be a risk factor, but glucose lowering may not be a benefit. Can you sort of rephrase that question?

MS. FLEGAL: I was just thinking of the analogy for obesity where it increases the incidence of some conditions but they also improve survival in some of those same conditions. Whether there's sort of a distinction like incidence and mortality don't have exactly the same risk factors.

Is there any suggestion of that in these data?

DR. GERSTEIN: Not that I know of. But clearly, if the glucose numbers -- I don't know of any data that would answer that question.

DR. BURMAN: If I may, I'd like to ask you a question.

DR. GERSTEIN: Sure.

DR. BURMAN: Do you think some of the differences between the ACCORD and ADVANCE related to the rapidity of the drop of the

glucose? Number one. And number two, do you have any information regarding the increased deaths in the patient from the ACCORD?

DR. GERSTEIN: The causes of the increased death? Is that what you're -- yeah, so within the ACCORD trial, we looked carefully at the reasons for death, and essentially people died for the same reasons that people die in the general population. So you know, a lot of the deaths were cardiovascular disease. And then there were a whole bunch of other miscellaneous deaths.

So in terms of the actual cause of death, nothing sort of emerged as being a particular cause.

And there was no sort of yellow flag that said, ah ha, this is what happened. And in terms of the reasons for the death, the analyses to date have not identified anything as being the particular reason why there was this mortality signal.

And so other analyses are being

done. More papers are being published. And in my mind, I think somebody else said it. It's likely that we will never find a smoking gun. It is something about the totality of the intervention that seemed to have the effect, in the same way that in the UKPDS and some of the blood pressure trials, it's hard to differentiate what is the component that was the benefit, or was it the whole strategy, per se, that did it.

In terms of the first question, Ken, was the differences between the trials? Is that --

DR. BURMAN: Yes, in terms of the rapidity with which glucose and HbA1c dropped.

DR. GERSTEIN: That is a hypothesis. So the ACCORD trial really did what we all recommend that people do. So intensively control people's A1cs in a safe way. At the time, it turned out that there was this mortality problem, but clearly, safety was paramount in the way that it was designed. So

it was designed to minimize hypoglycemia, safely lower A1c levels as rapidly as could safely be done.

And it did it within the course of -- a lot of it within the first four months. ADVANCE took three years to lower the A1c levels, and had a lesser difference between groups -- about half the difference between groups in ADVANCE compared to ACCORD. And the difference -- the contrast was maintained for a much shorter period of time.

So one could hypothesize that perhaps some of the differences between the trials relate to the speed of glucose lowering. That's clearly a hypothesis. And we'll never know at this point, but that was a difference between the studies. VA has not published their curves yet, so it's really hard to comment on that.

DR. BURMAN: Thank you. Any other questions? Dr. Parks?

DR. PARKS: Dr. Gerstein, you

mentioned that both ACCORD and ADVANCE are glucose-lowering trials, and clearly looking at differences between intensive and standard control. But the other difference here, the treatment regimens, do you want to comment on the differences in the treatment regimens to achieve the same goal?

DR. GERSTEIN: ADVANCE prespecified that they have to put people on glipizide, which is a drug similar to glipizide in terms of it's a sulphonylurea-type drug. And after that, people can add what they wanted to add to the intensive group.

ACCORD clearly regulated both -- so it was actually -- ADVANCE was (inaudible) as usual care. ACCORD was prescribed care in both groups. It was targeting the A1c with less than 6 percent with a menu of drugs, versus targeting A1c 7 to 7.9 percent with the same menu of drugs.

So it was pre-specified what the A1c was going to be in the standard group

versus letting it be what it would sort of end up being in a usual care sort of approach. And that's another difference between the designs of the two groups. But I don't think that there's anything particular special about the glipizide per se. I mean, it is a sulphonylurea and it seems -- one can always find unique properties of any one sulfonylurea. But I think it's just really part of the approach that they used. But again, one can speculate.

DR. BURMAN: Thank you. Any other questions before we move on? Thank you very much.

Let's move on to the next speaker. Dr. Steven Nissen.

Thank you.

DR. NISSEN: Thank you very much. First of all, it's really a privilege to be here. And I want to thank the Agency for the opportunity to participate.

As I think some of you know, I've

been critical of the Agency in recent years, and it takes a little bit of courage to invite a critic to come and talk about these things. And so I applaud you for giving me the opportunity to give you a perspective on this entire area, one that obviously has taken on increasing importance in the last year or so.

So hopefully we'll have a slide here. Very good.

I am really going to talk very specifically about a change in regulatory strategy.

And I'm going to outline what I think is a rational approach towards approval of these drugs. First, I wanted to show you a disclosure slide -- although you've obviously heard this already, please keep in mind, however, that companies are directed to pay any honoraria speaking or consulting fees to charities, so that I do not receive income or tax deductions from participating in

consulting or research with clinical trials.

So why are we here? The ACCORD trial demonstrated that a drug regimen designed to lower blood glucose is capable of increasing mortality in diabetic patients. That in and of itself I think is a clear signal that we've got a potential problem that we've got to find a way to address through regulatory policy.

Secondly, multiple rosiglitazone meta-analyses of CV outcome showed improved glycemic control but an increase in myocardial ischemic events. So we have two events that suggest that if you lower blood sugar, perhaps in the wrong way, you can increase morbidity and you can increase mortality.

It is also important to know that many agents to treat diabetes have failed during development, some due to cardiotoxicity. Most of which you don't know about because the studies that showed the

toxicity have never been published. And lastly, no robust cardiovascular outcomes data exist for any current diabetes therapies.

The problem is insufficient clinical trial data. In the absence of randomized CV outcomes trials, we are left with unsatisfactory methods to assess benefit and risk. These include meta-analyses or post hoc data dredging of randomized trials not designed to determine the benefits or risks of specific therapies.

And I thought this was particularly evident at the recent ADA meeting with attempts to determine the source of the excess mortality in ACCORD. And there's just no amount of torturing of the data that will enable risk assessment when specific drug usage was not randomized, as I will show you a little bit later.

What we do know is that a strategy in which the primary differences that were

observed were with these four agents. There was a lot more use of repaglinide, a lot more use of rosiglitazone, somewhat more use of insulin, and more use of an alpha-glucosidase inhibitor. Somehow or other, this regimen resulted in an increase in mortality. And I don't think that we'll ever be able to know what it was about this strategy that led to the increased mortality, no matter how far we delve into the data.

So here we are, 50 years after the initial introduction of anti-diabetic agents. And although cardiovascular disease is the cause of death in 75 percent of diabetics, there exists no well-designed, adequately powered, comparative effectiveness trials evaluating macrovascular outcomes for diabetes drugs.

Now, for those of you who may disagree, let me point out to you this very nice systematic review in the Annals of Internal Medicine, which looked at

comparative effectiveness and safety of oral medications for type 2 diabetes published several years ago -- actually last year. And this is the summary table from that analysis. This is Level of Evidence: Comparative Effectiveness Trials of Diabetes Drugs. All cause mortality, level of evidence available low to very low. CV disease mortality, low to very low. Non-fatal MI or stroke, low to very low. Peripheral vascular disease, low to very low. And microvascular outcomes, low to very low.

So we've heard a lot of data, and I think it's very compelling, about lowering blood sugar. But we have almost no information about how to lower blood sugar. And that absence of information has left us in the current dilemma that we find ourselves in. So we have a knowledge gap. And I think that the absence of information on macrovascular effects is unfortunate.

It's a consequence, however, of

current regulatory policy that emphasizes the importance of glucose lowering (inaudible) as a therapeutic goal. Now, I think we have four or five major classes. I heard there are 10 classes. I can't quite count that high, so let me just say that there are certainly many ways to lower blood sugar.

What we really need to know are what agents improve health outcomes, and what agents can we develop beyond current therapies that will improve health outcomes. We can lower blood sugar. If we have to, we can give insulin. So we have lots of ways to reduce blood sugar. What we really want to do is find the right way to lower blood sugar.

Now, it wouldn't be right if I didn't poke a little bit at my endocrinology colleagues, so please excuse this. But I have to point out that this sort of approach has led to a new disorder, which I've termed glucose-centricity. Now, if that's a new

word for some of you, I went to the Webster's Dictionary and I looked it up. And I found this definition of glucose-centricity. It's the irrational belief that lowering blood sugar using virtually any pharmacological means will produce a reliable reduction in adverse outcomes. I think what we've learned in the last year is that that's not correct. And so we've got to move beyond a glucose-centric approach.

There are major consequences of using glucose as the primary driver of drug approval. Preapproval studies focus on demonstrating maximal glucose lowering effects. Therefore, patients are selected with relatively high HbA1c levels because this enhances apparent efficacy, as I will show you.

The higher you start, the bigger the delta you see. And the purpose of many of these studies is to attain bragging rights. My drug lowers blood sugar more than

your drug does.

But that's not really what we want to know. In fact, patients at high cardiovascular risk are deliberately avoided. Sponsors say why take a chance of an adverse safety signal? So let's exclude these high-risk patients from our clinical trials. And that further compounds the problem, because it means we're not going to have enough events to actually see the signals we need.

The other thing that happens is that when safety signals arise, physicians stampede to the newest diabetes therapies for which we know the least about safety. And I found it really interesting that after the rosiglitazone concerns emerged, the fastest growing diabetes class is a new class -- the DPP-IV inhibitor, of which sitagliptin is the first agent.

However, this agent has very limited glucose-lowering efficacy, perhaps

about half the effective established therapies, and virtually no long-term safety data.

So what happens is, in the absence of definitive information, people just flip to the newest drug because that's the one that's least likely to have anybody worried about its outcome. I worry the most about new drugs, not the least.

So here's the dilemma that I'm going to try to propose an answer to. How do we balance the need to bring new diabetes agents to patients in a timely fashion? You know, the information this morning and early this afternoon is convincing. I know there's a microvascular benefit. I know lowering blood sugar is a good thing. But we also need more robust outcome data to inform physicians on how to use these drugs safely and effectively.

We have -- you have to as a panel -- I, unfortunately, don't get a chance

to vote, but you do -- to help the Agency understand how do we balance these two issues. Timely approval of new drugs versus having adequate information.

And so I will propose to you then a rational approach. And I'm going to tell you that although there are many people that would like this, requiring a large CV outcome trial prior to approval is undesirable because this approach would delay new diabetes therapies by five to seven years at the very least.

And so I'm going to propose what I think is a reasonable compromise. And it has two components: a pre-approval set of clinical trials designed to rule out a high level of CV risk; and secondly, a large randomized outcomes trial that must be underway at the time of approval.

And so this combination of an appropriate pre-approval set of trials and a large outcomes trial that's already enrolling

patients at the time of approval provides us, I believe, with what we need in the long run.

Now, what about the pre-approval development program? And again, this is a balance between what's desirable and what's feasible. I believe we need pre-approval trials of sufficient size and duration to rule out a hazard ratio of 2.0 for major adverse cardiovascular events. I will show you data for several other cut points for that upper hazard ratio, but I think that this might be a reasonable one. And 1.8 would also be reasonable, as I will show you in a subsequent table.

This would require pre-specified pooling of CV outcomes in all trials with adjudication by an independent clinical end points committee. That's not done now. What we get is kind of haphazard adverse event reporting in these trials, and that doesn't give us the clarity that we need. And so I'm suggesting that as part of the plan, we ought

to have the requirement that these events be carefully adjudicated during the pre-approval study. I think it would be useful to have at least one study in patients at high CV risk, perhaps 1,000 patients for one to two years.

Now, keep in mind that this is in the context of what I showed you earlier. That if we find the wrong way to lower blood sugar, we can harm people. And we have to rule out some level of harm prior to approval.

So this table is pivotal. And some of these data were data that Marv Konstam asked for earlier. And I sort of smiled when you asked the question because I knew I was going to show you this slide. So here's just a way of looking at this. If you have 50 events -- 50 cardiovascular events -- that can rule out an upper confidence interval of 2.5, which if you have a point estimate that's below 1.44 with 50 events, you end up with an upper confidence interval of 2.5.

And assuming either a 2 or a 3 percent event rate -- now, that requires studying higher risk patients to get those kind of event rates -- these are the numbers of patient years required. If you require 87 events, you can rule out an upper confidence interval of 2.0, which means that the point estimate needs to be below 1.31. And these are the number of patients years you've got to expose. At 122, 1.8 and 1.26. And then what I think is probably too stringent, 256 events will rule out 1.5 and a point estimate of 1.17.

But somewhere in here -- and the challenge for this Committee and for the Agency -- is where to set this point. We have got to know that we've at least ruled out some level of harm during the pre-approval testing. Later in this presentation, I'm going to apply this standard to some previous development programs, and I think you'll find the results

fairly interesting.

So here it is shown another way. If you have 87 events to exclude an upper confidence interval of 2.0 and they break evenly between the active and control group, that's the hazard ratio you get, and those are the confidence intervals. If you have 48 events in the active arm and 39 in the control arm, you've got a 1.23 hazard, and you still stay below the upper confidence interval of 2.0.

If the drug actually shows many fewer events, you get .67, and clearly you've ruled out the upper confidence interval. But if you have this excess of events, you end up with a point estimate of 1.56 and your upper bound of the confidence interval exceeds 2.0. And this agent would need more testing prior to being an approval agent. So this is one way to look at what kind of studies we might need pre-approval.

What are the positives of doing

this? Well, it encourages sponsors to include patients with higher level of cardiovascular risk. Look, everybody. These are the patients we're going to treat. I mean, I have a coronary care unit where 50 percent of the patients in the CCU have diabetes. Those are the patients that are getting these drugs.

To do development programs where you exclude all those high-risk patients because you don't want to see any signals, it's the wrong approach. And we've got to correct that now in the pre-approval process.

It provides more reliable pre-approval data by adjudicating cardiovascular events for pool trials.

It's not that hard. You know, a committee can do this with the number of events involved. Not expensive. You just need an independent group of people to look at the events and decide whether they're real or not real.

The negatives are that it would modestly slow development programs, perhaps delaying introduction of new diabetes medications by 6 to 12 months. If you really only need 87 events, you can get those number of events in a reasonable length of time with large enough trials in a high enough risk group of patients.

Step two in the approval process is an adequately powered cardiovascular outcomes trial. And so assuming no evidence for excess risk, an upper confidence interval of less than 2.0, a new diabetes drug would be approved based on glucose lowering efficacy if an adequately powered ongoing CV outcomes trial is underway already enrolling patients. I'm going to tell you why I feel this is necessary in a moment.

This outcome study should also address any other ongoing safety issues -- rental, fractures, skin toxicity, et cetera. This policy is a compromise

designed to balance speedy approval with the need to promptly obtain evidence of benefit or risk.

Now, what might such a study look like? I have deliberately not drilled down further because Rob Califf is going to talk about this. And I didn't want our two talks to overlap. But just to give you a few sample size considerations. If you're MACE rate, let's say a five-year MACE rate is anywhere from 11 to 17 percent, and you have a punitive reduction in risk of 12 to 18 percent, then these are the sample sizes per treatment group, approximately. And I'm sure Tom Fleming can get you more detailed numbers, or one of the other statisticians. But this is the ballpark.

In cardiovascular medicine, we do 10,000, 15,000, 18,000 patient trials all the time. These are not that daunting. And I believe that Dr. Califf is going to tell you how to do these on the cheap. How to make

them large, simple, and easy, and not having to cost hundreds of millions of dollars, but to be more reasonable in cost. If we could get this launched prior to drug approval, we would have a whole new era of solid data on what to do with these drugs down the road.

Why do we need these trials before -- to be started before approval? You know, it's not pleasant to look at this slide, but it is the reality. This is from the FDA report to Congress on September 30, 2005. This is the number of drugs with a Phase IV commitment and the number of commitments completed. It's about a 14 percent completion rate. If it's promised, it may not be delivered. And I'm going to show you that it often isn't delivered.

And so I believe the only way to guarantee the medical community that we're going to get the answer, but without delaying the approval of new drugs is to have such a

study underway at the time of approval.

The study must be high in quality. Now, I'm going to say I must compliment the ACCORD investigators. Because, frankly, I think they did a very good job of controlling.

88 percent of the patients in ACCORD were on statins, 76 percent on aspirin, and 71 percent were on ACE inhibitors. I cannot say the same thing for the ADVANCE trial. Why is this important?

If you want to claim an incremental benefit -- if it's on microvascular disease, for example, on nephropathy as it was claimed for the ADVANCE trial, you better have patients on those therapies that are of proven benefit. And I think it is not acceptable to have 47 percent of high-risk diabetic patients on statins in this day and age. 56 percent on aspirin.

We don't actually know from the trial report exactly how many were on ACE

inhibitors. They had similar levels of macrovascular disease.

And so part of the standard around these trials is they have to be good trials. They have to be done properly with adequate control of risk factors so that we can find out what the drug does on a background of reasonable, decent, medical therapy.

Now, those who cannot remember the past are condemned to repeat it, as George Santayana said. And so at the risk of beating a dead horse, let me go back and look at a few recent development programs and see what would have happened had these kinds of standards been in place.

Dual PPARs. Promising idea since both hyperlipidemia and insulin resistance appear to promote atherosclerosis in diabetics. Pharmaceutical companies have sought to develop dual alpha and gamma agonists. As I hope many of you know, many of these drugs have failed during

development. At least five that I'm aware of. They have fibrate-like effects, raising HDL and lowering triglycerides, and they have TZD-like effects in proving insulin sensitivity, and thereby lowering blood sugar. So it makes perfectly good sense why you would want to do this.

The first of these drugs to reach approval process was muraglitazar. September 8, 2005. Came to this Advisory Board, or a predecessor of it, for approval. This is the development program at the time. It met the current standard for what would be required of development. And here's what they had.

They had studies of 24 to 104 weeks again, mostly short-term. Several different doses of muraglitazar. Some doses were dropped during development. 23,704 patients. They compared interestingly enough to submaximal doses of pioglitazone.

Something that should bother

everybody here. If you're going to study a drug against an active comparator, you'd sort of like to study it against the optimal dose of the comparator. They studied against 15 and 30 mgs of pioglitazone, which I don't think is a very wise approach to development. But that's what they did.

And so here is the development program. And here's what happened. We saw at that advisory panel very robust reductions in HbA1c. And as I mentioned, if you study patients with really high HbA1cs, you can make a drug look really efficacious.

And so in this open label part of the studies with a HbA1c of 10.7, they got a whopping 2.62 percent reduction in HbA1c. But that wasn't all. Triglycerides reduced 27 percent. HDL cholesterol went up 16 percent. And no effect on LDL cholesterol. So a really favorable profile.

Now, there appeared in the database to be a higher incidence of major adverse

cardiovascular events. And the sponsor argued to the panel -- by the way, this is actually direct from the sponsor's slides -- that there was a lack of biological plausibility for cardiovascular risk with muraglitazar based upon the following: Beneficial effects on cardiovascular risk. I mean, how could a drug that lowers HbA1c by that much -- raises HDL 16 percent, and lowers triglycerides by 20 percent -- how could it possibly have cardiovascular harm?

There was a broad diversity among reported cardiovascular ones. It wasn't must MI, or stroke, or death. It was all those events that seemed to be in excess. And that didn't seem to make any sense, right? I mean, a drug shouldn't increase all those things. There was no dose response signal. The higher doses didn't look any worse. Obviously, the power to make that determination was very low. And that there was no cardiovascular toxicity in the

nonclinical studies. And therefore, there's no way that this drug could increase cardiovascular risk.

So the panel voted 8 to 1 to approve muraglitazar as monotherapy, and 7 to 1 to approve its use with metformin. They voted against its use in combination with sulfonylureas, because that study was one where there was somewhat more evidence for harm.

Six weeks later, we took that database from the FDA Advisory Panel meeting and we re-analyzed the data by pooling all the available data. And this is what we found. They were right. All the different components -- they were all increased. All cause mortality, the relative risk was 3.05; CV death, 4.57; non-fatal MI, 2.1; fatal or non-fatal stroke, fatal or non-fatal MI, and the hazard ratio for adjudicated congestive heart failure was 7.43.

If you then look at the

composite -- and I would point you to the one right in the middle -- all cause mortality plus non-fatal MI or stroke was 35 versus 9. Now, remember that I said that I thought you really needed 87 events. They had about half that number here. About half the number of events that would have been desirable. The relative risk was 2.23 and the p-value was .03. And so we recommended that this drug not be approved based upon these signals.

The FDA in fact agreed. Issued an approval letter requesting additional cardiovascular safety data. And after ongoing extension trials, confirmed the cardiovascular hazard. All development of the drug was halted. However, a risky agent came close to approval.

And it's really -- I think, it was a close call. A clear standard requiring an upper confidence interval of less than 2.0 would have precluded even the necessity for a cardiovascular advisory panel. When you've

got a signal of this intensity in the studies prior to approval, I don't think you want that to go forward without more safety data.

Let me give you a second example. This is again the slippery slope of surrogate endpoints and diabetes drug development. Now, as you all know, ezetimibe was approved to treat hyperlipidemia on the basis of reduction in LDL-C averaging 16 to 18 percent. So what would we do? What should we do with a diabetes drug that lowers blood sugar by 1.1 percent but increases LDL-C by 16 to 18 percent?

In other words, if a 16 to 18 percent reduction in LDL-C is sufficient to demonstrate benefit, what inference should we draw when a drug increases LDL by a comparable amount?

And this is the case of rosiglitazone. Here is from Joy Mele, the statistician, from her statistical analysis of the rosiglitazone approval package in

1999. And what you see is there's a 13 to 24 percent increase in LDL-C, but even more strikingly, when the LDL was below 130, it ranged from 23 to 32 percent. So this is statin magnitude LDL increases. This is from the Advisory Board package in 1999. So the people that had more normal LDLs are getting these very large increases in LDL-C. And Robert Misbin says patients treated manifest undesirable change in weight and lipids. And I agree with that.

Now, there's also then the cardiovascular event data from that approval package. And again, it's about the same number of events that we saw in the muraglitazar package: 36 with rosiglitazone, 10 with comparators. Here are the event rates. Here's the relative risk, 1.8. And here are the confidence intervals, from .9 to 3.6. So up to a 10 percent benefit and a 360 percent hazard.

The FDA reviewer says a

post-marketing study to evaluate long-term safety of rosiglitazone should be a cry for approval. Now, again, I would point out to you that with a standard of an upper confidence interval less than 2.0 there's no way this drug would have moved beyond this stage. It would have required additional safety data prior to approval. But what was done was they said, okay, well, let's approval it, but let's require a large outcomes trial post-approval.

And so the question is what happened to that mandated safety study? Well, it's called the adopt trial. But it wasn't a safety study. It was a marketing study designed to show greater durability of glucose lowering with rosiglitazone. Cardiovascular events were collected in haphazard fashion. They weren't even adjudicated. And because of the LDL raising effect of rosiglitazone, it turns out that more patients, p<.01 got statins in the group

that got rosiglitazones. And so they were able to neutralize some of the LDL disadvantage by giving more statins.

In spite of that, the hazard ratio for myocardial infarction is 1.33, with a 95 percent confidence interval of .8 to 2.21. This pre-approval signal never goes away.

And then a bunch of other marketing studies were done. Short-term studies to show glycemic reduction in various populations. There's no well-designed outcome trial to measure health outcomes. As Tom Fleming showed you, the RECORD trial is underpowered by a factor of about 3, even compared to the event rates that they had postulated. So it's got a 3 percent event rate when 11 percent was postulated. By 2007, 42 trials had been completed with 14,237 patients.

And by the FDA analysis in July of 2007, the odds ratio for myocardial ischemia is 1.4. With these confidence intervals of

1.1 to 1.8, the pre-approval signal never goes away.

I would submit to you that what we needed in 1999 was a standard like the one that I'm proposing. I think it would have protected us from this drug coming to market with a signal that never went away. So if an upper confidence had been required in 1999, the drug would have never been approved without more safety data. If a large, well-powered outcomes trial had been mandated in 1999, we would not have to wait until 2014, 1five years after approval, to determine if this drug is safe or not.

This is the target date for the ongoing trial that I believe is currently being discussed between the Agency and the maker of the drug.

Both of these approaches would have protected in this case against making what I think turns out now to have been a mistake. I think we have to recognize that PPARs are,

in fact, a special case. At least 50 INDs have been filed following the last approval of a TZD. Nearly all terminated due to toxicity. The toxicities observed in animals are also evident clinically. Cardiac, skeletal, muscle, renal, bone marrow. This is from a presentation by Jeri El-Hage, formerly of the FDA.

Most development programs were terminated without any publication of the toxicities encountered. This is the negative publication bias problem that we don't know why these drugs failed. All of these PPARs activate different genes and must be considered individually. This is not a drug class. This is a series of individual agents that do individual things.

Let me show you. This is a very nice paper published in 2004 that looks at pioglitazone, troglitazone, and rosiglitazone. And some genes that are gene expression is in common. But each of them

have genes that uniquely are activated or suppressed by the individual agents. And then what you'll see shortly in a publication is it turns out that one of the genes activated by rosiglitazone regulates a key matrix and taliprotinase (?) that's involved in plaque rupture. And this probably may explain why there is an increased risk of myocardial ischemic events. And it turns out that this particular MMP is not involved in stroke.

And so my prediction is in 2014, we get the data. What you're going to see is that there is a very substantial effect on plaque rupture and coronary events, but not necessarily on stroke. And it may relate to one of these genes. But you have to realize that these drugs all affect different genes, and they all have to be looked at individually.

So I believe that the goal of merely lowering blood glucose levels is too

simplistic. We must reduce the complications, including CV disease.

With respect to CV disease, it appears important how you lower blood sugar, as well as how much. I am not disagreeing with anybody who said that hyperglycemia is a really bad thing. But I am telling you that we have to think about how we lower it to get good answers.

Diabetes drugs, even in the same class, may yield dramatically different CV outcomes. Clinical outcomes trials comparing alternative therapies are essential to determine the optimal approach to prevent CV morbidity and mortality.

I am then proposing two components for diabetes drug development. Sufficient pre-approval data to exclude an excess of cardiovascular events -- an upper confidence interval not to exceed 2.0 -- and a robust post-approval outcome program to provide data in a timely fashion. That means an ongoing

outcomes trial at the time of approval. If we have those two things we can still get drugs to patients in a speedy fashion. We're not going to slow down innovation.

I don't want to slow down innovation either, but I want to make sure that the medical community -- that patients and physicians get the information we need to use the right drug in the right patient at the right time. We've got to get off of this glucose-centric approach and get onto an approach that says let's figure out the way to improve health outcomes, not just blood sugar.

Thank you very much for your attention.

Marvin? Oh, I'm sorry. I'm not supposed to call on people. You are.

DR. BURMAN: That's okay. Thank you.

DR. PROSCHAN: You mentioned excluding relative risk -- or hazard ratio of 2.0, but you didn't say relative to what. I mean, if it's

compared to an active comparator, that active comparator may have an increased risk compared to placebo.

DR. NISSEN: Well, again, there's no answer to that. And I think that because it's very hard to do these trials with placebo controls -- one has to then take agents that we believe to be reasonable and drugs that are widely used. And certainly, metformin is one of them, a class of drugs that at the very least in UKPDS looks at worst neutral, and maybe better than neutral.

We can't solve all those questions. But what we can do is at least know that a new agent relative to what we have is not a whole lot worse. And that makes a whole lot of sense to me.

DR. BURMAN: Dr. Konstam.

DR. KONSTAM: So Steve, you started out sounding very radical. Okay. Then you came forward with I think a safety proposal that in fact I think is fairly moderate. And then you

lost me. Okay. So where you lost me is that I understand the safety part. And I think that probably should be a lot of the focus of our discussion in the day and a half to come. But you acknowledge that hyperglycemia is bad. You acknowledge the fact that treating hyperglycemia seems to have a very clear-cut effect on microvascular effects. That those are important therapeutic targets. You know, and then you seem to go on to propose nevertheless you want to demonstrate cardiovascular efficacy.

So -- and in fact, in the statistics that you proposed about your follow on trial, you're really focused on efficacy. So you sort of moved on that.

And I'm sort of confused about why if you acknowledge that there is efficacy -- clear-cut efficacy associated with the glycemic effect -- why is it then necessary to demonstrate cardiovascular efficacy in a follow-on trial? And then the specific question I would have about that is

you're not -- the control group is not going to be untreated. So it's not like we're leaving untreated diabetes versus your drug treatment and then demonstrating cardiovascular efficacy.

And so, in that case, I guess I want to know what level of HbA1c wouldn't tolerate in the control group, because that's really the challenge that our experts have posed to us in demonstrating cardiovascular efficacy.

DR. NISSEN: Okay. Three things. First of all, let me take the last one first. I'm suggesting that people be targeted to the same HbA1c. I think the year is over. I mean, we've had a bunch of trials that have asked the question is lower better. And the answer is for microvascular disease, yes. For macrovascular disease, no. And so I think what you actually want to do is minimize the glycemic contrast between the regimens.

Now, with respect to this question

of efficacy versus safety, we have lost -- as somebody said, we have drugs to lower blood sugar. You know, people are not dying out there because we can't figure out how to lower their blood sugar. You know, we know how to lower blood sugar.

Between insulin and all these classes of oral agents, what we are lacking are agents that improve these macrovascular outcomes. So we've got to move now the next step. If we keep saying we don't need that data -- we don't need to look at that -- then we're never going to find out the answer.

And the last question is if you do an efficacy trial, even if you don't win on efficacy, you establish an upper boundary for the hazard. In other words, if you do a head to head trial of two different strategies -- if you do a head to head trial and you go for superiority and you don't get superiority, but you have a big enough trial, then when you're done you know something very

important about safety. And so my view is that I didn't happen to show you the data using it as a safety analysis, because I think it's better to set the bar up here and say, hey, you want to bring a new drug to market? Show me that you can help more patients with this drug than we can help with metformin, sulfonylureas, acarbose, and everything else that's out there. Show me something new. And that will then give us also the safety information.

DR. BURMAN: Thank you.

Other questions?

DR. NISSEN: Bob. Oh, sorry.

DR. BURMAN: That's okay. Please go ahead.

DR. FRADKIN: Are you proposing an upper confidence level of cardiotoxicity only for diabetes drugs? Or why wouldn't this be something that would be proposed for all drugs? I mean, when you think about the first slide that Dr. Gerstein showed where he showed all of

the conditions which are often treated with medications which are at substantially higher rates in patients with diabetes, and when the latest data shows that 24 percent of people over the age of 60 in the U.S. have diabetes and they have even higher rates of arthritis, and psychiatric disease, and incontinence, and erective dysfunction, and everything else that people are being treated with drugs for -- why would you have this as a particular requirement for a diabetes drug versus any other drug that a lot of diabetics are likely to get?

DR. NISSEN: Because we have priors. To use a term that Bob Temple likes to use. So I'll quote him and say that going all the way back to the university group diabetes program, the question of cardiovascular toxicity that is increased risk. And we've got ACCORD. And we have rosiglitazone. And we have muraglitazar. You know.

And when you have that kind of prior information that suggests that if you

pick the wrong strategy for lowering blood sugar, you can increase morbid and mortal events. And when you have a disorder that's the cause of death in 75 percent of diabetics, then you better know what the effect of the drug is going to be on that population. That's why it makes sense.

You know, if you want to ask the question or sildenafil, for treatment of erectile dysfunction it's a different question entirely. Different population. Different way of use. These are drugs to be used chronically to treat a disorder that ultimately is going to kill because of cardiovascular morbidity and mortality.

DR. FRADKIN: Just to follow up. Don't we have priors though also for psychiatric disease, and arthritis, and a number of other diseases that we have priors for? We heard about erythropoietin. And these are all chronic things, also.

DR. NISSEN: We do. And that's

exactly why we're doing the trial that Tom Fleming described for you. We're studying three different NSAIDs in 20,000 patients, establishing the upper confidence interval for cardiovascular hazard of 1.33. Because we have priors on those drugs. And they're commonly used in people that have cardiovascular morbidity and mortality as a prevalent risk factor. So you have to have a sensible approach. You can't do this for every drug, but you certainly can do them for those where you have some evidence that you might be producing harm.

DR. BURMAN: Dr. Temple.

DR. TEMPLE: This is sort of a follow-up on Marv's question. In looking at your slides I wasn't sure whether you were really asking for a demonstration of benefit or bringing the boundary lower than 2 for the larger, long-term study. And it sounds from your answer like that really is what you're talking about. But if you're doing that, the

most effective way to do it is, once again, to set an upper bound. That helps you pick the numbers, figure out who to put in. So have you thought about that? Would it be the same 1.33 that you're using in your NSAID study? Is that good enough?

DR. NISSEN: Well, Bob, I didn't make those calculations. But I agree with you that that's certainly one approach to doing that. You know, I think it has to be something that is reasonable. Now, here's a way of looking at it. We agonized over the 1.33, frankly, a lot. And we actually required more than that 1.33. And Tom didn't actually drill down as far as he could have, but what we said is that not only does the upper confidence level have to be less than 1.33, but the point estimate has to be less than 1.12. So if we got too many events, we wouldn't meet our upper confidence interval but still have an excess hazard.

And we said that we had to achieve that both in the ITT population and in a

modified ITT population where people were censored 30 days after stopping drug.

Now, you've been a proponent of the fact that a safety study should look at both analyses. So you have to meet four standards in the precision trial to be declared non-inferior. I think that's not an unreasonable level of risk. And I think that's very achievable. Because, remember that the event rates even in the current era in diabetics are significantly high, that if you go to people like we do in the precision trial, if anything, you're going to have higher event rates. I think it probably can be done in a study that might be in the range of 10,000 to 15,000 patients for five years. And we'll get answers. We'll get lots of answers.

But the reason I couched it for superiority is I want us to develop drugs to reduce the morbid and mortal events that are killing our patients in diabetes. I don't

think we should be satisfied with ruling out some hazard -- ruling out that a drug increases risk. We should be trying to draw up drugs that decrease risk.

DR. TEMPLE: Just one statistical point. We have said in ICH guidance and elsewhere that if you go for non-inferiority and win, that's okay. You still win.

DR. NISSEN: Yes. Okay, good.

DR. BURMAN: Thank you. I think Dr. Veltri first.

DR. VELTRI: Steve, just so I understand. In your pre-approval proposal you're talking about, even in the absence of any signal, whether it be preclinical, LDL, weight gain, blood pressure, anything. Is that correct?

DR. NISSEN: Yes, absolutely.

DR. VELTRI: Because to my knowledge there's no -- where you have, let's say, a drug that lowers improves glucose, hyperglycemia or dysglycemia, where that would be the case. And

as you said, I think improving symptoms and improving these microvascular events, which could be quite debilitating, and perhaps to some patients being on renal dialysis or being blind is worse than dying suddenly, which is quick, cheap, and painless.

I think it kind of just puts one perspective into it. And I'd like your comments on that.

And then in regards to the post-approval, since you're advocating a trial that would exclude harm, if you will, to some degree -- your 2.0, wouldn't it be better if indeed you think that your targeted therapy is going to improve macrovascular disease risk, that you would do that actually where you could exclude harm earlier on in a much larger appropriately powered efficacy trial as well so you can design it so that you can look at harm earlier on safety concerns, whatever --

DR. NISSEN: Before approval?

DR. VELTRI: No, no. After approval. In other words, if you have no signal and you do believe that you are gaining some benefit to patients based on glycemic control, that you're kind of relegating further testing where there may not be any biologic plausibility as you would call it.

DR. NISSEN: It's interesting you should mention that term because there's a quote -- one of our follows -- our cardiology fellows have a bulletin board. And they put quotes up there. And there's a quote from me that's prominently displayed. Seriously now. And the quote says, "the road to hell is paved with biological plausibility." And what it means is that -- and that's what happened with muraglitazar. It wasn't biologically plausible to anybody that a drug that raised HDL, and lowered triglycerides, and lowered blood sugar, could actually produce myocardial infarctions, death, and stroke. But it did.

And I want to say one other thing.

You raised another spectre, and I want to directly address this. If we require a higher standard of evidence for drugs in this arena, we are not going to cause people to go blind and have to require dialysis. We have 10 classes of drugs to lower blood sugar. We can lower blood sugar in people. We need ways to lower blood sugar that reduce the complications. We've got lots of ways to lower blood sugar. So we're not going to hurt anybody if we raise the bar here a bit.

DR. BURMAN: Thank you. Dr. Jenkins.

DR. JENKINS: I want to try to get a little clarity in the post-approval study. As I understand it you'd like for them to target a benefit showing study, but failure to show benefit, if they then excluded some predefined upper boundary of the hazard, it sounds like you would find that acceptable in moving the ball forward. If that scenario does play out, you have the signal pre-approval but it doesn't exceed the upper boundary of 2, they do the

well-conducted study targeting benefit. They don't show benefit, but they exclude some predetermined acceptable -- or unacceptable increased risk. What do you propose happens then?

DR. NISSEN: From a regulatory point of view, nothing. Because what you've done now is you've given the medical community what we need to know to make a rational decision. We've done a big study. They know what the point estimate is. They known what the upper confidence interval is. You know, we know what the drug does and what it doesn't do. And it will find its appropriate place in the armamentarium.

But what we have done is we've given increasing confidence to the people that prescribe these drugs on how to prescribe them wisely. And that's what I'm looking for. That's what I'm seeking.

DR. JENKINS: So just to be clear, while you would like to see the new drugs

improve cardiovascular outcomes, you're willing to accept the fact that they don't adversely affect cardiovascular outcomes compared to standard of care. You just want a good study.

DR. NISSEN: And the reason I'm willing to do that is I'm accepting that having choices -- having drugs in the armamentarium to lower blood sugar, given the fact that high blood sugar is a bad thing -- it does lead to microvascular events -- but that's good. You know, different patients will tolerate different drugs. Drugs will have different side effect profiles. You know, alpha glucosidase inhibitors are not drugs you want to give if you plan on riding on elevators.

And so there's lots and lots of issues here related to the overall pattern of adverse effects for a drug. But what I want to do is I want to make sure we're not sitting here 10 years after a drug is approved when it's being used in hundreds of thousands of people and just simply not

knowing, because that's not an acceptable place to be in this day and age. Not when we already have so many drugs out there.

DR. BURMAN: Dr. Fleming, did you have a question?

DR. FLEMING: Steve, can you put up that slide that does show the proposed 2.0, 1.8, 1.5?

While you're putting this up, I share one of the concerns or questions that Marv had asked a little bit ago about what your ultimate evidence would need to be. But then I think the screening assessment that you talk about -- here it is right here -- actually, I think has maybe a bit more merit to it than what you have particularly already formulated.

So specifically, a scenario that I could see would be logical would be to say the definitive trial would be this bottom line. The rationale for that being the definitive trial is you already have

substantial evidence of efficacy and microvascular complications, true clinical benefit. And the goal here, therefore, is to rule out that that is offset in an unacceptable manner by macrovascular complications. And also for patients and caregivers to be fully informed about what the benefit-to-risk ratio would be.

So given that you have substantial -- so, hypothetically, suppose this agent has substantial evidence for glucose controlled microvascular complication risk reduction, then the argument might be given that you could tolerate up to, let's say, a 50 percent increase. You have to rule out that you would have up to a 50 percent increase in cardiovascular complications in order for this to play out. Then, the ultimate assessment would be a 256-event assessment that could, as you would then state it, be underway at the time of the approval.

But what would need to be in hand before the approval would be something that would be -- let's say hypothetically a 122 event scenario -- 122 event trial. This has -- and in fact, a positive result would be any estimate that's no more than a 26 percent increase. And that has the property that if, in fact, there is no access, you have a 90 percent chance of getting a positive result.

It has the other property -- you said if there's an 80 percent increase, that you only have a 2-1/2 percent chance of getting that result. But even more to the point, if you have a 50 percent increase, you have only a 14 percent chance of getting that result. So you're factoring out 6 out of 7 unacceptable agents with this screening assessment.

So really what this would be -- this trial would be a screening assessment to rule out unacceptable safety

risks, ultimately confirmed by a confirmatory assessment. And what you want from a screening assessment is to have low false negative error rates, and you're going to have to give somewhat on the false positive error rate. You're formulating it here as you can only have a 2-1/2 percent false positive area. Well, that's true against an 80 percent increase, but this design also has the property that you have only one chance in seven of getting an encouraging result when you have a 50 percent relative increase.

And I think that's an added feature to motivate the elegance of this that you hadn't brought out. But it's all based on the assumption that this would be the confirmatory trial; i.e., you don't have to show benefit against cardiovascular risks in your confirmatory trial. You would only have to rule out an unacceptable increase in the context of knowing you have favorable microvascular complication effects.

The advantage of this also is this could be a somewhat longer-term trial than this one, such that if you get a different benefit-to-risk ratio over time, you're going to be able to recognize that with this confirmatory trial, which would only have to be underway at the time of the approval.

DR. NISSEN: Tom, just to respond for a second. I guess that that trial I would put down here, and I would make it a 508 event trial to rule out 1.33. In other words, what I'm suggesting here is that that's not quite stringent enough. And we really need to be a little more precise. So I would be more comfortable taking that to the next level.

DR. FLEMING: Well, and that's an issue that can be discussed. The point is, the study that would be randomized underway wouldn't be a superiority trial. It would a trial ruling out an unacceptable excess risk. And you might say that's 1.33. Maybe it is. An argument for why it could be 1.5 -- and this would have to be

well thought out -- would be an argument that you're getting benefits that have been established in other domains, such as microvascular domains. And therefore, that allows somewhat greater leniency or possible increases in cardiovascular macrovascular complications before it would be unacceptable.

But the bottom line is if that's where you draw the line, then this is, in fact, a screening trial. Not specifically targeting 1.8, but targeting 1.5. Saying if it's 1.5, you only have 1 chance in 6 of getting an encouraging result to go on, or 1 chance in 7 if this is truly the case. But you have a 90 percent chance of going on if there's no excess. So it has a very effective screening capability.

DR. BURMAN: Thank you, Dr. Fleming. And I think we have to move on in the interest of time.

There will be time for questions later. Thank you very much, Dr. Nissen.

We're going to take a break now. It'll be minutes. Let's reconvene at 3:35.

Panel members, please remember there should be no discussion of the meeting topic during the break.

(Recess)

DR. BURMAN: Why don't we get started in about a minute or so? Please take your seats.

Why don't we get started for the last session of the afternoon? We're going to end the lectures and discussions this morning by Dr. Robert Califf, who is vice chancellor for clinical research at Duke.

Thank you very much for coming.

DR. CALIFF: You guys looked really tired towards the end of the last session, so I'm going to -- what I'm going to try to do is provoke -- at least to try to keep you awake here -- being a little bit provocative as I talk about the issues at least I've encountered in trying to design some of these trials.

I also feel emboldened with you. I spent the morning with the National Cancer Institute with the problems they're having in oncology clinical trials. So that's a pretty tough crowd. You guys couldn't be any rougher than they are, I'm sure. But many interesting issues there.

When I think about this topic, this is the time of year in Washington and further south that we drink sweet tea, so I can't help but think about the problem that you're addressing here is unsweetening the blood. And whether that's a good or bad thing is really what we're here to talk about.

So I'm going to try to give a bit of a conceptual framework. Talk about the key tradeoffs.

In the midst of this, bring up some issues about barriers to implementation. And then finish with -- I guess I'll start and finish with a comment on the status quo.

Hertzel and I were sitting there

saying that we agree with about 85 to 90 percent of what Dr. Nissen said. But as I go into the issues and implementation of clinical trials, let me just say I can't imagine a situation, given what we know now other than approval based on a screening mechanism somewhat like what Dr. Nissen described. And then pragmatic clinical trials that really answer the question of whether the net balance of risk and benefit, not just for cardiovascular disease, but for a really true whole body net benefit versus risk is answered.

And as to the comment about should this be for all drugs, my personal belief is that chronically given drugs -- because of what we now know about the biology of what drugs do -- should all be studied if they're going to be given to large populations chronically with enough patients and enough outcomes to truly measure the balance, the benefit, and risk. And so that's sort of

where I am.

Now, I had the privilege a few years back of being asked to work with Dave DeMets, whom I admire greatly. Has also worked a lot with Tom Fleming to sort of think about what we've learned about therapeutics of cardiovascular disease. Here, we're talking about cardiovascular outcome trials. And these are some truisms that we came up with which are almost always true. And I think can be verified to almost always be true through any sort of systematic or non-systematic look you want to take at it. Many of these have been discussed already this afternoon in the part of the meeting that I've been able to listen to. And I'll talk more about them as we go through them.

These were published in circulation. And they're all pertinent to what one needs to think about in designing an outcomes trial in cardiovascular disease.

So I'm now going to just -- by way of background, I'm just going to give a commentary that I hope will stimulate discussion tomorrow as you get into recommendation making mode about what's behind all this. And I think what's really behind the change that's needed in the way we think about these clinical trials is that we've learned a lot about therapeutics. What we've learned causes cognitive dissidence with what we'd like to believe or what we wish was true.

And so we've continued to operate in a mode for regulatory approval, labeling, advertising, and prescribing, particularly in the United States, based on what I call an advertising mode. Which is take one concept that you believe to be true, focus on that concept, and sort of screen out cognitively all the other dissident information that's hard to assimilate and deal with.

So Steve's slide about TZD -- he

could take any class of drugs and show exactly the same slide and get exactly the same answer with regard to gene expression.

We now know whether we measure gene expression -- the proteome or the metabolome -- that drugs within the same class cause different patterns of response in whole organ physiology -- our whole body physiology. And the reason for that is that most of the targets the drugs are hitting are in systems that we don't know about yet. So on-target and off-target effects are important, and systemic therapies affect many targets at the same time.

Yet we behave as if looking at one parameter gives us assurance that the net balance of risk and benefit to the whole individual can be measured by that one target. And this is sort of repetitive of Dr. Fleming's slide, but it's been a heyday in the last couple of years, not just in diabetes but in almost every area of

therapeutics. Someone brought up psychiatry which may be the kingpin now where even a drug that we've been using for years, like intravenous Haldol now has a black box warning about cardiovascular risk.

So all of these drugs affect multiple systems. They all cause a balance of benefit and risk often in systems that were not intended. My favorite one by the way with TCD is not cardiovascular. It's bones. Something that was picked up by looking at clinical trials.

Secondly, we know that the effects of most therapies on humanly meaningful outcomes are modest, so randomization is essential with large sample sizes. And yet, we still behave as if doctors can tell whether chronic therapies are having a net beneficial effect by their memories of their own patients.

And, in fact, if you look at yesterday's New York Times, you'll see a

discussion of at least one prominent cardiologist advocating that the most important thing we can do is to get rid of the idea of evidence-based medicine, which I thought was a very interesting concept.

And we also sometimes behave as if looking at post-randomization database is going to tell us about post-marketing treatment effects. And I think there are many reasons, most notably that most treatment effects are modest -- that we really can't do that in pretending that doing multiple analyses of poorly controlled data will give us the answer as a mistake.

Now, this is not a new concept. Aspirin is probably a drug you recognize. It's been along for a long time. And you would probably agree that it has significant cardiovascular benefit. And yet, if we look at the direct-to-physician advertising that existed in medical journals and then direct-to-consumer advertising in the 1950s,

you'll notice at the bottom there are aspirin -- this is FDA-approved labeling -- does not affect the heart.

That is what we believed after millions of people had been treated with aspirin. It was only after proper trials were done that we really were able to talk about this. So this slide that had a critical effect on my career from Salim Yusuf just makes a point that we've got to measure a lot of events to detect the kinds of effects that for a dominant disease that is the leading cause of death and disability in the economically developed world -- those are the kind of effects that we really need to understand. Modest effects are critical. This is not my mantra. This really emanated from many others who have been preaching this for a while.

Thirdly, we know that the effects of therapies are context dependent. And one of the big issues we were discussing in the

oncology meeting today is the question of whether we're just so inept at doing clinical trials in the U.S. now, we should just do all of our trials in China and India where they cost about a tenth as much to do and import the results. Just like we get our shirts and shoes made in China and India now.

But I think there is ample reason to believe that is an inappropriate thing to do for the American public. By the way, I do think trials should be done in China and India for obvious reasons. People in China and India need good therapies, too. But we also know their interactions with other treatments. They're common and unpredictable. The length of treatment is important. You've had a very good discussion about that already. And the clinical environment matters.

And yet we behave -- and it's frequently said -- that we're doing testing of drugs for measurement of human benefit as

if we were in a laboratory over a short period of time. We're controlling everything instead of operating in the real world environment is the right way to do it.

We also know that therapies cause a mixture of benefit and harm often involving different organ systems over different periods of time. You've had a good discussion about that. And yet we behave, and still imply to the public, although the direct language is not this way -- this is still what the public often believes -- that short-term studies done pre-approval can actually provide assurance that a drug is "safe and effective."

I think we know now that's simply not the case. Because if you're going to give the drug over a long period of time, different things happen in different organ systems over time that you just can't anticipate.

We can -- I think as Steve said and

Dr. Fleming more elegantly and statistically pointed out -- we can screen and reduce uncertainty, but we can't assure the public that drugs are safe and effective based on small studies that don't measure integrated balance of risk and benefit.

This slide from Curt Furberg, I think, makes a point in terms even I can understand. We've always got this mixture of good things and bad things. Good things and bad things happen in different ways to different groups of people, and also happen in different ways to different groups of people over different periods of time.

There's an example that I wanted to give of sometimes good things happen when you measure long-term effects. And this is from a trial recently reported in The New England Journal that we coordinated looking at a bisphosphenate to prevent fractures. And lo and behold, it prevented fractures as expected, but no one had done a trial that

lasted more than 24 months. The data monitoring committee stopped the trial, not just because fractures were prevented, but because there was a 28 percent reduction -- 25 percent reduction in overall mortality -- total mortality in the trial.

We don't know why that happened, and there are many theories. But the point I want to make is this is not all about safety. I would suggest that with chronically given therapies, we will find a number like statins and ACE inhibitors that the more we look at them, the greater the benefit is that we observe in the broader population of people. But the bottom-line is we don't know unless we look and empirically measure because doctors' memories are not adequate to account for all this complexity. And post-marketing, uncontrolled studies can't possibly give us the answers about modest effects.

Another great example, of course, about varying effects over time comes from

hormone replacement therapy. And the point I want to make here is sort of the inverse of what we normally talk about. And that is if one had looked in the first six months at either HERS or other Women's Health Initiative, you would have stopped for terrible harm. HERS actually came together over time in both studies, which is kind of an interesting phenomenon.

I have no idea what it means, but the point here is that the treatment effects are not constant over time. This is sometimes the case, and sometimes not the case. We don't know until we look.

I put this slide in particularly because I knew that Ruth Day was going to be on the panel. And I want to pause here for a minute and just make the point that these disturbing things that we know about therapeutics that don't fit the way we've done things in the past cause us to want to block them out and continue with the way

we've been doing things. And that, of course, is as I've learned from our business school, is really the key to advertising. It's connecting things through a story or a picture that make sense. And shielding out all the other contradictory information that might cause you to question what you're doing.

But I hope this panel will really question what's been done in the past, not because there were bad people or there were bad ideas in the past. We've just learned that things can be different. As I'm going to talk about as we design these trials, if we cut out the ridiculous bureaucracy that we now have in many of our trials, we can actually do these trials and get the answers that we need.

And then the final few. This gets to the point I was making. Our current methods of implementing trials are harmfully and unnecessarily bureaucratic and expensive.

And so we behave in a manner that says we can't change the cost of trials, so we just have to find shortcuts, even if we're accepting a large amount of uncertainty for what these drugs do to people chronically.

And then we all have biases and conflicts of interest that prevent one sure answer. This is one of the more difficult things I think about the whole enterprise of clinical trials. Thousands of people put in millions of human transactions. Someone presses a button and you get a result. And yet, there's so many decisions that are made in designing a trial and even interpreting what the analysis is in a trial that there's not one sure answer.

So we behave as if companies can conduct their own trials "hiring" investigators without independent study management and analysis of the results and produce unbiased results. In fact, yesterday we just had an encounter with a company that

wanted us to sign a contract that gave them assurance that if the results of the trial were negative, we wouldn't mention it for two years. This is still going on. It's routine in the clinical trials business. And I think those that are thinking about the design of trials need to consider these issues, too.

But I would also point out that just because a trial is done by the NIH does also not assure that it's without bias. This is a human enterprise. We all have biases. And in fact, I would argue in the design of our trials, it's balancing the interests of people with different biases that really represents the key to a successful and well done trial.

Now, more evidence that we've got a problem. This is an old study from the Lancet -- Reasons Why Clinical Trials Are Not Published. Dr. Nissen has referred to it. I think clinicaltrials.gov and the World Health Organization are helping out now to make sure

that we do at least see the bottom-line on trials. But if we actually get into the details of why trials are not published, very often it's for a negative result suppressed by industry. But in our own analysis now, we're finding equally as often in investigator-initiated trials at our best academic centers, it's because there was a negative or unsatisfactory result from the point of view of the bias of the investigator.

So it gets to the main point. A balance of interest in the design of these trials is critical. And we have to recognize now important this balance is. Because if we don't have the balance as Dr. Ridker pointed out in review in JAMA, it's likely we'll only see trials that are designed to be positive in the first place. And that's something we definitely don't want to have.

But to blame the rest on all of us, also in the design of trials, the

acknowledgment of what our conflicts are is critical. You should know, if you don't, that all of us, including all the speakers as far as I know on this panel, routinely sign contracts to participate in clinical trials at sites that do not require that the results be published. We published this in the New England Journal. The good news is academic centers are willing to give us what they put in their contracts.

The bad news is none of them require that there will be a publication from the trial of the entire study results when participating as a site.

So in the design of trials, we have to think beyond just what's the question being asked? What's the statistical analysis? We also have to think about the societal balance that will insure that the results of the trial really get out in the way it should. JAMA has been focused on this. And we just did our own survey in my

field -- just to show you that I'm not picking on diabetologists. We reviewed coronary stent trials done in the year 2006. This paper was rejected by multiple paper medical journals, and you may see why in a minute.

What we were doing was looking at acknowledgement of conflicts of interest in the reports of clinical trials about coronary stents. What we found was that 83 percent of the time in 2006 there was no acknowledgement of a conflict. And equally as interesting to me, when there was an acknowledgement and an author had more than one trial reported in the same year, the acknowledgements disagreed the majority of the time.

So there is no consistency in the way that we're dealing with this on the academic side. This is not just an industry problem.

Okay, so what is a balance of power that can be had in an outcome based clinical

trial? It's one in which there is a sponsor who does participate but doesn't control. There is a steering committee that participates but also doesn't control. There's a balance of power in the way that studies are done.

And now that these kinds of trials that we're talking about here are going to be global by their nature, it's very important to have global participation of thought leaders that represent different cultures, different views of how things should be done.

You would think that in 2008 I would not have to mention that there should be an independent data monitoring committee. I know that's an FDA rule, but this is something that needs to be watched carefully. It's not the case in every field that this is being done even today.

Okay, so now what about specifically designing the trials? And we're in the midst of designing a few of these now,

so I'm going to confess the problems that we're running into. This is not easy, and at the end of a couple of hundred million dollar experiment with 14,000 people, you hate to find out that you did it wrong. So I don't have answers; I have opinions.

So these are the five, what I call, big ticket items. Trade off of generalized ability and validity; looking at target versus drug; looking at superiority versus non-inferiority; trial conduct -- what I think of as sensible of nonsensical; and what I now refer to as regulatory disharmony. There are also some important details. These are not details to people who design clinical, but they're sort of the second order that are critical. And I'll review each of these briefly.

So this is the simpleton's view of one of the key issues in designing clinical trials. We'd like the trial to be perfectly valid, and the tendency if you stick to that

is to go to the lower right hand corner, get as valid as you can, and do an experiment which is very carefully controlled and excludes many of the people that would be getting the drug.

On the other hand, a registry will give you something that includes the whole population, but without the element of randomization and some control you don't have a valid study. So the goal is to get to the upper right hand box as much as we can. And I would argue that the pragmatic trials, as it's now called, is really a way of finding the best compromise between those two things. And there is no single best answer. It depends on what the drug is intended for; it depends on who is really going to be using the drug; and it depends on what people think is going to happen when the horse is out of the barn and the drug is on the market in different countries.

So if we're focused on generalized

ability, I brought entry criteria. We allow any concomitant therapy. We embrace standards of care but avoid detail protocols, and we do the opposite if we're focused on validity. My opinion, as you might already know, is that if we want to know about outcomes in large populations that represent people that are going to take the drug, we really just need to stick to the common ground -- proper consent, randomization, measurement of whether people are taking the drug which is still important, measurement of the endpoints, and unbiased manner.

What really amazes me, and I think it kills a lot of the creative thought about this is how often industry SOPs and FDA inspectors fail to distinguish what's important from what's fundamentally irrelevant to answer any question posed by the trial. And this leads to hundreds of millions of dollars of waste as I'll show you that causes people to conclude that we just

can't do these trials.

What about target versus drug? You've had a good discussion about this already. I actually believe that we need to do both. The NIH is about to embark on a large target trial with hypertension to look at the upper limit of taking it from 140 to 130. Very similar to what ACCORD is doing in diabetes. And unfortunately, you can't answer both questions in the same trial. You just can't do it.

I agree with Dr. Nissen. Torturing the data leads to a lot of interesting thought, but it cannot answer the question. And you just need to do two different kinds of trials for two different reasons.

I'm not opposed to torturing data; I enjoy it myself, but we shouldn't be using it to make policy if we can avoid it.

So we need both kinds of trials. We need therapeutic target trials to understand whether, in general, it's

beneficial to drive a biomarket to a target. And we need drug-specific trials to know about this balance of risk and benefit. Again, because the risks are very often not in the target-specific arena. They're in a whole different biology that we don't yet understand.

In the end, and this is really important, there is no magic bullet. We're left with some uncertainty about this mix. And so as people think about this, I would urge you not to think about -- and I think one thing I would say to Nissen and Fleming maybe for future discussion -- depicting this as a linear pathway to screening and decision I think is a mistake. This is going to take a "patchwork of trials" carefully thought out that address somewhat different issues with each trial.

Then we have the superiority-non-inferiority construction. You had a good discussion about this already.

It's hard to imagine that a treatment that lowers sugar shouldn't decrease macrovascular disease. But the superiority is going to be a tougher and tougher hurdle as we find some treatments that actually do reduce macrovascular disease and are proven to do so because then you can't exclude them from use in the trials. And we'll get into the era of comparative effectiveness, which is now a mainstay of many other areas of medicine.

Noninferiority, though, has its problems. I still have found no one, including Tom Fleming, who can explain this to ordinary clinicians in a way that they can repeat it if they leave the room and come back a half an hour later. It just doesn't fit into the way we think about things, and yet I think it's really important because I do believe what was said already by several people here. While we would all love to have a treatment that reduced heart attacks, strokes, deaths, and microvascular disease,

if we had a drug that really reduced microvascular disease and didn't kill people, that would be a very important thing to know.

So I think superiority-non-inferiority is really a false argument. The real question is what's the estimated effect of the treatment on the net balance of risk and benefit. A trial can test for both if it's properly constructed. And that's the way a lot of trials are currently being constructed. The first test being can you show that you're not hurting people, and the second test being you proved you're not hurting people -- can you show that you're actually causing a benefit.

And so the real question is what's the minimally important clinical difference that should be excluded in non-inferiority trials or exceeded in superiority trials. And here I hate to sound like a radical compared to Dr. Nissen. This may be the first time in quite awhile, but we know from

20 years of intensive discussions, focus groups, questions to patients, questions to providers, that for a disease that affects tens of millions of people all around the world and is going to grow by more than threefold over the next 15 years, particularly in developing countries -- that a 10 to 15 percent relative difference or a 1 percent absolute difference per year is clinically important, and it will change the way people treat patients.

And you can do this with a simple thought exercise. If you have a treatment taken by 5 million people and it increases the risk of death by 1 percent per year, it would kill 50,000 people a year. I would argue that's a number we need to know about and something that should be excluded if possible in non-inferiority trials, even if it made your neuropathy better or reduced renal dysfunction.

At least then people could make an

informed decision about the tradeoffs that they wanted to make.

Now, I learned -- I did a little reading in preparing for this -- I learned there's actually a huge debate about where the quote came from. But I think this is sort of the key to me about the whole thing. A difference to be a difference must make a difference. And if you want to read a fascinating story of someone who spent a couple of years at Hopkins training in medicine and then did something entirely different, read the story of Gertrude Stein. I would recommend you look it up if you don't know the whole story.

So then we get to the design of sensible versus non-sensible clinical trials. So the goals of a medical intervention if you ask people would be I would want to do it if it caused me to live longer, feel better, avoid unpleasant events, and spend less money -- or spend less money and keep all the

rest constant. Except for a few people who need help, your average citizen is not interested in spending money on medical care if it's not going to be beneficial in a tangible way. And so since surrogates work for on-target and off-target effects separately -- this is a discussion I've had with Temple many times -- I don't argue that blood pressure is a good surrogate for stroke. But systolic blood pressure is not a good surrogate for off-target effects of any hypertensive drugs. And unless you know both, you're sort of stuck. And off-target effects we now know are ubiquitous thanks to being able to measure large scale genomics and proteomics.

I refer you to the Journal of Clinical Trials, where there's a six-part series on sensible clinical trials where we got together academic, industry, FDA, European regulators, and we did a bunch of thought exercises about if you accepted the

premise that the public needs to know what the long-term benefits and risks of drugs and devices are -- and behavioral interventions, by the way -- and you thought that the cost of doing the studies was keeping people from launching the studies that were needed, what could you get rid of where you could still get the same answers but spend a lot less money. And I'll have you read the details for yourself and see what you agree or disagree with.

But fundamentally, this is the bottom line. And I think people were in a state of shock this morning. The entire budget of the NCI cooperative clinical trials is $150 million. In outcome studies in cardiovascular disease, as some people in this room know quite well, people are spending $450 million a trial for a single trial. So our whole government cancer portfolio is a third of the cost of some single trials being done in this field. And

yet, when we got people together and said what could se stop doing and still get a valid answer, we came out with numbers that were at least a third as expensive. And if we went to a radical extreme, we came out with numbers that were about a tenth as expensive.

So you've got to ask the question, is it really worth $450 million to make sure that concomitant medications that are stopped and start at multiple times during a six year trial, or recorded every time they're stopped and started, and every time a patient gets nauseated some study coordinator at $80,000 a year has to record whether the patient was nauseated and when it stopped and whether they thought it was related to the drug. I would argue that's really stupid, but that's what's happening. And it's really putting an impediment to launching these trials and answering the questions.

So I refer you back to Janet

Woodcock, who I admire quite a bit, who gave a talk at the BIMO (?) meeting just last year, and reiterated what happened five years ago at an Institute of Medicine meeting about the quality of clinical trials. So the definition of a high quality clinical trial with regard to the data is one in which the data is good enough that the decision wouldn't change if completely accurate data were used.

And if there's one thing I want to implore you to do in designing these trials, it's get rid of the junk that doesn't help you answer the questions that the study is designed to answer. Save the money and do two or three times as many trials, at least.

And that gets us into regulatory disharmony. So which makes more sense? Doing a separate trial in every country or conducting global trials? I would argue it's obvious conducting global trials makes more sense. But how can we do what makes sense if

regulatory requirements are different in every country? There's a perception, which I think is true, that this is going the wrong direction in the last few years.

What do I mean by the wrong direction? Let's look a little bit at what's at stake here. These are slides from Bob O'Neill at the FDA from internal analyses. These were presented at a Pharma meeting. I had his permission to promulgate these widely, although to my knowledge this has still not been published in a medical journal.

This is looking at cardiovascular trials that are housed within the FDA. In looking at regions of the world as a factor in treatment effect, particularly with an interest in the U.S. You can see there are a number of trials. And on average, these were all beneficial trials. These are trials where the treatment effect was in the right direction. However, on average the treatment

effect was less in the United States than outside the United States. And when taken as a sum, this is actually a significant difference.

So that if you look at U.S.-non-U.S. as a stratifier, there's an interaction between treatment effect and whether the patient was enrolled in the United States. And this has come out in some individual trials. This is a fairly famous one in cardiovascular disease where the trial overall was dramatically positive. The study was stopped for benefit. You'll notice that there are many zeros in front of the first number and the p-value, but when the subgroup U.S.-non-U.S. was looked at, there's a slightly less than neutral effect in the U.S. and a grammatically positive effect outside of the U.S.

So getting in sync and understanding that there may be regional differences, we don't know what all this

means. I'm not giving answers here. I'm just saying if you're looking for simple answers, you're probably not going to find them now.

So here's what we face as we try to do this. Given differences of opinion by regulators in different countries, the sponsor has to either reduce the number of countries -- that is eliminate the ones that are demanding useless bureaucracy -- or revert to the most expensive common denominator.

This leads to what I think of as a very vicious cycle. And it goes like this. If the trials are too expensive, we can't do them. Therefore, we'll just have to accept or ignore uncertainty in order to enable development of new drugs. And I think this is a very dangerous way for us to go societally given what we know now about chronic therapeutics.

What we need is a virtuous cycle of

developing common methods to reduce uncertainty. And it's not enough to do this just within the United States. It has to be done on a global basis if it's going to work.

All right, so now to the details quickly. This is a list of what we commonly argue about as we're designing these kinds of trials. The enrollment criteria -- very important set of issues here. If we take patients early in the disease -- and I think Marvin was alluding to this in one of his questions -- maybe we have more of a chance to modify the disease. And the diabetologists commonly hold out for this approach. But the event rates are low so it takes forever.

In the Navigator trial, Cleveland Clinic is adjudicating events, I think we're now in Year 7 of the trial. And it's an endurance contest to see if we'll get to the end. But we enroll patients with a low event rate who don't have too much disease to start