Tom Fleming's comments, it is important to begin to look at the issues of practice patterns and whether a trial is driven by medical management, if one is going to focus on acute coronary syndromes, and what component medical management has versus percutaneous coronary intervention. So, thank you.

DR. PACKER: Questions?

DR. THADANI: Although you said that the patient population was similar, I think there are several differences. Patients might have been sicker. In the old days we used to call it, you know, acute coronary syndrome with pre-infarction.

The other problem, one of the issues is although you said angiography was not mandated, if my recollection is correct, almost 90 percent of the patients had angiographs. If 90 is not mandated, I don't know what mandated is. Almost every patient got angiography. So, if you are going to apply the practice guidelines based on PRISM, which I assure you nobody in my situation is doing, they are not getting the drug for 48 hours and then doing an angiogram. They give the drug and the next thing you know, the patient is in the cath lab. So, it is not applicable results to clinical practice at the present time.

The dose used in RESTORE was different than in PRISM-PLUS. So, that is another issue which comes into correlating the two because the dose was higher in the RESTORE, if I remember correctly, and in PRISM-PLUS it was different. Those are issues that I could argue with you --

DR. PACKER: But let me interrupt you. We have a real mission today, and the mission is not to get the sponsors to defend their data.

DR. THADANI: Certainly not. The question is

because can we combine your results with the PURSUIT or the totality of the data? That is why I am addressing these issues -- they are still there. Ray has pointed out that if the drug effect is so good we should see it early and, yet, in PRISM-PLUS, although it was a second endpoint, we do not see much at 48 hours. You know, you show a great effect at 7 days which is maintained but the composite endpoint at 48 hours is negative. So, that is again another area of are the drugs different, or do the procedures make a difference? I want your comment on that.

DR. SAX: I think the points you raised are correct about the specifics of the trial, and I think it just goes to illustrate that one has to look very carefully at what is actually driving the results in the individual trials, and that just makes it a little bit more complicated as one thinks about non-inferiority or comparability for composite endpoints.

DR. PACKER: Again, whether a certain trial has a certain p value, a certain time point, etc., etc. is not the point of today's discussion.

DR. DIMARCO: I would like to ask a question. You chose to go by clinical infarcts, the other studies have gone by enzymatic infarcts. It seems to me your approach has the disadvantage that you might miss things that may have prognostic significance but at least you have the same

definition all the way through, whereas those people, if they had enzymatic bumps, you know, the day after the patient left the hospital, they would have been missed if it was silent.

DR. SAX: Dr. Thadani has made the comment about that, that essentially infarctions are confined to the period of screening --

DR. DIMARCO: Why did you choose yours, and what do you think the advantages are?

DR. SAX: At the time, in RESTORE the steering committee was very concerned about the meaningfulness, as has been addressed here, of small CK elevations. The data, at the time the trial was initiated, were not available. I think that the data now are considerably more compelling to suggest that serial CK screening and, in fact, even low levels, as Dr. Califf has already mentioned this morning -- even low levels of CK elevation probably have some prognostic implications. But at the time that this trial was conducted that wasn't understood, and so the steering committee clearly favored going with a more clinically based definition. In unstable angina generally one follows a rule-out MI protocol and that is what drove the screening of CKs in the unstable angina setting.

DR. PACKER: Rick, does that suggest that if you were going to do it again you would have used the same

definition or a different definition? If you were going to do it in the year 2000?

DR. SAX: If we were to do a trial in the year 2000, looking at non-inferiority in the setting of angioplasty I think we would look at the amalgamated data that has been acquired and follow the trials that have already been done, and do serial CK screening. I think the data are now there to support doing that.

DR. DIMARCO: How would you look at the CKs after the procedure? Would you keep the people in the hospital for the next 24 hours or monitor them for an extra 24 hours?

DR. SAX: No, we would monitor 24-48 hours or until hospital discharge. Generally, these patients if they have uncomplicated procedures are discharged within the first 24-36 hours.

DR. PACKER: Following up on that question, how would you know the effect had persisted if you didn't screen CKs after discharge?

DR. SAX: That is the question Dr. Thadani has asked every advisory committee, that essentially your definition of infarction, unless it is clinically based and you get readmitted, your infarcts are basically limited to the time of the drug effect. It is the same thing as urgent revascularization. So, what one is seeing at 30 days in angioplasty trials, by the definition of the endpoints based

on CK screening and based on urgent revascularization is, in fact, confined to around the time of the drug effect.

DR. THADANI: Milton, on that, there is a more complicated issue. The patient comes at point zero; your enzyme is negative; he is asymptomatic at 6 hours; his enzymes go up. Now, since you have already randomized the patient you count it as an ongoing infarct which you did not pick up because enzymes don't increase until 6 hours, or the drug could have induced infarcts. In the trials, most of the trials are calling those as an infarct before actually it is not a reinfarction. It could be a silent reinfarction, but if you have a clinical endpoint of chest pain and then a bump there is much more reinfarction. So, there are always problems with even looking at infarcts with CPKs at any time point. So, when you decide you have a patient with a non-Q-MI in your trial, if his enzymes at point zero which is 12 hours -- say he comes at 12 hours, he has non-Q-wave. Other guys in the trial -- because he has come within 2 hours and his unstable angina with 6-hour enzyme is positive, he could have been a non-Q to start with. So, I think there is a major problem to analyze that data, although you hope that in a large sample size they are equally randomized but there are some difficulties.

DR. PACKER: Can we have some more discussion about this because if we are going to discuss the conduct

and design of positive controlled trials in this area, we can't do that without at least recognizing what an endpoint for a drug effect should be. What should an endpoint for a drug effect be? Should it be death? Death and MI? If it is just death and MI, realizing that MI is going to be much more common than death, is it enzymatically screened MI or a clinical MI? Is it death, MI and revascularization? And, if it is revascularization is it all revascularization or urgent revascularization, or is it death, MI and refractory therapy which may or may not lead to revascularization? Unless we feel comfortable in known what endpoint we should be looking at, almost all other questions in this area would be very difficult to pursue. Jeff, any thoughts?

DR. BORER: Well, sure since you ask. I don't think that anyone would argue about death, and probably nobody would argue about infarction although the point has been raised, that we have to define infarction. Personally, if I had to make a first cut here without any data in front of me to determine the implications of that, other than what Rick showed and what is in the book here, I would say enzyme screening is the appropriate way to go. The endpoint is more specific and probably has more important implications than merely the symptoms that go with it.

I think, and this gets back to a comment Rob made before, or Cindy made, the issue is not whether you do an

angioplasty or not but whether the conditions exist that require that something else be done because the patient can't be managed. So, that would lead me to suggest some definition of refractory ischemic symptoms with which the patient cannot go on unless you add something more. Maybe it is angioplasty; maybe it is bypass grafting; maybe it is something else. But I think it is the clinical condition that mandates the additional therapy that should be the endpoint.

DR. PACKER: If you add the third component, Jeff, recognizing that the third component was not necessarily collected in most of the trials, one would be facing an enormous challenge in terms of using the existing database to adequately design positive controlled trials.

DR. BORER: Well, a problem but one possible approach is to go back and look at the data and see if that issue can be captured. I think that ignoring that third, much softer endpoint is done at our peril because the fact that it occurs confounds what one might have anticipated with regard to the harder endpoints. I think that this would be an issue of informative censoring, just as Tom was mentioning earlier. So, I don't think we can ignore the other endpoint. The question is do we have the data and can we define it in a way that is useful?

DR. THADANI: I personally think that death is the

definite endpoint. Death never lies. And, in trials the definition of infarct, even by enzymes, is different in different trials. Some take twice the normal. Now, given the data on troponin INT, a lot of patients are getting infarcts. I have had patients with CPK of 50, MB of 4.8, troponin of 1.6. So, patients got infarcts and none of those infarcts were counted in the previous trials. So, that moving definition makes it much harder. Even if you define the previous data, it would be different numbers.

Regarding the need for revascularization, I think it varies from fellow to physician. I got a patient on nitro 240; Mike's patient is on 20 and the patient has 3 or 5 minutes of chest pain and the next thing I know he is in the cath lab. So, unless you can rigorously define -- and some of the protocols have defined it, you know, irrespective if the patient is on IV nitro. I think in PRISM-PLUS the definition was 20 minutes of anginal episodes and in practice nobody does that. So, I think it is a much softer endpoint. I don't mind this being a secondary endpoint, but in the composite it is difficult unless you really define it and everybody follows it. So, I think those are the issues one has to address.

DR. PACKER: There are lots of comments here, and it is a very important discussion because, to tell you the truth, if we don't get past this almost nothing else

matters. So, Udho, you are suggesting death and MI --

DR. THADANI: No, I am suggesting death on its own; maybe MI as a secondary because I am not convinced --

DR. PACKER: You are not suggesting mortality alone?

DR. THADANI: I am suggesting mortality alone because the small MIs driven by enzymes don't translate into late mortality.

DR. PACKER: You can't be suggesting mortality alone!

DR. THADANI: I am.

[Laughter]

Why can't I?

DR. PACKER: You can't suggest mortality alone because, one, there is no effect on mortality.

DR. THADANI: That is why I am suggesting large trials to show effect.

[Laughter]

Why not? I mean, you did trials in heart failure because you were not happy with hospitalizations for heart failure. You ended up doing trials to show a mortality difference. All the thrombolytic trials are based on mortality. Now we are going to lower the standards and use therapeutic agents -- we have not talked about negative risk/benefit ratios here. All we have heard in the two

trials is they talk about death and MI. Even in the first trial we did not talk about bleeding risk, you know, the grand complications, whatever else. So, I am saying what I am saying --

DR. PACKER: But how can one do putative placebo modeling for an endpoint for which there is no treatment effect?

DR. THADANI: Well, if you can't do it keep on doing placebo-controlled trials.

DR. CALIFF: But there is another issue with regard to mortality as the sole endpoint here, you would need trials of 100,000 --

DR. THADANI: So, your drug is not as good as you think.

[Laughter]

DR. PACKER: The trial doesn't do that. It doesn't reduce mortality. That doesn't mean it isn't valuable.

DR. THADANI: I realize that, but I think that is important. I could argue with you that I could tell a patient I am going to do angioplasty, I am going to give an infarct. Rather than waiting for an infarct on day 30, I am going to give it to you on day 1 by enzymes. Are you going to be happy with it? If the PCI study is the correct one, and as Rick has shown, if you go by symptoms limited your infarct rate is 6.1; in all the other trials it is about 9

or 10. Do you tell the patient I am going to give you angioplasty and I am going to give you enzyme rise on day 1 by 4 percent? I bet no patient would take it if you told him that.

DR. PACKER: We will do it this way, Marv; after Marv I am going to turn to Ray. Jeff, hold on; I promise I will get back to you. Marv?

DR. KONSTAM: Well, I disagree with Udho because I think MIs are important and because I think we are not going to see another study if we hold it to the standard of mortality alone because of the event rate problem.

I actually want to make just two sets of comments. One relates to this idea of acute urgent revascularization as an endpoint. I will say that for studies with acute coronary syndromes in this country in the year 2000, you know I have a major problem with that as an endpoint. I think there probably is a country somewhere where coronary intervention is not commonly performed for patients with acute coronary syndromes, but it isn't the United States. Furthermore, as has been alluded to before, we don't have a scenario where people are -- and I think Cindy said it well earlier, or someone said it -- we are not keeping people in a unit for a few days and seeing how they respond to medicines before we are taking them to the cath lab. It is just a fact of life. We can argue whether it is right or

wrong but it is just a fact of life. I think that in that environment I am troubled by acute urgent intervention as an endpoint in the acute coronary syndrome patients.

I would have no problem with it in an angioplasty trial, which actually, as others have pointed out, is really a stent trial these days. In a percutaneous coronary intervention trial, need for repeat urgent revascularization takes on a whole new meaning and there I would have no problem with it.

So, although I sympathize with the notion that there is meaning to patients having recurrence of their syndromes, I think for practical purposes I have a problem with using acute urgent revascularization as the indicator of recurrences as part of the endpoint.

So, to me, MI and death is really where you are on solid ground for the acute coronary syndrome studies. I don't know what the answer is with regard to clinical MIs versus enzyme screening. I defer to Rob and others in terms of the data indicating that small CPK rises -- or maybe we should be looking at small rises in troponin as having prognostic significance. So, I sympathize with that view and I think that there is something more objective about it than just relying on clinical symptomatology as part of the MI syndrome.

But I am very troubled by the Integrilin data

showing the difference in the MI endpoint between the investigator-defined endpoint and the endpoint committee-defined endpoint. Maybe we might have some discussion about that. Let me say it this provocative way, if you believe the drug worked, then the investigators did a better job of defining endpoints that showed the drug worked than the endpoint committee using enzymatic screening. And, I am very troubled by that finding. It has still not been explained to my satisfaction. It suggests to me that there is something happening in the environment of the intensive care unit with an investigator in front of the patient that is defining an event that really is meaningful. I will just leave it with that.

DR. PACKER: Let me ask Cindy, Cindy you referred to this in a previous comment today. Can you give us your views as to what would constitute an appropriate endpoint?

DR. GRINES: Well, I agree with Marv. I would probably eliminate refractory ischemia and urgent TVI. I think that is very subjective. I don't think that it is well controlled as to the degree of antianginal medications that patients are on to make that determination. The definition of urgent intervention could be just putting the patient on the table at the end of the day. It doesn't mean that you rushed him down and burst through the doors to throw the patient on the cath lab table. Furthermore, it is going to

be eliminated by stenting, which is happening in about 80 percent of all the interventions now.

So, I think that we should eliminate that as an endpoint. Clearly, death has to remain an endpoint. With regard to infarction, I think we all agree infarction is important but what we disagree on is how it is defined. I think that there are a lot of questions still as to the frequency of the screening, as to whether an MB elevation alone in the absence of a CK is an appropriate endpoint. Although it does have prognostic significance, it may in fact just be a marker of diffuse soft atheroma, similar to what Steve Nissen was discussing this morning. I think even data from Duke shows that the clinician's definition of an infarct does have more prognostic importance than the clinical events committee. I want to opt for allowing the clinician to define the infarct.

Now, I am not totally opposed to drawing enzymes, but what would be nice is if those enzymes were, in fact, blinded to the investigator and the investigator had to make an independent determination as to whether they thought the patient was infarcting or draw their own enzymes, as opposed to these protocol-determined enzymes. And, that happens in many trials. You send out your troponins so the operator doesn't really know. If they want to draw enzymes they order them themselves.

DR. PACKER: I am going to ask Paul Armstrong to think about this because I want to get your views on this. Ray?

DR. LIPICKY: I will be changing the subject --

DR. PACKER: Oh, no, we don't want to do that. Not yet.

DR. THADANI: On Cindy's question, can I ask Cindy the significance of enzyme rise post-PCI because I was at a meeting and they said the enzyme rises really don't mean that much as opposed to the clinical coronary syndrome, acute ACS, because there are differences because your lumen size is bigger when you are blowing a balloon up. Am I right in my thinking or is it a new trend which is developing, which might have important consequences for what the enzyme rise means?

DR. GRINES: I think that there is a lot of data that if you do have an enzyme elevation it is going to ultimately affect your prognosis. But enzyme elevation in the absence of angiographic complication or any clinical scenario -- it doesn't cause the patient to die and, you know, we talked a lot about -- at least not die early. They may die three years down the road but that is from a rupture of a second plaque presumably. If we all believe that the treatment effect occurs during the duration of treatment, then you have to question whether these MB leaks -- truly MB

leaks, not CK elevations -- whether that is clinically relevant. The only way you can answer that question is if you do have a mortality difference.

DR. THADANI: So, you are agreeing with me that mortality is important.

DR. GRINES: Well, obviously mortality is important.

DR. THADANI: Because why we started treating infarction was to prevent death. Originally why did the whole thrombolytic trial start was to prevent death. And, I think the same thing, my comments were applicable to ACS because when the Duke group showed the data that ACS is as bad as acute MI because you showed 12 percent, 14 percent, yet, surprisingly, mortality is not showing a difference.

DR. PACKER: We will have Paul and Rob.

DR. ARMSTRONG: I have two comments, Milt. Symptoms in this disease are obviously incomplete and often misleading feature of --

DR. PACKER: Closer to the mike.

DR. ARMSTRONG: Symptoms are often are late, incomplete and misleading feature of the disease, and contemporary medicine suggests that physician interaction with those symptoms makes it even more difficult.

Second, I think that the frequency with which we sample and the things that we measure will impact on the

definition of infarction, and it is a continuum without an arbitrary cut. All of the data that I am aware of suggest that any movement above a previously established baseline has prognostic implications.

The other issue, as has come up several times this morning, is the good old electrocardiogram is a good arbitrator at the time of entry into patient trials, and also helpful in terms of arbitrating the presence or absence of symptoms and their meaning. There are biologic ways, including more continuous measurements of those that have direct prognostic significance, and track treatment that has been established as effective.

So, I would throw out those issues as I think germane to the discussion at hand.

DR. PACKER: Rob and then Tom?

DR. CALIFF: I would like to make a couple of comments about the endpoint issue. First, I think no one is going to argue against death being included, but I think death alone is just out of the question because even a dramatically effective treatment -- I mean, there are just things that are important other than death.

For those designing trials though, it is important to understand that power based on a composite endpoint, leaving you with relatively little power for death, does leave you open to the possibility that by chance you will

end up way on the wrong side in your point estimate for death and that can lead to some unanticipated, unpleasant circumstances in explaining what you find.

My favorite endpoint is still death and MI, and I think symptomatic MI in the setting of a cath lab experience -- I think it is not really fair to compare that to symptomatic MI in a patient who comes in to the emergency department from home. I think most of us, if we had an angioplasty or a stent implantation, would rather not remember the experience in the lab, and there are a lot of other uncomfortable sensations going on. So, to say we are going to look for symptoms and then draw enzymes I think is not exactly a perfect way to do it either.

So, my favorite endpoint would be death or MI. I think what we learned from PURSUIT, that we are trying to get published if we can get the journals to understand it, is that when PURSUIT was designed there was concern that the high threshold of enzyme elevation used before was lowering the event rate and, therefore, leading to very large sample sizes. So, any elevation of MB was called an MI in PURSUIT.

What the analysis really shows comparing the CEC to the investigators is that there were a lot of examples, particularly outside the United States, where a single sample showed a minor elevation and, by definition, you had to call that an MI, from the CEC. But I think a lot of those

were laboratory artifacts and not real events. If we require 2- or 3-fold elevation of CKMB the relationship between the investigator call and the CEC becomes very concordant. So, I think a threshold greater than just any elevation would be needed. And, if there is a wide number of countries used in the study, probably a core lab, if you could do it logistically, would be ideal. That costs a fair amount of money up front, however.

In ischemia, I can certainly buy that if a patient has refractory symptoms that is a bad thing. No patient would want to have that, and it could be counted as an endpoint but, again, those designing trials need to be aware that we are not sure that your power is actually going to be greater if you include that composite on average. It may be, at least based on some of the databases that we have, that you actually reduce your power because the treatment may have less of an effect on that third component of the endpoint.

DR. KONSTAM: I understand what you are saying about the PURSUIT data but might you not draw a different conclusion about how the next trial should be done, namely, that in fact the investigator-driven endpoint did a very, very good job at distinguishing drug from placebo? And, so what does that mean? It either means it is partly by chance or the investigators were picking up something irrelevant

that the drug did, or, which I would favor, that the investigator-driven endpoint actually did better than anything else. And, why wouldn't you just conclude that you might want to head in that direction, perhaps with a check on it? That is to say, it would have to be CEC concurrence with an investigator-defined endpoint.

DR. CALIFF: It is a case in PURSUIT if you look at the endpoint where both the investigator and the CEC had to call it. That is where you see the greatest treatment effect. So, it is tempting to want to go in that direction, but that is just one trial. We don't have an empirical database to say that would happen every time. So, I am sort of swayed by the combination of the two, but there are just too many cases when you do a trial where it is an obvious MI and the investigator -- part of it I think is because study coordinators collect the data; the investigator may be looking at a case report form and not really thinking about it. There are just too many cases that are obvious where the investigator is wrong for me to be comfortable with that as the only endpoint.

DR. GRINES: But, Rob, you have monitors that go out to all these sites. So, it is not just the study coordinator or the investigator that has made that mistake; the monitor is supposed to catch that as well. So, then for the events committee to double the rate of diagnosis of MI

is pretty substantial. How do you explain the errors in monitoring then?

DR. CALIFF: We could have a long discussion about on-site monitoring and what it accomplishes and what it doesn't, particularly if you are doing large trials.

DR. PACKER: But not today.

DR. CALIFF: Right.

DR. PACKER: Let's see, Ray? Tom, I will not forget about you, I promise. Ray and Jay since it is important to get some regulatory input here. So, Ray?

DR. LIPICKY: Well, I wanted to disagree with the notion that mortality should be in the combined endpoint.

DR. PACKER: Say it again, Ray.

DR. LIPICKY: Atypically, I would like to disagree with the discussion and I don't think mortality should be part of the combined endpoint. What we are talking about is deciding what an endpoint should be for a positive control trial sort of based on what we see in the data. Just like you said, Milton, this stuff doesn't seem to affect that endpoint. So, it is sort of catchy to have mortality be in the endpoint; it is certainly not bad but I think it dilutes things. I can certainly measure mortality but I would not have it as part of the primary endpoint where, in fact, it is detracting from one's ability to detect the signal.

DR. PACKER: Ray, this committee has gone through

this a number of times. I am sure Tom will refresh our memory as to what the pros and cons of this are, but if I remember correctly, the reason for including mortality in a combined endpoint has little to do with the argument as to whether there is a treatment effect there, but has a lot to do with the argument that that is the worst possible outcome and, consequently --

DR. LIPICKY: It is a competing risk, but I think that is an erroneous thinking process.

DR. PACKER: -- and, consequently, we have operated in general with the concept that to do, for example, hospitalizations without including mortality as an endpoint can lead to a whole host of erroneous conclusions, not just based on competing risk, not based on what is a worst clinical scenario, but it would imply that fatal MIs can be excluded from an analysis that focuses only on non-fatal MIs.

DR. LIPICKY: Right, but I want to defend the position I took a little bit so I can argue this. That is, if you don't know anything at all about the treatment that you are undertaking, and that is usually the case in a placebo-controlled trial, you really do have to evaluate this new treatment for competing risk and all that sort of business. But what you see in this business, and you have a bunch of trials to look at, is that the mortality is

negligibly affected, if at all, so it, in essence, is noise. So, now it becomes a safety issue and I would evaluate that as an independent entity; certainly measure it, and I don't know that I would power the study to be able to make a distinction there but, for sure, as you said, since the studies are not powered sufficiently to detect reasonable magnitude of effect for death, if it is part of the combined endpoint it may trend in the wrong direction.

We, indeed, say everything has to trend in the right direction. So, we either have to stop saying that or we are really being schizophrenic. And, we can evaluate death and its trend, in the right direction or wrong direction, whether it is part of the primary endpoint or not. If the new trials behave like the trials we have seen, having death in that composite endpoint is just noise and I don't think it belongs.

DR. PACKER: Ray, realizing that schizophrenics tend to be very adaptable people -- Jay, hold for a second. Tom had his hand up. It was probably not to address this issue --

DR. FLEMING: Now it is.

[Laughter]

DR. PACKER: Jay, was it this issue?

DR. SEIGEL: Well, I rose to address a different issue I would still like to address, which is investigator-

determined MIs, but I would like to address this issue as well.

DR. PACKER: Hold for a second, Jay. Tom?

DR. FLEMING: On this issue, I see three reasons to include death. One is it could carry some of the signal and when I look at David's meta-analysis, in the meta-analysis there is a 25 percent reduction in the deaths at 30 days, which is consistent with the overall magnitude of reduction in death/MI. So, it could be carrying part of the signal.

Secondly, I agree with Milt's point. It is profoundly important and it is concerning to me to leave out from the primary endpoint an element that is the most clinically important.

Thirdly, from a statistical standpoint, to not count deaths does generate a very substantial informative censoring. If somebody dies without having had an MI we are going to assume -- they are not out of the analysis, they are still in, but their subsequent profile for when then have an MI is represented by other people who didn't die, and to presume that that person who died was like the people who didn't, in terms of their MI risk, is a very significant assumption that I would believe is probably easily proven to be wrong.

DR. LIPICKY: I can argue with each of those. I