Dan Perry: I want to start by thanking all of you for braving the weather. Believe me you'll be dryer and more comfortable here for the next three, four hours than you will be anywhere else in this town. So thank you for making the trip. Thank you especially to our speakers who are going to enlighten all of us this morning on thoughts about making Phase II clinical trials for Alzheimer's drug discovery a more efficient and more predictive process than it is at present and have been in the recent past.

For those of you who are not familiar with the ACT-AD Coalition, the acronym stands for Accelerate Cures and Treatments for Alzheimer's disease. It's a coalition made up of over 50 not for profit organizations that represent the interests of patients, their family caregivers, consumer organizations, seniors and women's health advocates. And we have this coalition in place now for four years and we've found that the most effective way that we can help facilitate a smoother passage for the cures and treatments that we so urgently want is to engage our colleagues at the Food and Drug Administration in meetings, workshops exactly like this one.

And each year we pick far in advance a topic, a subject, an area of development that is particularly open to discussion and possibly to some adaptation that will help the whole process move forward. And this year's topic was actually suggested by Dr. Katz himself, who told us some 12 months ago that Phase II clinical trials are beginning to look more and more like Phase III clinical trials should look: large, extensive, time consuming and risky. And so we have reached out to some of the clinical leaders in the Alzheimer's community from academe, from among clinical trialists, and an industry and we pooled the best thoughts that we can get this morning on how to make this process more efficient. And so that's what lies ahead of us.

We know that there are over 100 clinical entities that are somewhere in the developmental pipeline. The pipeline is richer than it has ever been. The time is drawing nigh when the baby boom generation is going to start moving into ages where they, themselves, are going to be at risk for this disease. They're already taking care of their older parents and older relatives. So we know that the time is absolutely ripe for this kind of discussion.

And the FDA's role is absolutely critical because they need to gather it from the scientific community the emerging consensus around how to best measure the effect of therapies and how to identify the patient population that is most appropriate for clinical trials. And so that's what this process is all about.

Again, I thank all of you for being part of it. I hope and expect that you will find the presentations and the discussion and the back and forth this morning enlightening, enriching and we're hoping that we move the needle a little further to the ultimate desired goal that's so important for all of us and for our country.

I also want to extend a thank you to two of the member organizations of the ACT-AD Coalition who are represented here today. There are a number of them represented. The two that have joined us for this morning's workshop as co-hosts, and that's the Alzheimer's Foundation of America and the Alzheimer's Drug Discovery Foundation. And I'm going to first ask Sue Peschin, who is the Vice President of the Alzheimer's Foundation of America, Vice President for Public Policy to say a welcome also. Sue?

Sue Peschin: Good morning everybody. On behalf of AFA, I want to welcome you here today as well. And I want to say that our organization is honored to serve with ACT-AD and the Alzheimer's Drug Discovery Foundation as a co-host for today's discussion.

For those of you who might not be familiar with us, the Alzheimer's Foundation of America is a national non-profit organization. We're based in New York and we unite more than 1,400 member organizations all across the country that range from non-profit organizations that provide services and supports for people with the illness and their family caregivers to some state entities like the AAA's and the units on aging, to a lot of public safety agencies, such as police departments and fire departments that engage in search and rescue for people that wander and become lost as a result of the illness.

Our services include a toll free hotline that's staffed by licensed social workers. We have a free quarterly caregiver magazine that has over a million subscribers. And I want to invite any of you that have new programs or research that you would like to get out there to the community to please let us know because we'd love to feature you in our magazine and we've done that quite a bit with things that CMS has done. This is for federal agencies as well as academic programs. So please reach out to us because we would love to feature you in our magazine and the work that you're doing.

We also have an AFA team support and scholarship program. We do every year in November for Alzheimer's Awareness month, National Memory Screening Day, and we also provide grants to service organizations as well as respite grants to families in need.

And last, we also do professional training for individual dementia caregivers as well as for dementia care units.

We believe that our strength comes from collaboration and, when it comes to advocacy, really only two things count. One is timing and the other is working together because, in my experience, you really can't accomplish anything alone, especially here. So being on the verge of the baby boom explosion, I think that we all have pretty good timing right now in terms of the moment that we're at in history. But, in terms of working together, that's sort of another story.

And I wanted to give a couple of examples of how many of the Alzheimer's related groups have been working together lately. One is that more than 100 national and local organizations have signed on in support of the National Alzheimer's Project Act, which is a bill that would create the Country's first national office that would be responsible for organizing efforts across the government to treat and prevent Alzheimer's disease. And due to these coalition efforts, the Senate version of the bill is scheduled to be marked up this Wednesday with House action soon to follow. So that's really, really exciting because we have not had an Alzheimer's specific bill go through really, so this would be a huge victory for us.

The second is that, just this past week, the Alliance for Aging Research, AFA, Cure Alzheimer's Fund, Friends of the NIA and a new coalition called Leaders Engaged in Alzheimer's Disease, also known as LEAD, came together to organize House and Senate Hill briefings to highlight the work of the National Institute on Aging and to call for more funding in fiscal year 2012. The briefings are part of a broader budget effort that we have going where we met with the Office of Management and Pledging and have done a sign on letter that was signed by more than 20 national organizations. We also have an upcoming meeting with NIH Director Dr. Francis Collins, and we're gonna push him.

So we want all of you to stay tuned for how these efforts play out and we may also ask some of you for your help because we're in the process of organizing another letter and we'd love support from the scientific community to call for more funding.

And the last thing I want to mention is something that's been unprecedented, which is happening this Saturday evening and it's a national telethon. And that's gonna be on NBC in about 15 major markets across the country, including in the D.C. area. In D.C., it's only gonna be on for about a half an hour from 6:00 to 6:30

PM and it's for the Alzheimer's Foundation of American and it brings together celebrities, major media personalities, our workers, actually our host for the event. I went to the taping and it was very inspiring and impressive and we highlighted early onset. We have a lot of local organizations that provide direct services and support in there. So there's a lot of good information in addition to entertainment as well.

And the point of it is really to show how it not only impacts the individuals with this disorder, but also places an emotional toll on families and a tremendous economic burden on American society.

So I do want to say that the work that you're doing today is more important than ever for people who are diagnosed with Alzheimer's and for those who have yet to be diagnosed. According to the National Institute on Aging, as many as 5.1 million Americans have the disease and these numbers are projected to increase with the aging of the population, in particular the baby boomers.

So I want to join everyone and Dan in particular in thanking Dr. Katz and his colleagues at the FDA for the work that they do and for being open to exploring solutions to existing problems that are ultimately, hopefully gonna help all people with Alzheimer's Disease.

We've established ourselves at the Alzheimer's Foundation of America as a leading organization for advancing care, support and increased research funding. And through our work and in coalition with other people, the Alzheimer's Foundation of America hopes to put an end to the devastation caused by Alzheimer's disease. And I thank you very much for being here and I look forward to today's discussion.

[Applause]

Dan Perry: Thank you Sue. And now, rounding out our welcomes from our co-host organization, Dr. Howard Fillit, President and CEO of the Alzheimer's Drug Discovery Foundation. Howard?

Howard Fillit: Yeah, thanks everybody, and I'll be really brief. Our foundation has given away 40 million dollars in the last 12 years. We've seen about 3,000 different ideas for new drugs for Alzheimer's disease. We've funded over 300 of them. Most of it's been pre-clinical and some of it has gone up into the clinic and we've funded a number of Phase II, really kind of Phase IIa trials.

I say this in our context here because there's so many ideas out there for new drugs for Alzheimer's and many of them are high risk and we really don't know where the next drug for Alzheimer's disease is going to come from. And it's kind of a shots on goal type of thing. There are so many drugs that are on the market today come from serendipity and science, and the examples of that are innumerable.

So we need good drug discovery science. We need good pre-clinical science. We need good medicinal chemistry. We need animal modelers that really know how to do that, that part of the work. We've cured animals with Alzheimer pathology over 300 times as we all know. But the real experiment that we have to do is to take these model pills and give them to you and see what happens. And we know in CNS drug development that the most common cause of failure today, we've gotten around a lot of the predictive safety issues, that the most common cause of failure in CNS drug development today is lack of efficacy.

So the key step in what we're all doing is called enterprise of basic research channeling into drug discovery and ultimately getting into units. This is really at the Phase II, and I believe particularly the Phase IIa stage where the first signal of efficacy is gonna come from. And that's why I think what we're gonna talk about today is so important because, as was mentioned, so many Phase IIa studies are looking today like Phase III trials and I think that there's gotta be more innovative ways to do Phase II, to get a signal on efficacy and then be able to convince the deep pockets of investors and to inform them to take on these risky and extremely expensive Phase II B and Phase III trials that really ultimately can lead to an FDA approval hopefully one day, particularly for a disease modifying agent.

So I want to thank Dan for the great work that Dan's doing in putting this all together. It's a great honor to be co-hosting this day, Dan. Thank you very much and I look forward to a really open discussion about how we can do Phase II trials in a cost effective way that we can afford that will increase our shots on goal towards you know Phase III ultimately, and approved hopefully soon for it to be a treatment.

[Applause]

Dan Perry: Before I introduce our first speaker, I want to echo what Sue said about our gratitude to Dr. Katz and all of the FDA reviewers and staff and officers that join with us on a regular basis for this kind of

a workshop and I appreciate all of you for being here and being part of this process.

We are going to have questions following the talks, and especially during the panel discussions, and I encourage all of you to jump right in. This is not a didactic session where we just have speakers up here speaking at you. This is really a roll up your sleeves and let's engage on the issues. So, when that time comes, either address the speaker or the panel from the microphones here or, if you're too far inland, just raise your hand and we'll get a microphone to you.

We don't have any regular break scheduled at present, so please help yourself to the facilities. There are refreshments in the hall and now we're going to get started.

And we're going to get started with a review, a clinical review of recurrent and recent clinical trials in Alzheimer's disease and we're very fortunate to have Dr. Rachelle Doody of Baylor College of Medicine to lead off that step. Rachelle?

Rachelle Doody: Thank you Dan. Well it's a pleasure to be here and I think a terribly important topic that we are all addressing today. I want to say that every time I encounter these three organizations that are hosting today, I think thank goodness they exist because they do so many things in so many different venues.

My charge today, and we're not gonna make up the time during my talk, but my charge is to just lay some issues on the table. I'm obviously not here to solve the question of Phase II clinical trials in Alzheimer's disease, but I'd like to put a little bit of a historical view on it and a little bit of a current issues view on it and see what you think.

So let me start off by something that, in some people's minds, is a provocative point. There are evolving concepts of memory disorders and dementia, and I'm supposed to represent some of what's being thought in academia. Clearly some changes occur with aging and I think everybody in academia believes that these may be targets for intervention.

So we're not just talking about Alzheimer's disease. We're also talking about aging. There are identifiable risk factors; that's pretty undeniable. And the fact that Alzheimer's disease neuropathology starts decades before there are symptoms is also pretty clear. But, as I will point out, we don't know the relationship between the risk factors and the neuropathology and that poses a huge problem for trials design.

And then finally, certain cognitive changes clearly constitute early Alzheimer's disease. And, in the centers of excellence, we no longer worry about this. We see people with amnestic MCI who have a certain cognitive profile and no ADL decline and they're Alzheimer's disease and that's their diagnosis and they're not told you're maybe Alzheimer's disease and they're not told well we won't know unless you die. They're told you have Alzheimer's disease. And, in practice, you know especially in a place where we follow about 80 percent of our patients until they die, this turns out to be the case.

So some of the risk factors; I mentioned that we clearly know there are risk factors for Alzheimer's disease and some of them are listed here, the ones that I think are really best supported by the literature: aging, menopause, the APO, lipoprotein E4, genotype, elevated glucose or hyperinsulinemia or diabetes or some combination of the above manifested in many different biomorphic markers like hip to waist ratio. It's all getting at the same thing. Elevated cholesterol, elevated homocysteine, major head injury with loss of consciousness and something or other about inflammation.

So to illustrate the point that I made before, and I know that this is a diverse audience in terms of scientific versus policy expertise, these slides come from a paper about Down's Syndrome or trisomy 21 and about the unfolding of neuropathologic changes really over the lifespan of these patients with amyloid deposition in the brain in their 20's; no cognitive impairment. About 10 years later, microglial changes or signs of inflammation associated with that neuropathology of amyloid. Another 10 years go by and you have neurofibrillary tangles and probably the beginnings of some symptoms.

In patients with trisomy 21, we're never really sure that Alzheimer's is present until they begin to lose function. So it's as if we can't diagnose Alzheimer's in a trisomy 21 patient until it's moderately severe. And then we get the neuronal loss, neurochemical changes that we associate with Alzheimer's disease, and then a clear dementia.

So if this is the case in Down's Syndrome, where the average age of onset is in the 50's, if we're talking about sporadic Alzheimer's disease where the average age of onset of the dementia is about age 74, you can just go back and say that well, this is probably starting in the brain in the 40's. And I think that there is neuropathologic evidence that that is the case.

So what are some of the strategies for the anti-dementia drugs under development that we're going to be talking about? I like to group them in as few categories as possible. So there are the drugs in nutraceuticals based on epidemiologic observations. And remember, these are the observations that we're all quite convinced have something to do with risk, but we don't know what they have to do with neuropathology. Then there are the neurotransmitter based therapies and still, if you look at all the drugs under development for Alzheimer's disease, a fair number of them fall into this category; glial modulating drugs; neuro protective drugs which, when I have a bigger slide, I might add in, and metabolic drugs and regenerative drugs all into this category because what is neuro protective in the test tube, we don't know what it does in the living being.

Cynthia told me her machine is always trying to synchronize with the server at home, so we'll see if we can't stop that. I can't.

Okay, the next category is Tau modulating drugs, of which now there are several, even immunization approaches which who would have thought we'd be taking that approach with neurofibrillary tangles, which are intraneuronal; and then amyloid modulating drugs, which are really the focus of much interest.

Now I want to say that, in my experience, at least when this whole field of real therapeutics in Alzheimer's disease began, Phase II used to predict Phase III. And, if you look at a list, and I think I tried to be pretty representative here, it was pretty much the case. Now you know you can quibble over things. When you talk to Larry Sparks, you know Atorvastatin was positive in Phase II. But, if you really look at the paper and you really hold people to the primary outcomes you know without sub-populations or post-hoc analyses, at best you could say it was equivocal. He said it was more of an effect in people of a certain age with a higher cholesterol level.

Donepezil, Galantamine – I didn't put Tacrine on here because Tacrine was really quite a complex story. That was the year I entered the Alzheimer's field was the year we began that work. And so, to me, Alzheimer's has always been a treatable disease, but, at the same time, it was a chance to observe well just how did people decide whether to take a drug like Tacrine into Phase III. And, at the time, that was all being just furiously negotiated. There were no anti-dementia drug guidelines, so they got written as the Tacrine story unfolded.

Okay, so Phase II used to predict Phase III pretty well. And what did these trials have in common? Well they expected people to get better. They expected you give a drug, you got Alzheimer's, you get better. They expected that we'd be able to see it. If they got better, we would know. And, in fact, Paul Leber insisted that we better know; otherwise they're not better. So we had to have a clinical benefit on top of a psychometric benefit.

The trials were short. We were just finishing the era of crossover trials. So we were into an era of parallel group, double blind placebo control trials, but they were three months or six months long. And, other than that, a couple of them were a little bit longer as time evolved. And they didn't have background treatments; that's the other thing to keep in mind.

So now we talk all the time about the divergence between Phase II and Phase III results. And I put a question mark there because you know sometimes some of us really didn't think that Phase II was positive. But here we were with massive, multimillion dollar Phase III programs growing out of Phase II. I think it's worth looking at some of the more recent ones.

So Dimebon: a Phase II trial designed very much like the other studies that I just showed you. Randomized double blind placebo control parallel group, six months. The added spin was a six month blinded extension, which we did not do in those early trials. Positive on all five outcomes. By the way, monotherapy at a time where many places would not approve or justify monotherapy; in other words no background treatment, and so it was done in Russia. Okay. So there's our first notable divergence from the older trials. And all five outcomes were positive.

Move to Phase III; multi-national. Same design as the Phase II trial. Still no background treatment and negative. And, without going into the details of any of these trials, no placebo decline. In fact, the placebo got better on two measures in this multi-national trial.

What about Rosiglitazone? It started out with a fairly standard six month parallel group study but, you know, some would say it was a negative study because it didn't meet its pre-specified outcomes. However, in looking at it more closely, it was observed that, if you just looked at the E4 negative individuals, it looked like you might have a signal there. So what begins to happen here is we begin to redefine the signal and I think that's the theme of Phase II trials in Alzheimer's disease. We're redefining the signal.

So we defined it as okay in E4 negative individuals and GlaxoSmithKline developed a really thoughtful, broad, huge program in order to explore this issue with enough power to stratify by E4 genotype. The background therapies were handled in an interesting way. In some of the – in one of the Phase III studies, patients could be on a cholinesterase inhibitor whereas, in another, they could be on a cholinesterase inhibitor and Memantine.

So you begin to get decisions about all of those branch points of what could happen in a clinical trial became a different clinical trial. And, in the end, they were all negative. So that program was ended.

So Migalastat: here we took a completely different approach. It was a two month long biomarker study. And based upon earlier data, it was presumed that it would move CSF-A beta. And, in fact, it did. In the Phase II study, it lowered – I'm sorry – it lowered serum A beta but not CSF-A beta. Okay, so it didn't do exactly what was expected, but we redefined the signal again. We said, well, there could be a lot of reasons for that, why different compartments would behave in a different way, and it could still be a biosignal and that's what we need in Phase II. And, of course, in that short duration of time, we could not possibly expect to see efficacy. And so, we went a different route.

And, as you know, the Phase III trial was long. It had a delayed start. It had variable backgrounds of stable treatment allowed; in other words, not a specifically specified treatment and it was a negative study.

Tarenflurbil or Flurizan: it was a standard Phase II study except that it was a little bit longer; not a six month trial. I think it was a one year trial on background. And, in this case, specified; it had to be a cholinesterase inhibitor. And again, many people said at the end of Phase II, "Oh, it's a negative study," because it didn't meet the pre-specified ADAS-cog criteria. The global was not positive. In fact, nothing was positive in the overall groups, either by dose or doses combined.

However, it was observed that, if you split the group – and here was the interesting design feature – the Myriad people said well this isn't a post-hoc analysis because we built that into our statistical analysis plan. If we didn't see an effect overall, we would look specifically at the mild patients because this is an anti-amyloid drug and maybe you really need to be mild in order for it to have an impact. They did pre-specify that, but they didn't

power the study to separately look at the mild patients. So we're left with an under-powered observation that, in the mild patients on the high dose, there seemed to be an effect on more than one outcome. So, again, we redefined the signal and did a long Phase III program, very similar to the Phase II, which was negative. And, by the way, the Phase III program here actually specified a mild population, followed up on that post-hoc lead, and limited the population to mild patients.

Tramiprosate or Alzhemed: again, the approach was a short bio-signal study. Let's look for Tramiprosate to alter, preferably CSF-A beta values. And, in fact, it did. It lowered A beta (1-42). So a standard long study on stable background treatment was designed.

However, when you try to operationalize stable background treatment in the Alzheimer's field, you cannot and should not look at the patient and say, "No matter what happens to you in the next two years, you can't change your therapy." Patients who understand their rights would simply drop out of the study, but patients who don't understand their rights would be upset by that and have to be protected by an IRB.

So even with stable background therapy, which in many patients in this study included a cholinesterase inhibitor and Memantine, there were drop-ins and dropouts and anti-depressants came and went. And, for a long time, Paul Aisen believed that that was the explanation for the negative Phase III studies.

So, final example: Xaliproden. The Phase II was a standard, but somewhat longer, randomized, double blind, placebo controlled study. Now, to my knowledge, the Phase II results are not in the public domain. I did not go back right before this lecture to see if they had been published. Cynthia? But the trial that was designed to follow up on Phase II was again a standard but longer study on stable background treatments. By the way, the very same approach to background treatment that we saw in the Alzhemed studies, and this was not the explanation for a negative study.

So we began to loosen up the criteria for what constitutes a signal in Phase II. We began to accept more in the way of post-hoc analyses, and it has not borne out that this improves our Phase III trials.

The strategy that people talk the most about – and when I say people I mean the public, the investment community and, to a large extent, academics – is the anti-amyloid approach. So the senile plaques that look like cigarette burns in the brain, we do

understand fairly well what produces them. The production or failure to remove A beta and its coalescence into dimers, trimers, fibrils, aggregations with an inflammatory response in the brain, leading both to the anti-amyloid and anti-inflammatory approaches.

And then again, as a review for those who don't think about this every day, if this is the culprit and we said, okay it's there early – look in the Down's patient; it's the first finding. Keeping in mind that you can have amyloid deposits in your brain and live to be in your 100's and not have Alzheimer's disease. So it may be unnecessary, but not sufficient condition. Nonetheless, in order to get it, you have to have this pink fragment, this beta amyloid; you have to excise it from the amyloid precursor protein through a beta secretase and a gamma secretase operation. And so beta and gamma secretase inhibitors or antibodies are conceivable approaches. If you can't stop it from being produced, perhaps you can stop it from aggregating or kelated in some way or have a protein metal interaction attenuation or immunize in order to remove it.

So what are we doing with these strategies right now? And this is, by definition, any of these slides are going to be incomplete. This is based on what I know. But there are a fair number of passive antibody approaches as well as IVIG as you can see; some active monoclonal antibody vaccines. The anti-fibrillization, anti-aggregation approaches, including scyllo-inositol and PBT2, both in pretty much the same stage of development, having both completed Phase II programs; beta and gamma secretase inhibitors that remain.

I always like to point out that there are still other active treatments in similar stages of development; it's not all amyloid. And so I selected several. The SB-742457, 5-HD-6 antagonist which has recently read out its Phase II and part of its Phase III program and still has a study underway. ST-101, the adeno associated viral vector or NGF study using this technology to enhance NGF production; an interesting representative anti-oxidant. There are several others as well as several PDE inhibitors. And then the latest versions of methylthioninium chloride or Rember, sons of Rember that are getting ready to get up and running.

So it's important to list these because, if you contrast these to what was on the last slide, you've got completely different drugs and there's probably no way that you could have the same Phase II trials.

Lest we be too complacent, we should remember that there have been a whole host of recent Phase II discontinuations, including the TTP-488 or Pfizer renamed compound that just failed in the ADCS trials. A high dose had to be stopped because of possible side effects and that definitely influenced the futility analysis of the compound.

So where am I going with this? Well, as I've seen it, the evolution of Phase II designs is an attempt by people to hedge their bets. And I have to be a little bit careful with this slide because you know all of you have worked very hard to come up with your own Phase II designs, and that's your intellectual property. And so none of these represent anybody's single approach, so they're a little bit grouped together to protect that.

So what are people doing? Well some are saying, "I absolutely believe the amyloid story. I absolutely believe that we've got a disease here that's already too far advanced by the time a person is sick. We're gonna have to treat this disease before you get sick." And they're moving to pre-AD dementia, which requires some type of biomarker with or without memory loss and different sponsors have made different decisions about that. Are you gonna treat the truly asymptomatic or the one whose memory is so bad that it's Alzheimer's? Okay, so that's one of the decision points. Or, if you're gonna define it based on the CDR because, thanks to Washington University, we've got a fairly good longitudinal history of what happens to the CDR. People with CDR 0.5 generally get Alzheimer's.

And then, what are the outcomes in these trials? Well one hope is that brain A beta levels – since you're really basically putting down all of your chips on amyloid, then look at amyloid in every way that you can. The onset of dementia might also be a significant outcome to measure and/or the clinical progressions. So, depending on how this pre or early AD population is defined, you can still measure clinical progression, understanding that you're giving up a lot of the range of cognition and, therefore, you need a very sensitive measure.

Disease modification: each of these approaches tries to at least open the door for claims of disease modification. So here we're using mostly neuro imaging and CSF measures. These Phase II trials are running one to two years, as Dan was saying in the beginning. They're long. And then other options, well even if treatment is allowed, in some cases they're actually making the statement that patients cannot go on treatment later. So, if you're a patient with an Alzheimer's kind of memory loss and nothing else,

and you agree to a two year trial but that you won't go on treatment, I think this is going to pose operational difficulties and it's going to lead to biases in what kind of patients get in the study at what kind of sites and/or it's going to require international trials.

So that's one general group of approaches. Another is to go for an early AD, somehow enriched population. I really should have put that in quotes because I'm not convinced that this means something different when one of these Dubois criteria are present. So you have 100 Alzheimer's patients in front of you who have mild disease and you do a PET scan on everybody and some percentage of them will have an abnormal FDG PET. That percentage that is completely unknown. And, if you look at the recent papers that are coming out, particularly in Europe, to address the issue of sensitivity, what you find is that you might be losing half the patients if you require two biomarkers, either low CSFA beta and/or high CSF tau and/or FDG PET or hippocampal atrophy.

You know whatever it might be, what you're doing is you're taking 100 people who look like they have Alzheimer's and throwing some of them away. And maybe that's to their benefit and maybe it's not. But, in the absence of data that convincingly says it's to their benefit, I think there's a note of caution.

In general, enriched criteria population studies use clinical progression as the outcome. They have taken approach that evidence of disease modification in Phase II is optional. I think many of the sponsors, when they begin running the numbers, they say sell, you know, if we don't have a drug, do we really need to know if it's disease modifying or not? Disease modification markers will not trump the loss of clinical benefit, so maybe we'll save those particular outcome measures for Phase III. Again, one to two years and you have to keep in mind that you will have high screen failure rates in these studies.

Interestingly, people don't want to pay for those either. They want to pay for 5 of them or 10 of them, but chances are that we're gonna see 40 percent screen failures. It depends on how you put the criteria.

We also have Phase II trials with standard mild to moderate AD populations. If you believe that the way we used to trials, that we expected a notable clinical impact on patients, if you believe that's still the case, even with anti-amyloid therapies and other therapies that we're using now, then this is the way to go because you have a broad population. You don't know in advance who's going to

respond more. You do know which placebo patients are going to decline more. And, to date, we have required placebo decline to show benefit.

Disease modification optional in Phase II: the trials again are going back to the theory that, well, it could be quite short. So there are some that are three months, some that are six months all the way up to two years using this population and this strategy. We may not see placebo decline and monotherapy would be difficult to insist upon in these trials.

Finally, I have seen some people who have elected to enhance the fact – which one do you click so I can do it myself? Just close it – who have decided to enhance the observation that placebo decline is more notable in moderate patients. And there are some companies designing Phase II trials just in moderate populations understanding that there will be a problem with generalize ability later, that they may have higher dropout rates, but they trade off a shorter study.

So, to end, I think there are a couple of salient points that come out of this type of thought process. Okay. One – by the way I just, I found this chevron slide totally by accident. I clicked on something and this slide appeared and that's how it became this type of a presentation.

So, anti-amyloid therapies may require asymptomatic or minimally symptomatic patients. They may require this. But we should not assume that all treatments will. And I'm beginning to see a generalization of this view.

Second: monotherapy requires global trials. You just can't get enough representative patients at sites in the United States and Western Europe to do it that way. We seem to have accepted that some treatments will not improve patients. If this is correct, we will need long trials in treated patients.

And finally, Phase II may have to allow for many different trial designs, depending on the mechanism of action, the concomitant therapies and the disease stage. In addition to that point, even one drug may need more than one Phase IIa study because, if you want to optimize a study to look for the biosignal, that might be a very different study than a study that looks for a clinical signal, a signal of efficacy. I think the two examples of studies that were optimized for the biosignal have shown us that.

And finally, we should not lose sight of the necessary goals in Phase II: safety data, the dose and the signal of efficacy.

So those are some thoughts to just sort of open the discussion, but our panel, later on, will address these and many other issues. Thank you.

Dan Perry: Alright, thank you.

[Applause]

Dan Perry: I'm going to contradict myself and ask you to hold your questions so that we can move onto our next couple of speakers and then we'll have a panel. And following the first panel, we will really open it up. So please bear with us.

Our next speaker is Dr. Jeremy Hobart, who has come all the way from the United Kingdom where he is at the Peninsula College of Medicine. And Dr. Hobart will talk about some recent research on refinements of cognitive scales. Dr. Hobart.

Jeremy Hobart: Thank you very much for inviting me to speak. I hope what I have to say is of interest. So far we've heard that the main problem seems to be lack of efficacy detection, and I would ask you to consider whether that is because the drugs don't work or we actually can't measure any changes.

The nubbings of the issue from our perspective is that you can make a big difference to rating scales but you have to systematically evaluate them and you have to move beyond current tests of reliability and validity into the new psychometric methods and that's what I'm going to take you into. And this is a piece we've been doing now for a little while on the ADAS-cog with the aim to enhance its measurement performance in clinical trials of MTI-4 or AD or, in fact, whatever where you might want to measure cognitive performance. These are my co-workers, all of whom are critically important to this work, but I won't discuss them in the interest of time.

So what I'm gonna do is show you what can be achieved. And then I'm gonna give you a simplification of how it can be achieved. And then I'm gonna take you through the measurement limitations of the ADAS-cog and build up on the program of work that we've done and then how this might be applied in a clinically useful way.

We have a goal, and our goal is to vastly improve the ADAS-cog, not just to modify it and make it a tiny bit better, but to vastly improve it. But we need it to be grounded within the solid neuropsychological conceptual model. We need to benefit from

the application of sophisticated measurement techniques and, ultimately, we need to meet regulatory requirements.

Why do I think that's possible? Because the specific limitations of the ADAS-cog, the benefits of applying new methodologies, the genius of ADNI, and that has enabled us to do this. And you have to get into the conceptualization of measurement instruments as rulers.

This is the ADNI sample in terms of its distribution of cognitive performance measured on the ADAS-cog. This is the ruler mapped by the ADAS-cog and the simplicity here is that you see that the notches of the ruler don't extend over the sample and there are great big patches, especially in the mild area where you want to pick up changes and differences.

There we've done – we can achieve this; that is we can stretch the breadth of measurement to cover the entire ADNI sample and we can produce precision across the whole range. And, in doing so, you get exactly what you need: the ability to detect group differences. So here – and if you look at this distribution in the kind of greenie color – this is people who have a mini mental state score of 29 or 30 distributed over a very wide range. And this is what we would expect to see clinically that there is a range of cognitive performance, even in normal people.

In doing so, we also improve the ability to detect changes at the group level. So this is ADNI over two years using the original ADAS-cog, very small effect size, non-significant. This is the developed version of the scale and effect size 35 times larger and a highly significant difference detected

But you can also move these issues to the individual patient level, which is where we need to be going. Here you can say that, with the improved instrument, greater numbers of people developed a statistically significant worsening and, in fact, you shift this whole distribution down.

Now I've used this concept of a ruler analogy and some people find that a little bit puerile, but actually it's exactly what we need because, whenever we use a rating scale, we are asking it to perform as a ruler of some sort. So, when we use the ADAS-cog, we're asking it to perform as a ruler of cognitive performance. But remember, we get numbers on pieces of paper and we want to derive cognitive performance measurements that we can locate on a line on which we have gradations.

How do you get those gradations? Well, if you look at any component of any scale, for example the commands item of the ADAS-cog, here you see that there are five different commands that people are asked to do and they either get them right or they get them wrong and you get a summation of the number right. And conceptually, if you think about this, when we're trying to measure cognitive performance in terms of on a ruler and we're gonna use this commands item, we would say that the people who get five commands right are kind of up the more cognitive performance end of the ruler and the people who have no commands right are down at the less cognitive performance end; and that we could envisage a situation that, as you get more commands right, your level of cognitive performance is advancing.

What we need to do is locate the points of transition between getting naught and one, one and two, two and three, three and four, and four and five. And those points of transition become the notches that we have on the ruler. Now, when you examine and analyze the commands item within ADNI, you see this. That is that you get your five gradations that divide your cognitive performance ruler into six chunks, but they have different sizes as you would expect and you don't have enough for the types of precision that we would want. Had you get that precision, you add in more questions.

So this is what happens when you add in the constructional PRAXIS item. You start seeing that you now move from having five different gradations to nine different gradations on the item, on the ruler. And, if we follow that logic through, this is what you can get. But the implication of that, that is breadth of measurement with lots of gradations, you need many items. So, having small amounts of items will never give you breadth and precision of measurement of the type that we need. And, of course, we need measurement precision across a very wide range so that, when we measure people who are essentially normal, that we get precise estimates likewise all the way down the ruler.

Now to pick up on the story and see how it works through, we have to start with the limitations of the ADAS-cog. And I think they can be summarized in its simplicity in one slide. This slide shows the distribution of person measurements. This is in ADNI; this is 3,401 measurements. So this is the distribution; the histogram showing the distribution of cognitive performance measurements. That's all people, all time points. The bottom distribution is the patchy scale that I showed you earlier. The green line shows the range of the scale. You will note that the range of the scale is mis-targeted to the range of the sample. This

is the milder end; this is the more severe end. The people who are off the end of the ruler essentially are those people who have milder forms of AD.

What we know about scales is they are at their best at their center point; that is their sweet spot. And by that we mean that is where they give precise measurements with narrow standard errors. For the ADAS-cog, that's a value of 35 out of 70; equates to a mini mental test score of 14.

What we also know is that you move away from this, the errors of measurement get larger. So what we have here is the people we're trying to measure are essentially off the end of the scale and associated with large measurement error and, therefore, your ability to detect differences and changes over time is very, very severely compromised.

So how do we make the transition from the limited ruler with patchiness, gaps and poor range to having wide range and precision over the whole area? Well to do this, we have to start and go back to the conceptual model of the ADAS-cog which, in its original papers, outline the measurements of four different areas: memory, language, constructional PRAXIS and ideation of PRAXIS. And these are the individual components of the ADAS-cog that represent those areas.

The first thing one needs to do is unpack the information that you were given by the ADAS-cog. For example, in the naming objects and fingers component, there are 17 different tasks. You have to name 12 objects and 5 fingers or 4 fingers and a thumb. That information is truncated down to a score of naught to five. So information is being lost. If you unpack that, then you start to unpack the ruler.

We also find that, in the word recall task, you have three tasks that are meaned in terms of their score. Information is potentially being lost. In the word recognition task, you can score a maximum of 12 from 24 scoring opportunities. Once again, information is being lost.

So the first thing is to unpack the ruler as it stands. But the second thing comes from the genius of ADNI because ADNI contained and collected a whole load of other bits of information at the same time as the ADAS-cog. And you can work with the original conceptual model of the ADAS-cog of memory, language, construction PRAXIS and ideational PRAXIS. And, within that conceptual model, you consider adding the additional measures or

at least some of the additional measures collected at the same time concurrently.

But one of the things you have do here is collect the information at the item level and ADNI, one, didn't collect all the information at the ADNI level. It tended to put in total scores for most of the information. So, with the help of the ADNI ISAB and the Foundation for the NIH, we've had the opportunity to collect the item level data and proceed to the next level of analyses.

The second thing we use is a criterion, a hypothesis test for measurement. So this isn't about modeling data; this is about taking a hypothesis and testing its performance against a scientific criterion for measurement.

And that helps us to move from this, which is the unpacked version of the ADAS-cog, which is wider than the original, to this, which is the full version using as much of the data as possible. Now this needs to be refined further, but there are two things you need to note here. Note how the distributions are changed. What we're starting to see here is normalization of distributions for people at all levels of what here is mini mental state scores. And particularly what you're seeing is a distribution of cognitive performance in people who are normal. And that's what we would expect to see and that's what we want to move back towards which is individual person mapping.

You'll also note the change in F-statistic, which increases essentially three times showing a greater ability to detect group differences. And in doing that, building range and precision, then you have the ability to detect greater differences at the group level.

And there's the slide I showed you earlier. It shows the comparison of essentially the change detected in about 140 people of the ADNI sample over two years. Using the original ADAS-cog as it is scored originally, there's a small effect size and a non-significant difference. Using the newer instrument, unpacking it and adding in the information also collected at ADNI, you increase the ability to detect change such that now we're picking up significant differences with an effect size change 35 times greater.

And also with these new methods, you move to legitimate individual person analysis. With standard instruments and typical ways of scoring them, they're really not for use at an individual person level because the standard errors are very, very wide around any one person's score. When you move to new methods and apply them, you can start to get bespoke standard errors for

individual people. This means we can take an individual person and they can track their own change relative to the standard errors that they have at each time point and then we can look at those people in terms of how they change. And I think this pattern is exactly what you expect to see within the ADNI sample.

Now the problem with that is, as you will all have noted, is that, to achieve that degree of range and precision, you have to administer an awful lot of tests. So how can it be used in a clinically meaningful method? Well, it can be done very simply. ADNI collects data. This can be a computerized version of this. This is simply an onscreen mechanism whereby people enter the data as they're going along. The data gets handled within an easy management system from which you can withdraw or use any data you want at that stage. It then gets beamed through the new psychometric method program, which is calibrated based on our original work.

Slight interlude.

And that then delivers you what you want; gives you the individual person measurements. It allows you to track them onscreen over time, in real time; it gives you their bespoke standard errors; you can compare them against other people and all populations.

So, in summary, we believe it is possible to notably improve cognitive performance measurement. And by that, delivering range of motion with precision, linearity and bespoke individual level standard errors. The work we have done has built on the conceptual model of the ADAS-cog. And that's important for a number of reasons because it means we can equate any person's measurement on the original scale or any modification of it with the expanded instrument that I've shown you a little bit earlier. But, by building on the conceptual model of an instrument that has been used as the primary outcome measure in over 127 clinical trials of Alzheimer's disease, it has regulatory potential for being acceptable.

This is a hypothesis driven exercise which tests performance against a very stringent scientific criterion for measurement. We use simply differently handling the same level, the same data that other people or the people have collected, certainly with ADNI, over the last five years. So it has usability. The end users have to do no more than they already do. In fact, they have to do less because the information can go straight into a computer based program. So the burden is the same if not reduced.

You get everything you've had and a lot more. You get all the scores on all the scales that you wanted in the format that you have been previously. But, in addition to that, you get linearized measurements of cognitive performance with individual person bespoke standard errors.

It's applicable retro and prospectively. So ADNI, one, can now re-analyze its entire data set in terms of the relationships between the biomarkers and a newly developed instrument and reconsider whether the instrument, the limited range in precision of the original ADAS-cog, has, in fact, reduced our understanding of the role of biomarkers. And the most important thing is that you can move onto targeted measurement.

What I've showed you here is a situation that is derived from 594 items worth of tasks for an individual person to face. Because of – the picture I'm just about to show you when I get to it – because of this, all you need to do when these instruments are calibrated is give the people at this end the items here, because these are all irrelevant. The people in the middle, the items here because these and these are all irrelevant. And the people down here, these items, because anything above there is irrelevant.

So the potential for this we believe to be, to assist in the understanding of treatment efficacy is profound. Thanks for your time.

[Applause]

Dan Perry: Thank you Dr. Hobart. And I know all of those who are participating in ADNI are gratified to hear the utility that's coming out of that. And I also want to thank back to Dr. Doody's and her very comprehensive, if sobering, presentation of some recent trial history.

Our next speaker is Dr. Les Shaw of the University of Pennsylvania, who also will bring us up to date on some of the new insights that are coming out of the ADNI and similar studies.

[End of Audio]

Dr. Les Shaw: Other tremendous contributions of other expert centers and studies as well. At the outset of the ADNI study, when John Trojanowski, Virginia Lee and I considered how would we approach the question of diagnostic utility, or neuropathological implications of CSF biomarker measures of A beta 42, tau and phospho-tau, we felt that it would be very important to use an ADNI independent

population of neuropathologically diagnosed at autopsy cases, where pre-mortem CSFs were available for interpretation.

So once we sharpened up the measuring system that we're using, which is ALS bio 3 x map immunoassay system, we then applied it to that question; can we let us define threshold values for these biomarkers, as well as for selected ratios, as well as for a logistic regression model that includes CSF tau, A beta and ApoE genotype. And that data is summarized on this slide, and what this shows you are several important principles that have followed through the studies as we progress. One is that sensitive detection of AD as compared to cognitively normal matched for age controls is achievable using A beta by itself. CSF, on the other hand, because of the prevalence of abnormal A beta levels, plaque burdened in the cognitively elderly normal population, with average age 75 or 76; the specificity is modest at 73 percent. So a sensitivity, then, a sensitivity of 96.4 percent – the nine is missing there – that's 96.4 percent, and a specificity of 73 percent for those reasons. On the other hand tau A beta seemed to be a bit more of a balanced biomarker measure, with more balanced sensitivities and specificities, in the middle 80 percentile as we move forward.

Now, when we take that information and then go ahead and measure these biomarkers in cerebrospinal fluid, at baseline – obtained at baseline in the ADNI cohort, we see detect – we see abnormal A beta, 42, abnormal tau A beta ratio or abnormal logistic regression model in about 90 percent of the AD folks. That's an important number, we think, and want to talk about the implications of that, because at least 10 percent of the folks who are diagnosed clinically with probable AD have perfectly – relatively normal biomarker measures; not only CSF, but also other of the imaging biomarkers as well. And then about 70, 75 percent of the MCI folks and about a third of the cognitively normal elderly population showing abnormalities in A beta or in the tau A beta ratio.

Now, so that's basic information about the prevalence of the abnormalities in the ADNI population. A second question that we're all interested in though is what about the predictive performance of these measures in predicting progression from MCI to clinical diagnosis of Alzheimer's disease. In the ADNI population, after the first year, it was 16 percent of the MCI folks progressed, and after two years, 40 percent. Whereas if we look at the control to MCI folks, a very – a very small rate of progression, 1.4 percent at one year, 2.39 percent in two years, and yet a third of those folks have abnormalities in these biomarkers.

Now, a number of different groups have done these analyses and this is just one representation of a survival analysis, but with some important implications. About 15 – looking at the tau A beta ratio, which is a bit better than the A Beta by itself, in predicting progression at the earliest time point, within six to twelve months after the baseline period, we see significant separation, based on in the MCI population of ADNI, the criterion of greater than or less than the tau A beta ratio. Approximately 15 percent of the folks with a normal tau-A Beta ratio are converting to a probable AD. About four times as many of those folks with a ratio above 0.39 are converting to AD. A very significant predictive performance, but nevertheless, we still have folks who have normal biomarkers converting to a diagnosis of probable AD; somewhere between 10 and 15 percent.

We'll come back to that. But clearly one of the important – next important points regarding what we're all discovering together from the ADNI and other studies, are important interrelationships between these biomarkers – biochemical biomarkers and imaging biomarkers. And this is work from Wash. U. (Washington University), Anne Fagan's paper, one of her several significant papers, showing correlations between tau and phospho-tau and decreasing normalized whole brain volume in the very mild to mild dementia with – of the Alzheimer's type subjects.

Another important, in terms of scientific basis of these measures, is the correlations with neuropathology. On the left is a very well-known study published in 2003, showing a correlation between CSF A beta 42, versus plaque counts in the neocortex and hippocampal regions. And two important points; one is the correlation itself, very significant, so the lower the A beta, the greater the plaque count in this study. But secondly, we notice there are both demented and non-demented folks, another underlying principle of the fact that this is a measure of plaque burden. It's not a measure of do you have dementia or don't you.

A second kind of correlation that's very important that came from the ADNI study and others are replicating this in different ways in different studies, is the correlation, a very good correlation, between CSF – CSF A beta lowering and an increased cortical PiB reading. This study has been expanded within ADNI and outside of ADNI and there is an emerging consensus, an emerging conclusion from the data that there's a really good correlation between CSF A beta measures and PiB, or – and soon to be further enhanced and developed as a type of plaque burden measures.

Again, a good measurement of plaque burden, but again, with serious significant interpretation aspects that I think we're all familiar with and comfortable with. And then just to re-emphasize the fact that in cognitively normal elderly folks, as we reach age 50 to 60, somewhere in there, we start to see a significant increase in plaque burden measured by PiB or PET, or in this case by cerebrospinal fluid lowering of A beta.

Another very important basic principle is that data integration in this field is at its infancy. It's beginning to accelerate, and these are just a few samples of the studies where biostatistical analyses are at a concerted level being undertaken to integrate imaging as well as biochemical biomarker measures, with the overall result that we do see better predictive performance, for example, in predicting progression from MCI to AD when combining two measures. For example, CSF A beta or a tau A beta ration and hippocampal volume, as an example of that.

Because of time, and I did want to be brief with respect to some of the highlight science that's going on in this field, no question the support provided by the NIH, the NIA, Dr. Sotos and Neil Buckholtz and others have been a tremendous part of this effort, as well as are collaborators. On that note, thank you.

[Applause]

Dan Perry: Thank you very much, Dr. Shaw. I know you'll be back a little later this morning on our second panel on biomarkers. Our last stand alone speaker is Dr. Ron Black of Pfizer, who will tell us how to pull all of this emerging new information together, how the industry is using it in the design and the carrying out of their clinical trials. Ron Black.

Ron Black: Okay, well, thanks Dan and thanks to ACT-AD and the AFA and the ADDF for putting together this great conference. There we go, yeah. Okay. So I'm gonna talk about – I'm gonna talk about how we've used some of these – some of these interesting findings about biomarkers – excuse me – in populations like ADNI in our potential disease modifying agents and trials of potential disease modifying agents in Alzheimer's disease. And Dan put it very well when he said the problem really with Phase II is that it's getting more – more and more large and expensive and time consuming and risky. And so I'm gonna share with you one of the results of our large, expensive, time consuming and risky Phase II trials that we've done within the Pfizer Janssen Immunotherapy Alliance, formerly the Wyeth-Elan Immunotherapy Alliance.

So what do we mean by disease modification? I'm talking about trials here that are designed to show that an agent is disease modifying, and basically this has a lot of different definitions, but it can encompass a change in disease pathology and also a change in disease pathology that relates to change in the progression of clinical disease, which is a little harder. And why is this important? Well, we feel it's important to know the mode of action of a drug, certainly, 'cause that informs the prescribers. It really demonstrates the value of therapy, whether the taking of a drug will have long lasting consequences on the progression of the disease or whether it just represents an improvement in symptoms. And also you would think that these disease modifying therapies are the ones that particularly might want to be selected for treatment in early patients, patients with very subtle symptoms or even in the prevention of symptoms in patients who are at high risk.

So what are the role of biomarkers in these disease modifying trials? Well, first of all, the biomarkers may reflect target engagement or proof of effect, and I say proof of effect because proof of concept and proof of mechanism, they all have different meanings to different people, especially in industry. They're usually pretty precisely defined. And because of that, they're really useful for decision making in early phase studies. You know, when you're asking the question do we have anything at all.

Also, they may be useful for Phase II studies because the change in the biomarkers might precede the clinical effect, but we don't know whether this is because the biomarker change is really, you know, upstream from the clinical effect or it may be just related to measurement precision, and some of these biomarkers could be very precisely measured. Certainly the biomarker change may indicate an effect on disease pathology, but not all biomarkers correlate well with disease pathology and it's not always known, although, you know, certainly the – you know, the genius of ADNI, as Jeremy said, is contributing greatly to this knowledge. And also, of course, you know, we're interested because biomarkers might provide evidence for disease modification. But in fact, to use a biomarker to provide this evidence, you really need to assume that you have an understanding of disease pathophysiology, and that clearly is – the pathophysiology of Alzheimer's disease is not completely worked out, I think.

So these are the biomarkers that we've used in the Pfizer Janssen Immunotherapy trials; the volumetric MRI, fluid biomarkers of

various kinds and also PET imaging. And just to go back to the original findings that sort of launched the whole field of immunotherapy, here you can see this is – this is the pathologic findings in PDAPP mice, and these are transgenic mice who have a – who have been given a mutant – a mutant gene for APP. And this is a good model for Alzheimer's disease, but it's an incomplete model. The mice don't really get much neuronal loss and they certainly don't have tau pathology. But what they do get is they get amyloid plaque burden, astrocytosis, and dystrophic neurites. And you can see here, treatment with AN1792, which is just the A beta 1-42 protein given as a vaccine, actually prevents and can even to a certain extent reverse these three neuropathologic hallmarks that you see in these transgenic mice.

So the drug that I've been spending most of the last 10 years working on is bapineuzumab and bapineuzumab is one of the many monocle antibodies that has the same effect as the vaccination, in essentially reducing the amyloid burden in these transgenic mice. And you've already seen this – certainly the seminal findings of Dale Schenk and his co-workers at Elan, launched this giant effort, maybe enrolling tens of thousands of patients in these Alzheimer's disease immunotherapy trials.

So just to sort of go over, as I promised, the biomarker results from these – from these early phase trials. First of all, volumetric MRI, and you've heard some about this, but, you know, again, the loss of brain volume really is the best correlate of clinical disease progression, and it's also good – one of the other reasons that we use it a lot is you can measure it very precisely. In measuring what we call the boundary shift interval, you actually take advantage of the complexity of the convolutions of the brain to register images in time – you know, a baseline image and a follow up image, in a very precise way, that you can – so you can measure the change in volumes. And just – this is from way back in the AN1792 trial, you can see here, this is the baseline, this is at the end of 12 months and the difference is very precisely measureable and you can follow this – the difference here between the first scan and the 12 month scan is all in red, and you can quantitate that very effectively.

So just to give a little bit of the background of this AN1792, the first sort of large, expensive, risky Phase II study that I'll be talking about, this was the study in mild to moderate Alzheimer's disease patients with this vaccine candidate, and this is the trial that was stopped because six percent of these patients developed encephalitis. And what we did with this study is we planned on –

we enrolled about four times as many patients on drug than on placebo, because we knew that AN1792 only was immunogenic in a relatively low percentage of patients. And we had planned multiple dosing sessions, but we stopped after usually the second or the third dose in patients when we saw the encephalitis, and then we analyzed patients at month 12.

And basically, the standard psychometric tests were all negative in this, so this is another trial in which the primary outcome measures were negative, but we did do a test – our own attempt, really, to create a more sensitive test and a test that was more responsive to the changes that you see in early AD. We created this test called the NTB, and what we saw there is that the patients who responded with antibody to AN1792 actually stayed about the same and the patients on placebo declined. And the other interesting thing that – is that we had a dose response in that effect, and that is, you know, all the patients got the same dose of AN1792, but they responded with antibody titers very differently. And you can see here the effect on the NTB was highly correlated with the antibody response.

So what happened with volumetric MRI, this test that clearly tracks the clinical status of the patients, well, as you can see here, there was a profound effect on volumetric MRI. Highly statistically significant, but as many people know, this was in the opposite direction of what was expected. So patients here on placebo, their change in MRI volume over the year that we followed them was about 2 percent and the antibody responders was about 3 percent. And this also was highly statistically significantly correlated with anti – with the titer in these patients. So that was our – sort of a first attempt to use biomarkers to make a decision in a relatively small trial, and in fact, we got the clinical and the biomarker effects going in the opposite direction.

So just a couple of – we did a lot of work to try to explain this effect, and in particular, we collaborated with Nick Fox at the National Institute of Neurologic Diseases in London, and Rachael Scahill, who worked in – who is working still in Nick Fox's laboratory. And what we wanted to do is we wanted to find out what parts of the brain were responsible for this sort of paradoxical change in the MRI brain volume, and first Rachael looked at white matter and she went through all of these scans, traced out some white matter structures such as the pons here and found that the white matter did not change at all. Basically, it became the placebo in the responders. So didn't seem to be an effect on white matter.

And then as they do so well in the laboratory, they did voxel-based morphometry and statistical parametric mapping, and this is in particular to show the regions where, first of all, the placebo – the placebo brain loss was greater than the responders here. Basically nothing came up on this statistical parametric mapping. But in terms of the response, loss of brain greater than placebo, these are the regions that seem to be responsible for this – for this paradoxical effect, and in fact, this map here includes a lot of the regions of the brain that tend to accumulate amyloid and accumulate amyloid early, such as the hippocampus. So just a few hints into sort of why maybe we saw this paradoxical effect. And in fact, you know, there's been a number of – you know, this study was started in 2001 and finished in 2002, and in fact, there's been a number of patients who have come to autopsy, and those autopsies have consistently shown a reduction in plaque in antibody responders.

So what about bapineuzumab, that antibody that, again, has a similar effect in mice to the AN1792 immunization. Well, this is the volumetric results – volumetric MRI results from bapineuzumab Phase II, which was that ascending dose 18 month trial of bapineuzumab versus placebo in patients with mild to moderate Alzheimer's disease. Just to go over the clinical here, the clinical results of the bapineuzumab Phase II were probably – you know, it's a discussion in and of itself, but just to summarize, what we saw was a difference here in these post-hoc subgroups.

And again, the post-hoc subgroups with bapineuzumab were maybe particularly suggested by the side effects, which were side effects of a condition we're calling vasogenic edema, for want of a better term, and that side effect was different in the ApoE genotype groups. The non-carriers got substantially less vasogenic edema than the carriers in that study, and in fact, the non-carriers had a different effect of bapineuzumab here in another analysis. Again, that's entirely post-doc.

You can see here that the placebo patients declined far more than the bapineuzumab patients in the ApoE for non-carriers. But the ApoE for a carrier – carriers, there was little difference between the placebo and the bapineuzumab patients. And that was actually reflected in the brain volume results, so that the ApoE for non-carriers had a statistically significant decrease in the decline in their brain volume compared to the placebo patients; whereas the ApoE for carriers had no difference at all. And again, is this reflective of the clinical change or, again, because this is a post-

doc analysis, you don't know whether this might be due to just sample bias. So again, this awaits confirmation yes or no in our larger trials.

So now some of the other fluid – some of the other biomarkers included the fluid biomarkers. Well, the plasma A beta, as you can see – and this is actually a slide from a gamma-secretase inhibitor which has been under development at Wyeth and now Pfizer. You can see here you've got this initial decrease after a dose of this gamma-secretase inhibitor, and then an elevation.

Well, what we see with antibodies in general is we just see a massive elevation. This is about 30 or more fold, and basically plasma A beta has a very short half life, the half life is altered when the antibody binds to plasma A beta in the plasma, essentially, and with the binding of the antibody, the A beta in the plasma takes on the half life of the antibody and you naturally see a substantial increase in the plasma A beta content with time. And you can see here it closely reflects the time course of the antibody in the subjects.

So what about CSF A beta and tau? Well, we've never seen a difference in CSF A beta in patients in these immunization trial, but the results for tau are interesting. With AN1792, back in 2001 when we started this trial, we didn't have good measurements for phospho-tau, so we just measured total tau, and you can see here we got CSF measurements only in a small subgroup of the patients in that trial, but we saw a nearly statistically significant difference between the – or actually that is statistically significant, I'm sorry – between the responders and the placebo patients; whereas the patients who were antibody responders to AN1792 had a decrease in their CSF tau. Remember, CSF tau is elevated in patients with Alzheimer's disease, so this is a relative normalization in that.

In bapineuzumab, we looked primarily at phospho-tau, and you can see here we did two studies. I'll talk about study 202 in a little bit. This is a small study of patients where we used a PET tracer, and both of these study 201, the large – the sort of large or really mid-size study of bapineuzumab over 18 months, and the smaller study 202 showed the same trends in phospho-tau. When you combine them together in data that was presented at the ICAD last summer, you can see here that the pooled analysis shows a nice, sort of robust statistically significant difference between bapineuzumab and placebo for the effect on CSF phospho-tau.

So, finally, the last biomarker that we've sort of taken from these longitudinal studies and applied to our Phase II trials is positron emission tomography. And we looked both in this small study of about 30 patients in three dose groups of bapineuzumab, an 18 month study in mild to moderate Alzheimer's disease that we did in the UK and in Finland, in parallel with the larger Phase II study. In that study, we found actually that, well, here at last is a study that was positive on its pre-specified primary outcome measure, where the bapineuzumab patients, their amyloid content by PiB is being measured here, and the bapineuzumab patients have a decrease in their amyloid content, here measured in SUV. The placebo patients have an increase in their amyloid content, a treatment difference of about 0.24 and a highly statistically significant p-value.

And just to give you an idea of what that represents, the standard uptake values which this study is measuring, basically measure the ratio of the amyloid content in the brain regions of interest. In the cerebellum, a value of one means essentially that there's no amyloid. A value of two means that there's twice the signal. Alzheimer's disease patients in these regions of interest range around between two and three, and so a difference of about 0.24 is anywhere from 25 percent to a little bit less of the amyloid content in these brains, and that's the magnitude of the difference between patients who have been treated for 18 months with bapineuzumab over patients who have been treated with placebo. And again, this – the final analysis of this was based on 19 patients in the bapi group and seven patients in the placebo group, so a very small study. No time consuming and risky.

What did we see on FDG? We didn't see any difference on the FDG. And just to give you an idea what this looks like, this is the estimated change in baseline over time. The placebo patients going up, the bapineuzumab patients going down a little bit, and a statistically significant difference at the end, and the PET people always like to see the pictures, so this is also published by the principal investigator in Lancet neurology. You can see here this is sort of above average patient and a below average patient for both the bapineuzumab and the placebo patients, and you can actually see the difference, really, at the end of 78 weeks, in the amyloid content here decreasing in these bapineuzumab patients and increasing in these placebo patients.

So just to conclude, so what are our findings, in terms of the evidence for disease modification? Well, you know, the brain volumes may correlate best with the clinical progression, but the

drug effects are hard to predict, clearly, in our experience. Plasma A beta changes for monocle antibodies reflect target engagement in the periphery, but in fact, they don't really predict, we think, what the drug is going to do in the CNS.

We like to think the tau and phospho-tau levels in the CSF. We use them and they may reflect a reduction in neurodegeneration, since in fact tau is the – tau and phospho-tau are the products of neurodegeneration. Amyloid imaging by PET, I think we've documented – that combined, you know, with some of the extensive work with these amyloid tracers, showing that they really reflect the pathology in the brain of Alzheimer patients.

We think that we've documented the clearance of fibrillar A beta by immunotherapy, but certainly the results of our definitive clinical studies are necessary before we really have put all these interesting biomarker findings in perspective. And basically a lot of acknowledgements, a lot of people who've worked on these, certainly, you know, more than 400 clinical investigators and experts and a lot of patients who've dedicated their time to it.

[End of Audio]

Dan Perry: We're going to go right to the first of our two panel discussions. I'd like to ask Dr. Doody to come back. She will be moderating the first of the two panel discussions, which will focus on the design of Phase II clinical trials, and she will be joined by Dr. Martin Farlow of Indiana University School of Medicine; Dr. Howard Feldman of Bristol-Myers Squibb, Dr. Anthony Gamst of University of California at San Diego; Dr. Katz who, as we all know, is director of the neurology division here at FDA; and Dr. Rachel Schindler of Pfizer. I'll turn it over to Dr. Doody.

Dr. Doody: Thank you. Well, thank you, panel. You've got a big job ahead of you, trying to refine and synthesize this information, and basically put forth your ideas about Phase II clinical trials. We have a guideline list of questions and I need to ask, Dan – do you still want us to finish on time?

Dan Perry: No. We'll make it up by brief closing remarks.

Dr. Doody: Okay. So what time do you want us to end?

Dan Perry: Take an hour.

Dr. Doody: Okay. So to begin with, what we have basically, we want to cover issues related to the design of trials; the outcome measures; what's a clinically meaningful effect; some sort of consensus about biomarkers, which may be a little bit premature because we haven't heard from the biomarker panel yet; and some aspects of statistical analysis. What we don't want this to be is a whole bunch of questions to Rusty, because somehow it seems that in many instances what we do is ask you the same questions over and over again; I don't think that's terribly productive.

First of all, let me just open it up and be a little bit undirected here, and ask you as the panelists to start off by stating what you think are some of the big issues in Phase II clinical trials. Where do you think the problems are, or where do you want it to go, or what do you think we need to cover in this discussion? Marty?

Dr. Martin Farlow: I think the preceding presentations are very instructive, and particularly the data that you presented, Rachelle, with regards to this. There is this tendency in early Phase II trials to want to escalate these studies into mega-trials that have 500, 600, 700 patients that are going for a year or a year and a half. I think the look or the question is wrong. There's such an urgency in finding drugs, and there are groups of people in major corporations and small biotechs that are organized, that have enthusiasms for these drugs, and they tend to champion this drug, as it were, and in the process of these trials they tend to grow and they tend to have these biomarker assessments attached.

The problem you get into is they won't take no as an answer. They need to essentially recognize that there is a very high probability that any given drug or agent is going to be negative, and I think what is needed or what is desirable – and I take as the purpose of this conference – is as much finding a drug that treats Alzheimer's disease, of establishing guidelines or rules or understanding how to more efficiently determine that a drug has very little probability of being useful for treating Alzheimer's disease, with as little expenditure of resources and time – do it is quickly as possible – and then be willing to take no for an answer.

From that standpoint, I think obviously you need, be it a clinical measure or a biomarker measure, you need something that will change in a patient with Alzheimer's disease over a set period of time, reliably – say 90 percent or 95 percent of the time, over six months, you'll see a change in that measure and of a placebo population, and you have confidence that that will occur. If you get 100 groups of patients together, you're going to see that

change. You then need to ask the question, I think, "Is there a signal?" – and it can be a clinical signal.

If a company or a group of people wants to ask the question, or believes very strongly that there is a process going on – let's say, for example, deposition of amyloid in the brain – that they need to ask that question, that single question, and they need to then do the study and determine with the agent that they have employed has the action that they desire. If there is a signal there, then I think it's reasonable to go forward into these big trials. And if there is not a signal, they need to be willing to take no for an answer, and they need to abandon that particular agent or drug, and they need to move on and expend their resources elsewhere.

Dr. Doody: Let me ask you one question before we move to the next person. When you say "take no for an answer," let's take the time clock out of it, the patents as time clock, okay? Let's say you get no for an answer about this drug in this population at this point in time with this trial design. Does that mean the drug is thrown away and you go on to something else, or might you pull that drug out later when another similar compound has had a different set of decisions related to its Phase II development program? Are we too constrained by the time clock?

Dr. Farlow: I think that's possibly right, and there's a huge danger in employing biomarkers in these populations, particularly genes, in dividing subpopulations. What happens is instead of asking the one question that you have the power in your given study to ask, even though everybody says we're asking this one question, the subgroups are looked at in these early Phase studies afterward, and with a power that really isn't there, decisions are made to go forward with studies that involve 500 or 1,000 patients, because people really believe in this particular line of research in terms of amyloid or tau or whatever it is, and they really haven't answered the question. There's sort of a mission creep, as it were, in terms of getting away from the fact that you really have limited power with a small number of patients, to answer a question.

Dr. Doody: I'm sure we'll come back to that. Howard?

Dr. Feldman: So, Rachelle, I take the view that, like all good research, Phase II should begin with the question of what are the goals and what are the aims of the program. What do you hope to accomplish in Phase II? As Marty has indicated, and I think as everyone who has designed a Phase II program understands, you can't do everything in Phase II; you cannot answer all questions. And so I would boil

this down to the things you absolutely must know in order to even consider Phase III. So these are the things you must learn in Phase II.

You're coming into Phase II with only a very small amount of information in either healthy normals or in patients with the disease, but the exposures are vanishingly small, from maybe 5 to 10 persons at a dose, maybe an exposure of a week, two weeks, a month; that's all you know. By the end of Phase II you have to understand what the potential toxicities are, and whether they can be put in front of you to manage in clinic. So you've got to know what they are. You may not know the full extent of them, and one of the points you have to come to terms with is not every toxicity will be fully explored in Phase II, and you have to carry that risk into Phase III.

You have to understand that when you're talking about disease modification, you're talking about safety intolerability with a prolonged exposure, and so then the question is, how do you balance, buying down the risk, when you're going to be using a drug for many years within a Phase II profile, where you're going to have a shorter snapshot in time? So you must establish safety intolerability. You must establish dosing. In many ways, it's imperative that you understand what your effective dose is, and then you get to the question of dose range.

There are probably two models worth considering for that development question. One is, should we be using the oncology model of drug development, or should we be using a primary care model of drug development? In oncology, I don't think the oncologist would ever aim for a minimally effective or non-effective dose. They would say, "We want a dose that works. Show us a dose that works and we'll work backwards from there to lower doses." But in neuroscience we may take a different view. We may take a more primary care kind of model, saying, "We really need to know what the non-effective dose is. We need to try and limit the exposures," and so on. So to some extent, a program needs to decide, are we going after the oncology model or are we going after a primary care model insofar as dosing goes?

I think the other thing that's very important and often not well-reflected in a lot of the discussions, even this morning, are PK/PD relationships. We undervalue the importance of things like, is there a dose-response relationship? Or, is there a really wonky dose-response relationship, where we're looking at something that looks unusual? Are we trying to rationalize a U-shaped dose-

response curve? Are we trying to rationalize something that doesn't have an intuitive biological plausibility about it, and I think that's very important as a goal for Phase II, and needs to be a lot more emphasizes.

And then the 800-pound gorilla is the question of efficacy, and how much efficacy do you need, and what type of efficacy do you need? In my view, I think that you need a relevant biomarker to move on a disease-modifying therapy, and I think if you look at that list of Phase II trials that you have, and you ask the question of themes, I think one of the themes is target-engagement biomarker activity, although I did note that in the tramiprosate you did have an apparent biomarker signal, the right biomarker signal, and it didn't buy down much risk; you still had a failed Phase III. So you say it's imperfect but at least it speaks to a process which gives you the grounding to support larger investment.

But invariably, when you come to efficacy, you start to talk about larger samples and then you start to get into this more difficult question of, how far do you want to go showing efficacy in Phase II, or is that really in Phase III, and is the biomarker enough to leapfrog into Phase III?

Dr. Doody: Howard, when you say the biomarker, do you really think there can be one? In other words, let me just stop you. The tramiprosate example is a good one. If you expect it in the CSF, you get it in the plasma – does there have to be a package?

Dr. Feldman: No. Thank you. That's a great point, and it gets to really a critical aspect in development, which is your decision tree, and what you have to have at the end of Phase II is a decision tree and you have to understand how to put a decision tree together, and are you going to bank the whole thing on one dimension? Is it going to be safety? Is it going to be biomarker? Is it going to be MRI? What are you going to do?

And as someone who has been confronted with that in the last year or two, you need a lot of wisdom to construct the right decision tree, and then to Marty's point, you have to be prepared to respect your decision tree. I can tell you, it's easy in hindsight to look at your decision tree. It's not so easy, prospectively, to say, "We will live and die by these decisions, but it can be done and should be done."

Dr. Doody: Thank you. Anthony?

Dr. Gamst: I'd like to start off by saying that I strongly agree with the outline offered by my colleague, Dr. Feldman. I think the goals in Phase II are crucially important, and one separation that you can make is whether you're interested in showing some effect on a presumed mechanism of action, or whether you're interested in predicting what the results of a larger Phase III trial might be. If the former, then I think the outline offered is perfectly reasonable, and if the latter, I think one of the crucial questions is, can we do better with outcome measures, either to understand essentially clinical effects of the drug, whether there is any actual observable clinical effect in Phase II? As far as shorter trials are concerned, and just study mechanism of action, I think the biomarker trials are more reasonable.

And then the question is can you cheaply and reliably identify populations of interest? So, I mean, if you have an Abeta agent, can you identify a population subject in whom that's likely to work, and can you do that cheaply? So I think the major questions really revolve around identification of the appropriate population and refinement or understanding of outcome issues, and in particular, the disconnect between a potential mechanism of action and a potential clinical outcome. I don't think we understand the biomarkers well enough to know which way they should respond in response to a drug, and even if the drug does have an effect on relevant biomarkers, whether it will then have an actual clinical effect, whether the biomarkers are related appropriately to clinical outcomes.

Dr. Doody: So, Anthony, we've heard that it might just be all about the statistics – that there is a clinical effect, that we just didn't see it. How much do we have to delve in to the outcome measures themselves before we accept no for an answer? Can you tell that by a quick look at the Phase II data?

Dr. Gamst: In short, I don't think so. If you believe strongly in a mechanism of action, a very particular mechanism of action that you believe will have a clinical effect, then I think you can accept no for an answer as soon as you see that the drug doesn't seem to be working as expected. The danger in that is that you may be studying the wrong population, the wrong stage of the disease. You may also not have the biological effect that you expect, but in the end you would have had a clinical effect if you'd gone and studied…

Dr. Doody: But the question is, post hoc – you've done your Phase II trial, you didn't get a clinical effect. Is there a way to look at the data and

say, "Well, we didn't analyze the ADAS-cog right"? Or not? Does that have to be pre-specified?

Dr. Gamst: I don't think so. I think it has to be pre-specified. There's an extreme danger with post hoc analyses. So if you can find a group of subjects in whom the treatment appears to have worked, and the trial is null, there must be an equally sized group in whom the treatment extremely didn't work. Snooping through the data is an extremely dangerous practice. Having said that, I think that a lot of work should be put into improving clinical outcome measures so that you can see a clinical effect, either over a shorter period of time or in different populations, for example, early or preclinical AD.

Dr. Doody: Okay. Thank you. Russ?

Dr. Katz: Yeah. I agree with a lot of what's been said. I think from where we sit, and I think from what we heard this morning, you can't help but be impressed by the fact that Phase II seems to be just as large as Phase III, the studies are just as long – in some cases, if not longer – almost just as large, and I think we're sort of, potentially anyway, losing sight of ways to do it differently. I think we really need to start thinking about, how do we get to shorter, smaller Phase II-type studies? I think you can't learn everything, but I think you can learn useful things: target engagement; I think there's a place for short studies that just look at target engagement and maybe dose response for that sort of thing.

I was impressed by one of the slides that Ron showed – I guess it was bapineuzumab and the time course of PiB-positive scans. There was statistical significance, I guess, by Week 72, or whatever the end of the study was, but you started to see differences at Week 20. They weren't statistically significant, and I'm not sure that Phase II is the place to look for statistical significance. I know we sort of live and die by that throughout the development program. It's worth thinking about whether that's a useful benchmark to decide whether or not to go into Phase III from Phase II.

I think maybe we should be thinking about in-and-out, quickly and some of the value of what Dr. Farlow was saying, about just get in, get an answer that there's nothing going on early, and maybe if you did a longer study, something would be going on with that agent. Or, as you suggest, maybe down the road when we learn more, maybe that agent will be positive. But given what we know

now, it's not a bad thing to make your rejection decisions much earlier. You may lose some useful agents but I think the history is that that's a better thing to do.

I'd say we really need to think about how to do these things shorter and get useful information, so I think target engagement, when you really think you know something about mechanism – I'm not the big one for mechanism, at least when we get to the point of approval – but in early development. If you think it's an amyloid drug, you ought to see if it's engaging the appropriate target. The question of post hoc – we've seen many ____, and I would say one of the biggest problems we see, or in the sense of forging ahead with a treatment that ultimately turns out to be negative, is this tremendous optimism at looking at a study, at a Phase II study that's just negative.

It's been mentioned and I think it's true – you can always slice the data and find some subset in whom it works – and maybe that's true. Maybe with that particular agent it is ApoE4-negative people; that's possible. But I think, again, the history is – not that the history is vast – but I think the experience is that usually turns out to be a problem. So post hoc optimism, I'll call it, is often, I think, just turns out not to work out. On the other hand, Phase II is the place where you're supposed to just sort of cull through the data and see what you've got, but maybe that can be done in shorter, quicker studies. So maybe post hoc learning – maybe that's not even post hoc – but can sort of adapt a learning phase of Phase II into something that's shorter and smaller. Maybe.

Anyway, I think it's worth thinking about. It's worth thinking about, I suppose, adaptive designs – whatever that means. That means different things to different people, of course, and the particulars and the specifics are very, very important, but I think that may be a way to go. And safety. And again, I'm throwing these things out. These aren't official positions – let me get that on the table. One of the goals of Phase II, presumably, is to get safety information and that sort of thing. You know, some of that stuff might be incorporated, if done very carefully, in later stage studies.

In other words, if Phase II can be shorter from the point of view of getting all the safety data that we typically sort of get in Phase II now, maybe shift some of that later, and if you think you got something from a well-designed, small, short Phase II study. You know, with very, very detailed, real-time surveillance centrally on the part of a company in a sort of a Phase III-ish study – you know, before you have a lot of safety data from Phase II – where

somebody is looking frequently at safety data that's emerging, maybe that could take the place of extensive Phase II. Maybe you could get your safety information from there.

So I just think these are things worth thinking about, but I think where we've been – and I think the experience bears this out – that Phase II is long, it's expensive, it's under-powered. A long, big, expensive, under-powered study. At the end, what do you have? So I think we need to be just thinking very differently about this. It may turn out that this is the only way to do it and get useful information, but my suspicion is that we could do it quite differently.

Dr. Doody: But the trend has been, if we go shorter and quicker, biomarker instead of clinical. Do you think that we have to make that distinction? We're doing short, quick biomarker or long, drawn-out clinical?

Dr. Katz: Yeah. Again, I think, if nothing else, I suppose if you take a short, quick, let's say biomarker-based Phase II study – if you can even do biomarker studies with smaller, quicker studies – but if you can and if you go into Phase III and it basically turns out to be negative, at least you saved the extra year and a half that you would've been spending on your Phase II, which half the time leads you to a big, large Phase III study anyway, based on the underpowered results. I mean, there may be a total savings in time, in some sense, if you think of it that way.

Again, this is obviously not typically how it's done, but I think it's worth thinking about alternate ways to do this, because I think what we're doing now is not ideal. And again, more sensitive clinical measures. We heard from Dr. Hobart on that. I think that's incredibly exciting, at least in terms of signal generation. It may be that sort of thing, in short, smaller Phase II studies you'll get a trend, and even in a clinical outcome. So I think we can start, we should be starting to think about doing things differently.

Dr. Doody: Okay. Rachel. And as we transition to you, maybe somewhere in your remarks, can you work in the idea that you tell upper management, "You know, we're not going to really look for that clinical efficacy in Phase II. We're going to hold that off for Phase III and do some other things in Phase III." Your remarks, please.

Dr. Schindler: It's always interesting to be the last person to comment. Sometimes it's the easiest and sometimes it's the hardest. First, I agree with most of what's been said. It's kind of interesting

because before I joined industry I had a perspective that, I would say, was much more of a clinical perspective, and thought more about things, some of the questions that have been raised by some of the other colleagues on the panel, about things like, are we studying the right population, and is it the outcome measures, and what's wrong with what we're doing? Since I've become an official drug developer, I've been tasked with trying to do it in a much more rigorous way, and I would say of all the comments so far, that Russ actually touched upon a lot of that greatly.

First of all – well, let me actually go back. For those of you that don't know, I work at Pfizer which is, as you know, a very, very large company, and for someone who is dedicated to Alzheimer's research, given all of the failures that the field has had recently, I get worried at such a large company that they are simply going to say, "You know what? This is too hard. We're going to go elsewhere. Let's put our money into cardiovascular, or something else." So really, there's a lot of interest in making trials more efficient.

One of the first principles – to extend what Russ said – is to have clear proof of mechanism before we move into a proof of concept study. Now, sometimes that's not really possible but sometimes it is, and I'll use a terrific example. Today, in Ron's data, the bapineuzumab data that Ron showed, that Russ just referred to, it's beautiful – interestingly obtained in Phase III – but it's beautiful evidence of target engagement. And that's also one of the principles that has been stressed to us, in my company, about using all available information, and bringing it all to bear on your design. For example, we now know that you can demonstrate target engagement with PiB, with CSF measures, so that would be a very critical thing to do in Phase II, before marching on into a proof of concept study.

I can't remember who commented on this, but we have been less and less focused on statistical power. Statistical power is very important, but there's often this comment that, "Well, the study was too small. We didn't have enough power so we really can't say." We've been really asked to look at that differently and we have been. We've been looking at it more in terms of the probability of technical success, and the probability specifically of making a correct go/no-go decision. So in other words, not just statistical significance but based on the truth, if you knew the truth, of whether your results were real or not, or I should say the right results, what is the probability of making a correct decision? What

is the probability of making a false go/no-go decision? And what we've done is we've adjusted those confidence limits.

In different studies we've loosened them up and we've made them more stringent. In a case where we really don't think we have enough information, because we're going into uncharted territory, we make the hurdle relatively easy to pass over. But in any event, we create decision criteria, and we are bound to stick to those decision criteria. We cannot go back and say, "Oh, well, the study didn't show something for it because of such-and-such and, in fact, the reason we didn't hit the confidence intervals." We are basically bound to defining these probabilities of success, defining what would be advancement criteria for a go/no-go decision and sticking to them.

I chair something called the Technical Review Committee, and I'm responsible for actually enforcing those, helping create those decision criteria and then enforcing them and not going back. And some of that – how do you figure that out? You want to identify efficiencies in study design to maximize the probability of getting to a correct decision in that go/no-go, and that means looking at the size, things that have been mentioned so far, the number of doses. One of the things that's not been mentioned is the terrific use of interim analyses, and very carefully planned frequent interim analyses, which everyone always says, "Oh, you pay a penalty for looking." But if they're carefully planned and done in the right way, they can be extremely useful for stopping a trial before you get too far down the road into something that is not going to be fruitful.

We also look very carefully at the number of patients per group and the number of doses. If we're concerned about the operating characteristics of the design and making that proper go/no-go, or making the best go/no-go decision, we want to look at our data analytic methods, and again, those limits that we're going to allow, those confidence limits. And lastly, I would say, again, what I had mentioned earlier, the power and importance of using all currently available knowledge. One of the things that we've started to do frequently is modeling, and when I first heard about modeling I thought, "That's really kind of typical statistical stuff and that's not looking at the patients. We treat patients; we don't treat numbers. We don't treat laboratory results."

And I've actually started to really understand the power of modeling, and the importance of running simulations. Again, you're not always going to get the right answer. You may kill a

certain number of drugs that may have been good drugs for the wrong reasons, and you may advance a certain number of drugs that you shouldn't have advanced. But the modeling can really help you in guiding those decisions. So, actually I said that that was the last thing; there is actually one really last thing. That was the second to last thing.

A lot of the people that I've learned from outside of the field of Alzheimer's disease, in terms of drug development, have said that good drugs declare themselves early, and that you can tell relatively early. Somebody mentioned, for example – not that this means it's a good drug, and it doesn't mean it's a bad drug – but I think Russ, it was you who mentioned the data that Ron showed on bapineuzumab, that very early on, even though it wasn't statistically significant, there was already separation; there was already a signal. So putting all of those things together –

Dr. Doody: On a biomarker.

Dr. Schindler: On a biomarker. Right. On a biomarker. Yes.

Dr. Doody: On a really good biomarker.

Dr. Schindler: Right. On a really good biomarker. But in other words, there are signals of things, and that would be one thing to be taken into the proof of mechanism phase.

Dr. Doody: Right. Okay. So these are great comments, and now we want to hone them with respect to certain issues. There are three issues that I want to try to get discussed in this panel: one related to populations, one related to endpoints and clinical meaningfulness, and one related to biomarkers. We're obviously going to stray between those different designations in any discussion. So let's just talk a little bit about populations. There are strong beliefs that certain drugs should be tested in very, very early or asymptomatic or pre-symptomatic patients.

If we do that – there are even, within these questions that were given to us as guidelines, there is even an implication that maybe we should be redefining everybody with new guidelines, and not just for the purposes of clinical trials. But if we shift around and try to utilize different populations, do we feel that there are some downsides to this, or is it only upsides? Who wants to advocate for moving drug development into pre-symptomatic patients believed to have Alzheimer's disease? And if you do make your remarks about that, would this be mechanism-specific? Marty?

Dr. Farlow: Yep. I always get to be first here because I'm sitting next to you. You know, I would advocate for that but mechanism-specific only, and the mechanism in question obviously would be Abeta, and deposition, and it would relate to, and sort of in an emerging consensus, that the amount of damage done in moderate to severe, or even mild stage Alzheimer's disease, may already be too great that the various processes going on are sort of catalyzing themselves, and if you take amyloid now or not it's, at that stage, not going to make a heck of a lot of difference.

Dr. Doody: But would you hold your tau immunotherapy to that? Your transmitter drug? Your neuroprotective drug?

Dr. Farlow: You know, I'm not sure that I would, and there's a reason. The reason is that there's been so much effort, and particularly with regard to adding in the development of the imaging methods, and in terms of the various studies that have been done with the biomarkers, there's the beginnings of some understanding of how amyloid is deposited in very early stages, and some suggestion of an association of "pathological amyloid"; there is a cognitive penalty in supposedly normal, elderly individuals, so there is that association there.

So I think in terms of therapeutic equipoise or whatever, I mean, if there is a cognitive penalty, I think there is a rationale to treat with an agent. Obviously, an agent that has tremendous negative potential, there's still going to be questions about damaging people versus the benefits that might be gained. But if the safety studies look very safe, I think it's reasonable to approach and to ask the question at that stage. Just as you take cholesterol out when you're approving cholesterol drugs, you're looking at cholesterol. You're not necessarily directly asking a question about heart attack or stroke.

Dr. Doody: So for amyloid drugs, early, preclinical, good idea. How many people on the panel, in addition to Marty, believe that there is a cognitive penalty to amyloid? Just a show of hands.

Dr. Katz: Let me ask you this, to clarify that. Meaning what? Meaning people who appear to be entirely asymptomatic really actually do have a cognitive deficit?

Dr. Farlow: Well, no. What I mean by that is if you have a population of people and you give them psychometric tests and they are

demonstrated to have amyloid in their brain by PiB, they will have scores that are definably lower.

Dr. Doody: The change scores. I think the one paper that we have to go on would say that the change scores, the annual change scores within the normal range, might be different in those who have an amyloid burden in the brain. So my question is – because a large premise of what you said there was that there's a penalty – how many people on the panel believe that there has been demonstrated, that you're comfortable with the idea that there is a penalty in cognition to having amyloid build-up? And I'll make this a little easier by saying abnormal amyloid build-up. Okay? We get a technique that tells us it's abnormal.

Dr. Feldman: So Rachelle, I would just respond to that – I don't have a full answer for it – but I would share an observation that has troubled me for many years, which is that when you look at autopsy and you go and look at brains, and the neuropathologist does not tell you the clinical history, in a not-infrequent percentage of normal autopsies, maybe 20 percent, 25 percent can be higher depending on the series. There's lots of pathology. There's almost enough pathology that you would diagnose the disease.

Dr. Doody: And you're talking about cord plaques?

Dr. Feldman: Cord plaques, and then the pathologist says, "Well, that's fine. There's nothing wrong with this person," and they were tested, and they may be in the study.

Dr. Doody: Well, this is the reason why I'm asking the question.

Dr. Feldman: Right? So that exists. It does exist. But it's also true that on average, people migrate in groups less well when they've got a [inaudible], so both things are true. I think the thing that troubles me is that we're missing a piece of the understanding. You touched on it this morning. You said, "It's necessary but not sufficient." In some people, having a lot of amyloid I suspect is going to be not only sufficient; it will cause the thing. I think in other people, they're going to have a different set of circumstances, and I'm afraid that we're missing some understanding of the intervening piece that drives the amyloidopathy, either to a tauopathy or to an inflammatory disorder, or to a mitochondrial – whatever it is.

Dr. Doody: But do you accept that there's a cognitive penalty, whether or not it derives to an Alzheimer's case? I'm going to let you qualify in a minute, but you got this started.

Dr. Feldman: Well, it's not a hill to die on for me.

Dr. Doody: It's not a hill to die on. Before we get back to Marty, is there anybody who feels strongly either way? What most of you are saying is what I would expect. You're kind of examining both sides of the issue – Is there or isn't there a cognitive penalty? Are there circumstances where there is? Can we identify those people for whom there is? Do we have the right biomarker? But is there anyone who feels strongly that if you're building amyloid and that you can pick it up on PiB or AV-45, or whatever the tracer is, you've got a penalty there and therefore you've defined a class of people in whom the risk-benefit ratio has changed?

Dr. Feldman: You need to ask Reesa.

Dr. Doody: Reesa. Well, she'll get to – do you want to speak now?

Dr. Feldman: No. She should speak, too.

Dr. Gamst: I think there's reasonable evidence of a correlation at the population level between people with amyloid deposition and cognitive decline – rates of conversion, rates of decline. One of the things that I think we're missing that's relatively crucial is, suppose you were able to find these people and reduce their amyloid deposition. Would they then not experience the cognitive decline?

Dr. Doody: Yeah, but that's the clinical trials question, so we're okay with that. Okay, Marty, go ahead.

Dr. Farlow: Well, that was actually the point I was going to make. I'm sort of being painted in this firm amyloid, and it very well may be that the amyloid that's being deposited at least is spackle, more or less. It's sort of being painted around to fix something else that's going on. But I think at this very early stage, where you're just beginning to see maybe there's a difference, it has a clinical effect at earliest stage, I think that's a reasonable place to ask the question: Does it make a difference, if you have a drug, a methodology that will take the amyloid out, is that going to make a difference for these people?

Dr. Doody: Right. So is there anybody who just disagrees? You don't think, in an amyloid drug you need to move to an early population? I don't expect controversy about that. Okay. So we've defined that. Now let's move it a little bit differently. One of the things that we are doing is by specifying a population we are specifying a degree of risk that is allowable in interventions. If we move to a population who clearly has a cognitive problem – they have episodic memory disorder, with or without amyloid; we've moved to a different person.

Most of us think that person is on the pathway to something not good. How does that change things? Do we think that that is a population of interest for any and all anti-Alzheimer's therapies? Or is that a better group for anti-amyloid therapies because now we really understand the risk-benefit a little bit differently? How does that change things, to move to a group that has an episodic memory disorder? Rachel?

Dr. Schindler: I think it's the weight of the evidence. I think that population does trouble me. We all know that there are some that don't theoretically progress, but again, I think it gets back to the weight of the evidence. That in addition to, say, and sort of along the Dubois criteria – I'm not advocating that per se – but the idea of the weight of the evidence, so having a clinical marker as well as a potential biomarker. You know, just because you – well, let me back up. Just because you have very bad arteries does not mean that you're going to suffer an MI, and what causes clinical symptoms is often external factors, and those are often related to the threshold of what the biological changes are.

So for example, we define dementia by loss of function, and how do we define loss of function? So for somebody who doesn't really do very much, and whose system isn't stressed very much, we may not see a lot of loss of function. In somebody with significantly narrowed arteries who does not physically stress themselves, we may never really see symptoms of chest pain or symptoms of ischemia before an MI. And I don't think this is necessarily that different. I think that's a reasonable framework to start to think about this. So yes, episodic memory concerns me tremendously, but I think I would want to see other signs, much in the way that an amyloid load would not be enough, although the episodic memory would be much more concerning to me.

Dr. Doody: Well, what if they've got episodic memory and nothing else? Do they have to have a biomarker?

Dr. Schindler: So the biomarkers that we have now, it's not even that they're imperfect, which they are, but they tend to look at only a few things, and most of them look at amyloid. So that doesn't say a whole lot about other markers of disease, and there are also questions even about the amyloid and fibrillar, and the whole issue of what type of amyloid we're looking at. If it's just episodic memory from a clinical basis, based on what we know today, I would not feel comfortable with a risk-benefit ratio of exposing that individual to a drug that could be somewhat toxic.

Dr. Doody: Marty?

Dr. Farlow: Do you want me or Howard to go?

Female: I'm sorry. Howard?

Dr. Feldman: I would just say that we tried that; we did a generation ago. We tried this thing called MCI trials. We had lots of them. We anticipated that we would achieve consistently high rates of progression to the dementia diagnosis of AD. I don't think we succeeded in the enriched approach that we hoped. We achieved sometimes 4 to 5 percent progression per annum, instead of the 15 percent. And that was a clinically, phenotypically derived population. So I don't think that works, and I think there's emerging consistent evidence that applying a biomarker on top of that clinical phenotype will take you to a much higher rate of progression and, in fact, is, as you said this morning, identifying the right patients.

Dr. Doody: Okay. So Howard. If you have MCI and an FDG-PET that has posterior parietal temporal, posterior cingulate involvement, are you the same as a person who has MCI, low CSF, Abeta, increased CSF tau?

Fr. Feldman: I think those biomarkers are telling us different things, and I don't know that we know enough yet to be unambiguously clear on that point, so to some extent the decision may be based around what your therapy is meant to do.

Dr. Doody: Okay. I agree with you, but I don't think that's how these things are being operationalized.

Dr. Feldman: …how they differ, we're learning. We know, for example, from MRI, that there's less specificity as you get older. So if you were applying an MRI biomarker in an elderly population, you'd make

more mistakes than if you apply it in a younger population. So there are issues.

Dr. Doody: So maybe our phase 2 trials need to select a very precise definition of the population at risk even more so than what has been proposed.

Okay, we need to move on to a couple of these other points as we're running out of time. We should talk about the outcomes and endpoints in these trials. The question that comes up and that I guess I've heard Rusty ask this multiple times, so we'll ask him again, since it's on here. You know, if we are going so early that there's very little clinical that we can pick up in the person, and we've carefully specified on what basis we've decided this person is at risk -- they have memory loss and X, some biomarker -- do we have to be held to multiple outcomes, multiple endpoints, or is there a place for one comprehensive endpoint and/or delay to diagnosis?

Dr. Katz: Well, the easy part of that is delay to diagnosis. That's always been an acceptable outcome for us. I don't know if you're asking if that should remain the only one, and it doesn't have to be the only one, but that's clearly always been an outcome that we have endorsed as being totally appropriate for drug approval.

I do think that if you're talking about patients who are so early that they have no clinical symptomatology at all, they just have the beginnings of some amyloid deposition or some other biomarker that's emerging as positive but they have no clinical signs, and are not likely to have clinical signs for ten years, I think that's a very problematic population on which to base an approval of a treatment. Because in that case, at least at the moment anyway, let me say that, because in that case the only possible outcome you could rely on would be a biomarker and just a biomarker, and maybe many biomarkers. But if there's nothing clinical to measure, there's nothing clinical to measure.

Dr. Doody: Well, presumably there's a baseline rate of cognitive decline with aging. Suppose you measure a difference there?

Dr. Katz: Here's where I'm going. I think it's important to understand who you're treating. There may be some people who have normal aging and they'll also have normal cognitive change related to just normal aging, and then they will have amyloid on their pet scan. The question is, are they Alzheimer's or pre-Alzheimer's, or are

they just normal people? So I think some attention has to be paid to who you're treating.

Obviously the diagnosis of Alzheimer's is moving earlier and earlier, but I don't know how early we think we can make it yet. But I think that you have to know who you're treating. And let's talk about just trying to treat Alzheimer's patients.

I think if you can detect – and this is something we've also I think said recently – I think if you can detect a biomarker, but they're asymptomatic and yet still have some subtle cognitive changes on a sensitive measure, and people believe and are convinced that we're really talking about Alzheimer's, it's possible that a change on a subtle cognitive test, you know, without a global measure and that sort of thing, that might be enough. Again, we have not been faced with this, but I would say it's worth thinking.

Obviously discussions are going on about the dominant Alzheimer's patients, and I guess we're beginning to know that those folks have cognitive changes years and years and years before they become symptomatic, but on sophisticated cogno testing you can detect changes. So that's a possibility.

Even in that case, we would like to have information from perhaps elsewhere that the drug really does have some useful effect. But I take everybody's point that the populations that we've been studying just may be too late, so the only place you may be ever able to see something is early. But I would say as long as we know who we're talking about, that these are people who have Alzheimer's, or will have Alzheimer's, and all they have are subtle cognitive changes, a drug that had an effect on a biomarker and some subtle cognitive change, that's worth thinking about.

Dr. Doody: So what if they're MCI multi-domain, clearly amnesic, does that change your thinking at all? Is that a more clearly Alzheimer population?

Dr. Katz: Well, I'm not the person to answer that question. The field has to answer that question. I think the field has to decide who has Alzheimer's and when do they have it.

Dr. Doody: So panel, you've got MCI multi-domain, you're using measurable criteria standard deviation, half from the mean for age and education, maybe you'll use the NINCDS domains, memory and something else. Are we confident? You mentioned, Howard, that we tried that in clinical trials, but there was a big difference

between what happened in some of those clinical trials versus what happened in others, in terms of rates of progression, because of how the criteria were operationalized.

Dr. Feldman: We tried to slice and dice this issue of the psychometry, and depending on populations that you look at, you can define large proportions that have multi-domain, some that have only memory. If you take the aggregate of it, I just don't think it works. I think a biomarker is a much more powerful way of identifying the molecular pathology of Alzheimer's than trying to take a high-level view of psychometry and saying, "Well, there's a bit of executive dysfunction, there's a bit of this." I just don't think that's, personal view, based on the data that we've acquired and had in our lab, don't see it.

Dr. Doody: But unfortunately, there's a kind of fuzziness that goes with the term biomarker. It's biomarker for what? And if you're talking about a biomarker for amyloid pathology, you could have that, and as we discussed earlier, never have disease. So we can't really use the term biomarker in contrast to clinical picture.

Dr. Feldman: Right, so we should be specific that if you look at the Hanson study on lancet neurology, then looked at tau to Abeta data ratio, and you looked at survivorship free of Alzheimer's from MCI, you end up with five percent at the end of five years.

Dr. Doody: Right, in MCI.

Dr. Feldman: Well, that's what you're talking about. I presume, and maybe I misunderstood the question, I thought the multi-domain was in reference to mild cognitive impairment.

Dr. Doody: Yes. Okay.

Dr. Feldman: So not every study is like that. It is the poster child for that belief system.

Dr. Doody: So you're saying to enhance the MCI likelihood of conversion, what you need is the biomarkers, which have clearly been shown. We have several biomarkers that clearly do seem to predict conversion or risk, APOE e4, low CSF tel, high beta, the ratios. If you take an MCI population and apply those biomarkers, you do increase the likelihood of seeing disease. But that sticks us in an MCI population again.

Did you want to say something, Rachel? No, Okay. Well, we're coming to the end of our time, so I think what I'd better do, I did kind of steer away from the biomarker questions as much as possible because there's a whole panel to come on that, but any closing remarks that any of you would like to make about either where we are at this point, something that you want to clarify from earlier in this panel?

Sometimes panels are a little bit disruptive because they go in so many different directions, and at the end you might be left with less than what you started with.

Dr. Feldman: I'll just make one comment. Dr. Katz mentioned the idea of adaptive designs and I do think that it's one of the things, as I ruminate on better trial designs, it's one of the things that I think about a lot, which is, time is relevant. One wants to shorten the time. One wants to make the best use of patient samples. Rachel touched on it with frequent interims. I think there really is something that we're not deploying fully, which is ways to integrate Bayesian statistics, real-time data, get multiple arms as quickly as possible down to the right arms, but it does relate more to the oncology model of the disease, like going for the kill as opposed to going for everything you need to know. So it's a philosophical switch, but I would encourage people to think about this, because it's an area I don't see much publication on that could really drive forward.

Dr. Doody: Of course Bayesian statistics were not designed for that. Do you think that Bayesian statistics are the right statistics for adaptive design?

Dr. Gamst: I'm not a Bayesian by training or for a belief system …

Dr. Doody: You're the closest among us.

Dr. Gamst: I think Bayesian statistics as a field and of itself can mean many, many things. So there is a right way to apply Bayesian methods, and I think the way to apply them is to make sure that they have what's called frequentist validity. This is perhaps a bit too technical, but there are good group sequential designs, many of which based on Bayesian ideas, and I think we should be making more use of them in Alzheimer's disease trials. It's certainly a good way to narrow a field of candidates down to a smaller subset of potentially useful ones.

Dr. Doody: In other words, a trial that involves multiple different drugs?

Dr. Gamst: Yes, there are group sequential trials that involve multiple drugs, or the same drug in multiple doses. There are also group sequential designs that involve trying to identify a population. If there are sub-populations of relevance that can be specified in advance, trying to identify that sub-population in which the drug works.

Dr. Doody: Marty, do you want to have the last word here?

Dr. Katz: A topic that we haven't really hit on that much in this panel, presumably the biomarker group will, is I do think smaller trials, more limited goals, and defining the population that is more homogenous in terms of looking at probability of response using those biomarkers with, maybe it's an APOE genotype, maybe it's an age range, they're very old and they're acting differently. Maybe it's APOE 4 or just the APOE 2-3 group. It tells you to look at Alzheimer's and it's this broad spectrum, it's a very heterogeneous population with the adenine, the things that have been done genetically, and with understanding from the various studies, I think you're looking for a signal and I think it's reasonable to pick a smaller group of patients that is very homogenous and apply your drug.

Dr. Doody: Does the patent time clock allow you time to do the E4s first and then to do the other people and then to go to the very mild early? No? Okay, well, on that note, thank you very much to the panelists, great discussion.

Dan Perry: A few questions from the floor.

Dr. Doody: I'm sorry, you told me to do that and I didn't, from the floor. I had hoped you'd just pipe up. So please speak.

Dan Perry: But they have to be brief.

Audience Male 1: Yes, very brief. It was a wonderful talk and discussion. You know, recently, well, just today actually, I was very impressed by Dr. Shaw's presentation, and one of his slides he reminded me about a data integration. I'm still using the term, we call MIT thinking, multiplex integrated thinking. Do we have a road map integrating all the new data produced in very exciting 2010 publications in various journals that we haven't touched on today. Well, we don't have time to mention that.

But I was also looking to one aspect of our hospital, Brigham Woman, Rudy had a very elegant, a very different model. I don't know whether it's validated or not, but at least 42 genes in the whole process of the Alzheimer's. And in that kind of complexity I was thinking about the possibility of using two compounds instead of one. We've been focused on monotherapy. If we borrow the model from oncology, I'm jealous of our colleagues, they're so successful in the synergistic data they show, drug A plus drug B do better than the single agent alone.

This question leads direct to Dr. Case in terms of logistics and regulatory. And how do we encourage some biotech style of company. From the perspective, we only talk about the major pharma and one compound or two, but one or two compounds may be just the bread and butter for small style of biotech. We need to encourage them. I'm not from the biotech any more. I left Johnson and Johnson, Novartis and Pfizer. I went back to Harvard for academia, so I can say whatever I like to say, because I have no financial contrary interests. But I'm very sympathetic to those people, they really put a lot of effort in the licensing compound, either from NIH or from some major institute, they want to develop something to supply the pipeline, and what do you think about the combination therapy, two agents, what should they start on there?

Dr. Feldman: Well, certainly if somebody thinks two particular agents would be synergistic or useful together, of course, there's no regulatory bar whatsoever to proceeding that way. There are rules about ultimately how you would decide whether to approve two agents together. You'd have to show a contribution of each, but it increases the complexity of various things. But we've certainly endorsed that view in other settings. We've even asked people, again, in other settings, to expose their animal models of their particular disease to poly pharmacy because people talk about it but you rarely see that. So we would certainly endorse the idea if it had merit in it. From our point of view it's not really an issue.

Dr. Doody: Can we move to the other question? Follow up quickly, okay.

Audience Male 2: First, one observation. I'm from the venture capital management side, so last time I took a biology course I think Darwin was off the boat, so take that with that perspective. I'd say one thing that when we're in the boardroom, for this drug, nobody's worried about the patent life. You may hear about it, but nobody's worried about the patent life. Any drug that's actually a meaningful therapy is going to generate so much revenue so fast that the return

on investment will be ample. So I don't think that plays in this field as much as it does in incremental improvement drugs, although it's still in the background.

So my question is this, except for Dr. Katz, and thank you again for doing this, Dr. Katz, none of the other panel talked about the kind of work that Dr. Hobart talked about, which to me, with my particular perspective, is probably some of the most interesting novel stuff I've seen in some time, given all the good work that was done at ADNI to complement it with the other discipline. So I was curious why the other speakers were not incorporating that as a possible way forward, that's my question.

Dr. Gamst: Well, I think we all meant to incorporate that, but in light of focusing mostly on phase 2 designs we were interested in doing things like drug screening and quick screening for some level of efficacy or target engagement. If you're interested in clinical outcomes, I think we all believe that improved clinic measures would be extremely useful, and not only refinements of measures that we have now, combinations of measures, but also adaptive testing.

Dr. Doody: And there are other approaches, for example, to ADAS-cog. There's IRT analysis, which also takes into account the items and their latent construct with respect to cognitive loss. There are other things being talked about. The problem is that they're really not ready for the phase 2 trials yet. So an approach like Dr. Hobart's that adds more in and makes more units on the ruler, sure, that's ready to go. But really jumping to alternative analyses or even alternative models of disease progression, in my view, the things that drive disease progression are completely ignored in clinical trials, things like premorbid IQ, things like persistency of background treatment. I think we have models that say these are important and we might have to stratify based upon those factors. So there are many other issues, but they're not ready for Phase 2 design.

Dan Perry: Let's give this panel a big hand. *[Applause]* And our second panel is taking their positions, and then we will have more time for questions from the floor, so I hope you all engage with this panel.

Rachelle Doody said that biomarkers as a term has some fuzziness to it, and the question is, to be used to measure what? I think that's an issue that this panel that is now coming forward will address.

Dr. Reisa Sperling of Harvard and Brigham Women's hospital will be chairing this panel, and the other members include Dr. Cliff Jack of Mayo Clinic, Dr. Les Shaw who spoke earlier, from the University of Pennsylvania, Dr. Eric Siemers from Eli Lilly and Company, and Dr. Norton Walton of the Food and Drug Administration. This is our panel and I will turn it over to Dr. Sperling.

Dr. Doody: Not yet, we're having some technical issues. Do you want to take a five-minute break?

Dan Perry: We're wrestling with technology up here. If you can discipline yourself to take a five minute break and come right back, because we still are trying to catch up our time. Be back at 25 minutes to the hour, thank you.

[End of Audio]

Dan Perry: Please take your seats. The sooner we get started, the more time there will be for questions.

Thank you very much for promptly returning. I'm turning it over now to Dr. Sperling

Dr. Reisa Sperling: Alright. Okay. Well thank you very much. And I have the honor of chairing a panel of a very distinguished group to discuss biomarkers and hopefully really pick up where Rachelle's panel ended up and try to figure out how we might use this to define better Phase II trials in particular to help us design Phase II so that we don't have as many Phase II failures.

So today we wanted to really focus on two points in terms of biomarkers, the first being selection of optimal steady cohorts, which is how can we use biomarkers to move trials towards earlier stages of Alzheimer's disease if at least for some therapeutic avenues. This is our better target. And, in particular, touch a little bit on the role of biomarkers in the new criteria being proposed for both prodromal and pre-clinical AD and how we might start to use these in clinical trial populations.

And then the second major topic will be using biomarkers as outcome measures in Phase II, particularly how could we use them not only as Russ said for target engagement, but to look for early signal of efficacy that would aid us in Phase III design; but importantly, as you've heard from Rachelle, we need to somehow link the biomarker outcomes to eventual clinical efficacy because this is critical.

Now it's already been discussed a bit by Rachelle's group about why we might want to move earlier in the course of Alzheimer's disease and I will make an argument that, right now, we're treating at the stage of dementia and that we need to be treating perhaps 10 or 15 years before that at some point in the disease. And I agree that we're not there yet, but I don't think it's too early to start thinking about how we would design studies to help us get there.

I'll make an argument that every disease that the FDA thus far has had success in treating has been in treating at earlier stages of the disease. So cardiovascular disease, cerebral vascular disease, cancer, diabetes, osteoporosis; any of these. We don't wait until someone has already had multiple MI's or metastatic cancer or multiple hip fractures. We treat earlier and we treat primarily on the basis of looking for evidence of the disease with biomarkers before symptoms.

Now the problem is, of course, in Alzheimer's disease we don't yet have the perfect biomarker that can predict. And although I would argue with Rachelle, I almost stood up, that there is evidence that amyloid is associated with some changes in cognition as well as downstream neurodegenerative markers on average, it is also clear that some people with Alzheimer's pathology will live their lives and never manifest symptoms. And so we desperately need to understand which markers will help us determine who is moving towards the clinical syndrome and who somehow is at peace with their amyloid.

[Laughter]

So, I will also – you heard some wonderful work by Les Shaw, and you're gonna hear from Cliff Jack. I think, at this stage, we've actually come quite a long way in being able to understand individuals who are at the stage of MCI who are gonna move further down on that trajectory. And, as Rachelle pointed out, if you use multiple biomarkers at that stage of MCI, you will, in fact, screen fail a number of subjects because it is only about 65 or 70 percent of MCI subjects who have evidence of amyloid pathology. And if you acquire both amyloid and evidence of downstream neurodegenerative markers such as atrophy or FDG hypermetabolism, you will screen out potentially 30 or 40 percent of subjects. But I'll argue that, particularly in Phase II, when what you want to do is have this homogenous population on which to detect a signal of efficacy, you want the group that really has Alzheimer's disease and not the 40 percent who don't have Alzheimer's disease as their cause of MCI and who will not progress to Alzheimer's disease over time.

Now it's much more difficult as we move back to this stage of pre-clinical Alzheimer's disease because, again, not everybody with Alzheimer's pathology will move down that trajectory. And as it has already been mentioned by Dr. Katz, you know if we're talking about a clinical outcome 10 years from now, how can we use biomarkers at this stage of the disease to help predict what will happen clinically in 10 or 15 years from now. But again I think here's where Phase II design will be critical because we can't fail in Phase III with 5 or 7 or 10 year long trials. We need to be sure what we're doing before we move into prevention trials. So Phase II will be particularly important there.

Now this – I had the honor of chairing the NIA and Alzheimer's Association panel recommendation working group for pre-clinical, and we worked very hard in trying to come up with a hypothetical model. And you saw a bit of this from Rachelle earlier based on Down's syndrome but, in fact, there's evidence from all of the genetic forms of Alzheimer's disease and even some evidence from autopsy and biomarker data. And you'll hear from Cliff Jack in particular how we might use biomarkers of each of these stages of the illness to help us in looking at disease progression and eventually as outcome markers.

Now I've been accused of living in amyloid land and I do. However, you will note that there is a question mark between amyloid accumulation and the rest of the pathophysiologic sequence of disease because we don't fully understand whether amyloid is the cause, how it may lead to these downstream neurodegeneration markers of Alzheimer's disease. And we must be very aware that there are factors that modulate whether people are going to move along this trajectory pathophyisiologically, eventually manifest clinical symptoms or not. Rachelle already mentioned this idea of premorbid I.Q., brain reserve, cognitive reserve. We all hope that there's some environmental factor that can help us and, importantly, there are also comorbid diseases, several vascular disease and other neurodegenerative diseases which may modulate how people move.

Now importantly today, we're not gonna focus on the needs for biomarker validation and standardization. This is critically important and obviously this work is necessary and I think is ongoing. But I think, right now, rather than get stuck in that morass, we want to say if we do validate markers and how, if we're able to come up with markers that are consistent, how would we use them in helping us in Phase II design. And importantly acknowledge that the issues in validating a biomarker for clinical diagnosis might be somewhat different than issues in validating or

using biomarkers in clinical trials where one is primarily looking, at least as an outcome marker, in within individual change.

And again, I've asked our panels to particularly focus on the use of biomarkers to improve the utility of Phase II trials with two questions: Can we utilize biomarkers to define study populations who are most likely to demonstrate a signal of efficacy and could we use biomarker outcomes in Phase II to help us better predict clinical efficacy in Phase III.

Now Marilyn Albert, who chaired the MCI working group, the NIA Alzheimer's Association, this is still a draft, but they're working very hard on defining MCI for research purposes using biomarkers to increase the certainty that MCI is due to AD or the cognitive syndrome of MCI would be due to AD. And she and Cliff Jack kindly lent me this slide. I'm not sure; it might be too small to read. But here, importantly, what it says is that the highest level of certainty that MCI is due to AD would involve evidence of both an A beta biomarker and a marker of neuronal injury. And that, of course, you may not always have this in the clinical setting, but you certainly would have the opportunity to look for this in use for a clinical trial population. And that, as you have less evidence of biomarkers or have ambiguous biomarkers, one should be less clear that MCI is, in fact, due to Alzheimer's disease.

Now I'll end with this slide, which I love for two reasons in thinking about biomarkers. First, up there at the very top is where I would say we're looking at Alzheimer's disease now at the stage of dementia. And, unfortunately, this is what's looming underneath the surface in terms of individuals at risk for developing Alzheimer's disease and also I like this as an analogy because we need to use biomarkers to be able to peer below the surface. So, if we only look at Alzheimer's disease in terms of what we can see in clinical symptomatology, I think we're actually missing the bulk of the problem.

So I'll turn this over now to Cliff Jack, who will tell us about some recent biomarker data and dynamic change over time.

Dr. Cliff Jack: Okay. Well thanks Reisa for the kind introduction and thanks for inviting me to talk here today.

So I apologize in advance a little bit. There was another meeting I had to attend downtown for a different study earlier today, so I missed a lot of the earlier discussions. Les' presentation I missed. But so I apologize if some of this is redundant. I did hear a lot of

the discussion from the discussion session from the last session though.

So I think it's safe to say that there are five sort of major biomarkers of Alzheimer's disease right now. By major I mean these are well validated enough to be used, to be considered for clinical trials and they're also well validated enough to be considered for use in many large natural history sorts of studies. That's not to say that there aren't other potential biomarkers. Reisa, for example is very interested in functional kind of activity measures. These may come along. They show a lot of promise, but, like I say, I wouldn't elevate them to the status right now of being sort of a major biomarker that one would consider for a clinical trial.

And one can divide these – and I think this conceptually has already been broached here earlier today. One can divide these five major biomarkers into two sort of major conceptual categories. One, biomarkers of A beta, amyloid deposition. These are low CSFA beta 42 and positive PET amyloid imaging. And the second category would be markers of downstream neuronal degeneration or injury. And these are elevated CSF tau, a certain topographic pattern on an FDG PET demarcating synaptic dysfunction. And finally atrophy on structural MRI, again in a specific topographic pattern that recapitulates Brock staging very well. Brock NFT neurofibrillary tangles staging very well. So these are the five major biomarkers that I think should be taken into consideration here.

Now what I'm gonna propose is, describe is the idea that these don't become abnormal simultaneously in a subject who develops Alzheimer's disease but, rather, become abnormal in a staged or ordered manner. And just so you won't think that I'm making stuff up, I'll show some data that I think supports this idea.

So this is data taken from Mayo and what we did in this study was look cross-sectionally at two of the – only two of these biomarkers. One was PIB imaging and the other was hippocampal volume. So hippocampal W score is another way to think about hippocampal volume. And these data were scored by or laid out in this scatter plot by clinical diagnosis.

Now there is a limitation in this – there's a real limitation, an obvious limitation in looking at this data in this way and making inferences and that is that not all of these control subjects are going to develop Alzheimer's disease, clinical Alzheimer's disease. And, as Reisa has just pointed out, not all of these MCI subjects – in

fact, if they were to come to autopsy, would have Alzheimer's disease as the primary pathologic diagnosis. About 20, 25 percent would have something other than Alzheimer's disease as the primary pathology underlying their cognitive impairment if they were to come to autopsy, and this is based on well done autopsy studies. Nonetheless, this is sort of the best we can do at this time.

So if you look cross-sectionally at patients, you can see that the MRI data makes sense. So the hippocampus in the brain is least atrophic in people who are cognitively normal; most atrophic in people who are demented and intermediate on average in people who are intermediate.

The PIB data looking at brain amyloid deposition as a marker of brain amyloid deposition likewise makes sense in the same way. Highest levels in demented people; lowest in controls; and intermediate in MCI subjects. It looks like this is a bimodal distribution, but that's just because of the small sample size of the study. If you were to do a lot of people, this would fill in as would this.

Now, but there is a major difference between – so the biomarkers make sense. But there is one exception. There's one important difference and that's indicated in these red circles here. And that is that – so a number of 1.5 is generally taken to be the cutoff between normal and abnormal PIB PET scans. So this would be if a reviewer was looking at a scan ratio, a quantitative ration of 1.5 generally corresponds to what would be considered a positive scan. In about 20 percent of these – in fact it's a little higher than that – 20, 25 percent of these normal controls have positive PIB scans. In contrast, there's no analogous region of this scatter plot where people who are cognitively normal have shrunken brains, have shrunken hippocampi.

There is this one exception here, you can see one dot; there's always one outlier in any real data set. This was a PIB negative control subject and, in fact, as I looked at the hippocampi in the subject, it looked to me like a developmental anomaly. The hippocampi were under rotated, kind of like you see in kids with developmental anomalies, which I think that's the reason the hippocampi were small.

So the data generally makes sense in the sense that worst biomarker maps onto worst cognitive function, but there is this notable discrepancy so that cognitively, people can be cognitively normal and have a head full of amyloid, but you don't see the

analogous area over here of people who are cognitively normal and have shrunken brains.

Now we look at rates of change. So the same three groups here, control MCI needy, this is combining Mayo and ADNI data because, at the time, that's the most data I could get my hands on was to combine the two studies. Rates of change data were not so prevalent at this time. And you can see that the annual rates of change again on the MRI side makes sense. Highest rates of shrinkage, brain shrinkage in patients who are demented; lowest in cognitively normal subjects; intermediate in MCI subjects. But, if you come over to the PIB data, you see something that's very striking and that is that the annual rate of change, in fact, is not different between these three clinical groups on average. It's not zero, so the annual rate of change is greater than zero, but it's kind of small. But it is not different between these three groups.

So, putting these two sets of data together, the cross-sectional data and the longitudinal data, you come up with the following sort of general conclusions. One is that it does appear to be the case that you can have – elderly subjects can be cognitively normal and have a positive PIB scan; so a head full of amyloid. But you don't really see marked brain shrinkage in those subjects.

Secondly, from this data, the time in the disease – in the stage of the disease, where the clinical change is maximum, the rate of change in MRI maps onto or correlates with the observed simultaneous rate of clinical change. That's not the case with the PIB data. So the PIB data is, the amyloid deposition data is uncoupled from the rate of clinical change in these subjects.

So how can we put this altogether in a graphical way that kind of makes sense? So this is what we came up with at this time. On this axis is magnitude. It's a little bit confusing, because the magnitude of amyloid increases here and they may be due to the brain shrinks and the magnitude of cognitive function shrinks as well. So it's a little bit confusing in the sense that the different metrics are going in opposite directions.

But this axis right here represents a pre-symptomatic subject's – this prodromal and dementia. So this would be the cognitive trajectory that any individual subject would follow if they were to progress in the normal way from normal to MCI to Alzheimer's disease. And this, then, is a graph of the analogous, the two biomarkers we looked at, amyloid and MRI, that is suggested by the data we had on hand which is that, while subjects are still cognitively normal, they can develop a head full of amyloid.

These dotted lines represent hypothetical back projections; we don't know if it's a linear back projection or should it be a non-linear function. But we presume that subjects can, indeed, develop a head full of amyloid while there are normal controls at which point the rate increases but at a slow, relatively constant rate.

While subjects are cognitively normal, the rate of change of MRI brain shrinkage is negligible. But, at some point in time, it begins to accelerate. And as it accelerates, the subject moves from normal to MCI and then ultimately dementia. And this is just a general line of cognition which is depicted as paralleling but with a right word temporal offset in relation to the brain shrinkage align.

So, from this, you can – we kind of propose this idea that Alzheimer's disease is sort of a biphasic pathologic process and, hence, a biphasic biomarker process where the initial phase is a phase where amyloid biomarkers are gonna be most active and, later on, is a phase where neurodegenerative biomarkers are going to be most active. And this maps onto a phase where the patients are clinically asymptomatic.

And so this is the sort of first cut at this model we made. I should say that at least three other groups have independently come up with sort of a very similar notion. One is Dave Holtzman and I forget the name of his Fellow, but at Washington University, Bill Jegust, the Fellow who published the paper there was Elizabeth Ormeno, who now works with Reisa, and Brad Hyman, looking at past data and the Fellow who was the first author of that paper was named Ingelsson. But all really four different groups of independently have sort of come up with this general idea.

So we talked about five biomarkers and I've only shown you data from two of those biomarkers: MRI representing neurodegeneration and PIB PET representing amyloid deposition. So this is the next cut at this sort of model and this should say a hypothetical model because it is really hypothetical.

Here what we've done though is every biomarker is scaled from normal to abnormal. It eliminates some of the confusion of some of the lines going down, some of the lines going up. So every biomarker is scaled from maximally normal to maximally abnormal. The X-axis here are the same. So this is a hypothetical model of what happens to the biomarkers in A subject as that subject transitions from normal, technically normal, through MCI through dementia.

And this line right here is a reference line if you will. This demarcates single domain cognitive function, most typically memory. You could see it swings away from the X-axis just as the patient transitions from normal to MCI. This is almost by definition and, similarly, this line here represents general cognitive function which swings away from the X-axis as subjects move from MCI to dementia. Again, a definition.

Then we have three categories of biomarkers here. And this is the amyloid biomarker, which here we plot CSFA beta 42 and PIB PET or amyloid PET imaging together, not because we're absolutely certain that they go together in every subject, but just because, at the time this was created, there was no hard data to suggest anything other than this. We know that they correlate with each other very well in individual subjects.

And this line is depicted then as a curve or linear function that decelerates as subjects approach MCI and then dementia, again obeying the data that we had earlier.

This line right here represents tau mediator neuron injury and synaptic dysfunction. So, with the idea that, of the neurodegenerative biomarkers, it may be the case that biomarkers of tau mediated neuron injury and dysfunction become abnormal earlier than this right here, which is frank neurodegeneration or shrinkage of the brain with MRI. At the time this was created, there was some evidence that, at least FDG PET because abnormal in subjects destined to develop Alzheimer's disease earlier than MRI. And so this graph reflects the data that was available at the time. We also looked at some of the ADNI data and there was a suggestion that total tau became abnormal before MRI. So we bundled these two together as a line that preceded MRI, again not because we had hard evidence that tau and FDG PET moved simultaneously, but, rather, that there was a lack of evidence to suggest anything else.

So this is the graph then that, the working graph that we came up with modeling the trajectory of biomarkers as a function of disease stage in an individual subject who would progress through these clinical stages.

I won't go over this. Reisa's already gone through this. But then – so the last part of the presentation I'll just a couple of slides illustrating the roles of the biomarkers in clinical trials. Reisa's already indicated that there are two major roles: subject selection and its outcome metrics. Safety monitoring is also important but I won't mention that at all here today.

So one can then superimpose this notion of the trajectory of biomarkers and how one would categorize patients using biomarkers and clinical data, what the biomarker data should show at each stage, and then what the appropriate treatment or options might be. So, again, this is totally hypothetical. But it makes sense I think.

So, at this stage, this might be someone who is at risk, say an Apley 4 homozygote. The odds are almost certain that, if that such a subject lives long enough, they will at least develop brain amyloidosis. All biomarkers would be negative at this stage and the only logical treatment option at this point would seem to be A beta prevention.

At this stage, the A beta biomarkers become positive. We might characterize this subject as pre-clinical. In fact, in Reisa's working group, this is sort of the first stage of the pre-clinical diagnostic – this is where subjects would enter the pre-clinical diagnostic paradigm. Here the amyloid biomarker would be positive, but a tau marker and a degenerative marker might be negative. Here, logical treatment options would seem to be A beta prevention, arrest or reversal, and then tau prevention and neurodegeneration prevention.

At this stage, subjects are amyloid positive, tau positive but the degenerative biomarker is still negative. Again, this is still based on this hypothetical notion that tau biomarkers should precede shrinkage of the brain, frank shrinkage of the brain. Again, such a subject would be classified as pre-clinical because they're still in the clinically asymptomatic phase. Now it is the case that notice, these subjects are subjects who are in this phase by definition at some point are going to make this transition from normal to MCI. So it is the case that this sort of subject, at some point in time, while they might still be classified as cognitively normal, are going to be demonstrating subtle deficits on cognitive tests. And that's indicated by this line here, which you can see the signal – this is a memory test – but this indicating deflection away from the baseline while the subjects are still classified as cognitively normal. Here the treatment options would be A beta prevention, arrest, reversal – same as tau – and then degenerative prevention.

And then finally, this stage, which would be MCI and dementia. Here all the biomarkers are abnormal. Here patients are symptomatic whether they're in the MCI or the dementia phase of the disease. It's just a matter of time. And here, logical treatment options would be prevention, arrest or reversal of A beta, tau, and degeneration and symptomatic treatment.

And so, not to beat a dead horse, I realize this has been mentioned a number of times here this morning probably, although I wasn't here for a lot of the morning. But you know, so why has everything failed? Well, it's because the – oh, I just screwed this up. So why has everything failed to date? Well, it's because modern therapies targeting A beta, which would logically be directed at this stage of the disease or maybe this stage of the disease, have been applied to people in this stage of the disease. So I'll leave it at that.

[Applause]

Dr. Sperling: Alright great. I think Les is gonna give some comments. He already spoke with slides this morning. So just, if you have additional comments on those two points, it would be great.

Dr. Les Shaw: So, in terms of the broader question of selecting optimal cohorts, what can we say about the biomarkers that we measured and I think Cliff has really very nicely talked on these points and some of our earlier speakers as well. But, taking the example of the MCI cohort, the amnestic MCI cohort, clearly some of these folks will be stable for a very long period of time. Some will go on to develop other neurodegenerative diseases and the majority, let us say more than half, are likely to progress to Alzheimer's disease.

So, to the extent that it's helpful to at least provide a measure of risk for progressing to AD, we think that the imaging and the biochemical biomarkers can help that process. For example, using the tau A beta ratio as an example, we saw about a fourfold difference in folks who, between folks who had a normal ratio of those two biomarkers as compared to the rate of progression in the folks who had an abnormal tau A beta ratio. So to that extent, we feel comfortable to say yes, we think that these biomarkers and, in combination with imaging biomarkers, can segregate groups, this particular group, MCI folks, into high risk/low risk in that sense.

Interestingly, what about the 10 to 15 percent of folks who have normal biomarkers but who progress to probably AD? What about those folks? Is this a somewhat unique group in certain characteristics, certain other risk factors that aren't present for example; is clearly an area requiring further investigation.

So we feel that the standardization efforts that have been made worldwide certainly with ADNI for both biochemical and imaging biomarkers, have helped sharpen the tools – they're not perfect tools – and feel comfortable to say that these tools could be used in this way if that's going to help in the design of studies.

As far as using these as outcome measures, I think the most profound conclusion that I can make, and it's not very helpful unfortunately, is we need a lot more data to start to understand these as measurements of treatment effects, measurements whereby we can further interpret what is going on in terms of a particular treatment and effects of that treatment. So I really don't feel I can help there other than to plead that we incorporate these in more trials with larger numbers so we can determine this aspect.

So, those are my few comments.

Dr. Sperling: Great. Eric Siemers. Would you like to speak now?

Dr. Eric Siemers: Sure. So I guess just to cover both points. So, in terms of using these biomarkers to define the population better, I think that's certainly a place where they can be utilized. Now the one thing that – probably not appropriate certainly for me to get into – is that's also, well at some point involves some labeling discussions and that sort of thing.

But, with that aside, just as a clinician and from a biological standpoint, if you want to use these in a trial as inclusion/exclusion criteria, then I think – and people might disagree with this – but even for mild to moderate Alzheimer's you may be pushing the envelope a little bit to say, "Well I'm just so good that I don't need this scan to tell me whether or not they have the pathology." A lot of people will say that and I think it is true though; you're gonna be wrong in mild to moderate Alzheimer's 10 to 15 percent of the time.

Now is that 10 or 15 percent important? Well then it depends on a couple of different things; 10 or 15 percent is not that big of a deal. But, if you had a drug that was either particularly not well tolerated or, the other thing you have to say, a drug that's very expensive, then I think eliminating that 10 to 15 percent of people will end up being quite important. And so I think we can – certainly they will make the trials go better because you don't have that 10 or 15 percent muddying the data. But, beyond that, it's what medicine does. You define your patient population better over time to make your treatments more targeted for the actual pathology that you're doing and that's been true since x-rays for pneumonia for instance.

Now, if you back that out a little bit and you go to mild to moderate, in that case your biomarker – and just to simplify this, I'm of the thought that if you have clinical evidence of MCI plus a biomarker – we can call that prodromal AD or whatever. But, in that case, rather than eliminating 10 or 15 percent, you're gonna

eliminate roughly 20 to 30 percent. But the same issues I think come up. I mean the conventional wisdom is if you're moving earlier, then the drug has to be even better tolerated. Cost is again going to be an issue because they're going to be treated for a longer period of time. So it's important to be able to eliminate those people. And here I think it's crucially important for the success of the study because, if you've polluted your population with 20 or 30 percent of people who don't' have the biology, it doesn't matter how potent your A beta related drug is. It's probably – you're probably not gonna hit statistical significance in Phase III.

And then finally, if you back it all the way out to pre-symptomatic, now depending on exactly what age group you look at, you're gonna eliminate 70 percent maybe of the population; 30 percent let's say roughly have positive scans.

I think there's a lot of work that needs to be done in this and I don't want to spend too much time talking about it because it is complicated. I think there's some interesting data coming out of DIAN, the autosomal dominant cohort, that your development of clinical symptoms that you can pick up on without some sophisticated testing happens at about the time when tau goes up. And so that might be something that could be considered.

There's a study that I just read about that came out of American Heart just within the last couple of weeks. They actually said maybe statins are over utilized because it's not just people with elevated cholesterol, but it's people with elevated cholesterol plus calcification of their plaques. And I think that's what we might need from a treatment trial design in this pre-symptomatic population. They're at risk; not everybody will bet the disease. But, if we can define maybe staged biomarkers that really puts you at increased risk, that would be a way that we could do these trials.

I do think, especially when you talk about pre-symptomatic and the MCI trials, that this conversion, especially for MCI, it's not gonna work very well. I guess this has already been mentioned. One thing I just encourage people to look for; that sort of outcome can be looked at in the autosomal dominant Alzheimer's cohort, but those are hard, long studies to do and, maybe to save the field a little time, there's some Huntington's – there's a Huntington's study that will come out next year. It took 10 years, 1,000 patients. It's a binary outcome. I mean it's a binary biomarker. Either you have the gene or you don't. And the bottom line is the conversions don't work very well. And so I think we really have to look at these longitudinal continuous measures rather than this conversion.

So, to move then to the Phase II point, and how do we use biomarkers to be more successful in Phase II. One thing: Rachelle mentioned that people would quibble with their slides. I'm gonna quibble a little bit with one piece of it.

[Laughter]

But I wanted to bring it up because I think, actually I hope that this is an important point is that, for Semagacestat – I guess the quibble part is we didn't actually ignore the spinal fluid A beta. We did another study to show that which actually cost us a year and we would not have gone to Phase III. So we didn't decide it wasn't important 'cause we didn't get it in the first Phase II study. We did another study because there were some technical issues in the first Phase II.

So we have plasma A beta. We had spinal fluid A beta. And then, as I think most people know, the study was stopped prematurely because people in active treatment got worse. And the important point here I think is that, on the right hand column of Rachelle's slide – and this is everybody's got a slide like this and most of them look about the same – but it says all those studies are listed as negative. As far as I'm aware, that all the other studies other than semagacestat, there was no separation between active treatment and placebo. So you just didn't see any effect at all. For semagacestat, that was not the case.

Now unfortunately, the direction of the change was not the direction that we wanted, but it wasn't a negative study in the sense that there was no drug effect. So, and mechanistically what that means we don't know yet. It could be something related to the gamma secretase inhibition; it could be something just related to semagacestat. We're gonna try to sort that out.

But the point is it didn't do nothing. It did something and so, what I hope and maybe other people will comment on this, is that that means that we're kind of on the right track with this biomarker strategy for Phase II. Now what we have to take from that I think is should we have looked at something in addition to just plasma A beta and spinal fluid A beta. Can you look at volumetric MR and try to make that a six month study so it's not an 18 month study. I know out of ADNI, there's a lot of numbers thrown around in terms of powering there for a year and you get roughly maybe 80 patients in arm. You know we'd really like to make it even a little shorter than that. So I'd like to think that maybe this biomarker strategy for Phase II, we're on the right track. And, hopefully, the worsening with something that's not gonna happen with other

compounds, but, anyway, maybe some other people have comments on that.

And then finally, just the last point about Phase II, Reisa mentioned, and I think this really is a good point and one that I hadn't thought a lot about, is for these pre-symptomatic studies, Phase II will really become crucial because you know we've already talked about the fact that there isn't really anything clinical to measure without some really sophisticated cognitive testing. So, if, in your registration trial, you're looking at a biomarker plus a sophisticated cognitive measure, I mean what do you do in Phase II that's sort of simpler than that. I certainly think you need to get proof of mechanism and proof of target engagement, but what can you do beyond that? And so that will be difficult and so maybe that's a place where adaptive designs could actually play a role. So you do a Phase IIa that is proof of mechanism and then your IIb, maybe you do an adaptive design, have frequent interims, or somehow try to make that into a Phase II/III. But that would be difficult I think. So thanks.

Dr. Sperling: Thanks. Marc.

Dr. Marc Walton: Since I don't actually work on developing any specific biomarkers or specific drugs, I get to talk about things from a higher level perspective. The answer – the first question you asked is can we use biomarkers for patient selection in Phase III studies. I think the straightforward answer is well of course we can. But the real question is well how? And, before that can be even considered, an earlier question has to be asked and that is what is the purpose of the Phase II study?

The earlier sessions talked about Phase II studies and different designs and brought out that there can be different objectives that one is looking at for them. It's really a – that is really getting to the point that there are different purposes for Phase II studies that might be selected and what purpose do we want the Phase II study to serve? There are a whole variety. And what purpose we want out of the Phase II study is going to drive both how we use the biomarker and which biomarkers we're willing to use.

Phase II studies can be just a clue that the drug is getting to the receptor or anatomic location where it's supposed to act, which is important but it's a somewhat low burden for Phase II studies. Or we can be asking for Phase II studies to really give us a pretty solid basis for belief that the drug is gonna be efficacious and that we have the right does and we have the right population. But asking

for the sun, the moon and the stars out of a Phase II study is a big task.

The advantage is that it lowers the risk for the Phase III studies not showing what we want. And those decisions have to be made. It may be that what sometimes is necessary is several Phase II studies to answer these different questions or perhaps, more efficiently, as was mentioned earlier, a sort of a staged design that has interim analyses, adaptive designs where the different purposes of a Phase II study might be looked at during different stages of the Phase II study to make the appropriate decisions for these different purposes about population, about dose, about quit altogether or go on.

So the question is really going to be what is the purpose of the Phase II study, how does it fit within the entire drug development program, and then we can begin to talk about how to use the biomarker for selection of population. And I think it's, for that kind of a purpose, it's really gonna be important to also distinguish between two categories of enrollment criteria biomarkers. One, which we're calling, and is widely called, prognostic biomarkers, which range from everything from those that identify patients who have the disease versus those that don't have the disease. And clearly that has a lot of importance and it gets very difficult. The earlier one goes in Alzheimer's and includes also identifying patients who have the disease but who their natural rate of progression is going to be different.

In the extreme, perhaps it's identifying truly different subsets of Alzheimer's disease patients, almost subtypes of Alzheimer's. In other cases, it might just be within some continuum of the same pathophysiologic processes; for whatever combination of reasons it's occurring at different rates. And we might want to use that kind of a biomarker as enrollment criteria in the Phase II study so that we have the sensitivity to see an effect in the Phase II study.

Alternatively, there are biomarkers that are called predictive biomarkers. Those are biomarkers that are going to tell us which patients have some intrinsic susceptibility to the drug that we're testing. And those are going to be rather more drug or drug class specific because it might relate to the mechanism of the drug. Patients who do not have a certain aspect of pathophysiology very prominent in their disease are not going to respond to drugs that are dependent upon acting upon that aspect.

So that highlights I think that there are different purposes to the biomarkers. We're gonna have to decide what the purpose of the

Phase II study is and match those. It was – Eric brought out that he didn't want to talk about it, but it's important to actually think about where we're going in the whole drug development program and what patient population we are looking to ultimately have demonstrated we have an effect in. If we're going to use a biomarker to select out some patient population that is unique in some way and has some particular subcategory of pathophysiology and then use the Phase II study to identify those other patients in whom the drug is active and go on to develop just those patients, then clearly that has a lot of implications for what the labeling would be should the drug prove effective in the end.

Alternatively, if we're just selecting out patients that have somewhat faster rates of decline but we have confidence is not really selecting out a different type of disease, then that's a population of convenience, of enrichment for sensitivity to showing a drug effect of any kind. And that has much less implications for labeling of the indicated population.

Which type of a biomarker we are using though depends upon what information we have about that biomarker; what has it been established to demonstrate. It's important to remember that, although we want to use biomarkers to accelerate our drug development program, and it's clearing so that established biomarkers can do that. It can make the drug development program more efficient, more effective. That's with well established biomarkers.

But getting that well established biomarker is not necessarily a quick process. That does require a fair amount of effort in and of itself and can't be expected to necessarily be a quick process in a disease such as Alzheimer's, which has limits on our understanding of the disease and has a time course that is, itself, slow.

So the second question you asked was about outcome and the short answer is pretty much the same as I gave for using the eligibility criteria. They can; they can be very good at that, but what's the purpose of the Phase II study? What are we trying to learn from the Phase II study? How does learning that fit into the overall drug development program?

And I think that, to keep it a little bit shorter, I want to bring out the idea that, in both the eligibility criteria and the outcome measure of biomarkers, what we're really talking about can be thought of not as biomarkers, as a laboratory measure, but, rather, as a drug development tool. And this comes to my mind because of the guidance that we have put out from the FDA about

qualification of drug development tools, not because these things necessarily need to be qualified, although we hope that there will be interest in doing that and thereby demonstrating their generalizable use, but rather because, in all the discussions we've had, it really has brought out internally how much we rely on these things and want to rely on these things for drug development, as a drug development tool. It doesn't matter what it is in and of itself. It matters what does it do for us in the drug development program. And that's why, in this particular case, we see both patient reported outcome measures and biomarkers as fitting within the rubric of a drug development tool because they serve a purpose of no value in and of themselves, but they can serve a real valuable purpose in drug development.

And that, I think, nicely leads into another point that I wanted to bring out, which is going back to the talk that Dr. Hobart gave about the cognitive measures and the refined scales and something that he did bring out. But I want to emphasize as well is when he talked about the benefits of his – the approach that he is developing.

A lot of what he talked about was the statistical, the mathematical characteristics of it, about how it can more finely differentiate patients from each other, how it can more finely show small changes. He didn't talk about how those fine changes are automatically assumable to be of clinical importance because that's not the aspect he was bringing out and I think that's very important to recognize. It means that he is looking to develop a tool that can be sensitive to effects in patients and that tool can have a lot of uses as an outcome measure in Phase II studies. That can be really valuable as a drug development tool, not as a clinical efficacy scale necessarily, but as a tool that identifies that, on this – think of that scale perhaps as a biomarker. It's a biomarker that is sensitive to the very late aspects of human cognitive physiology. It's testing sort of integrated neuronal pathways and telling us how well those pathways function, which is really very close to the clinically meaningful outcomes. Whether or not the scale he develops or has employed actually is clinically meaningful is less important than believing it really is at least close to it and therefore has much less of chance of misleading us. So more sensitivity on those sorts of things has a lot of power in Phase II studies. I think I'll leave it at that.

Dr. Sperling: Great. So let me push you on the last point and get other people's opinion in terms of, at the stage of Phase II, if the issues of selecting a highly characterized population who may be less generalizable but might be particularly useful for – I don't want to

use the term of futility trial because I know that has a formal meaning – but a way of actually not only looking for a target engagement but specifically looking at a kitchen sink of biomarkers to say if you have a highly leveraged group of MCI due to AD people who are all biomarker positive and you don't see any change on any of these biomarkers, is that useful at this point for saying this isn't the right drug to take forward? And I wonder that obviously is a different context for labeling, 'cause that's not a generalizable population, but is that a useful way of thinking about Phase II as proof of concept on at least at biomarker stage.

Dr. Walton: Yeah, oh I think very much Phase II studies – ideally it would be nice if we could get that kind of information in Phase I studies. And with – and it may be that with a panel of biomarkers that are strongly believed in, even a Phase I study might be able to get us that information. But, if not, that proof of concept that a drug is engaging in the pharmacologic sites we needed to, that is having some physiologic effects, is really important. If we believe those biomarkers are telling us what we hope they are and we believe that it is a necessary element of action of the drug to impact those biomarkers and they fail to show an impact, I think most people would then feel that it's probably not worth proceeding on it and to stop the drug.

Dr. Sperling: We haven't really addressed but, in terms of the upstream biomarkers, can we use CSF A beta and PET amyloid imaging interchangeably for the two purposes? So one for selection and secondly what are the differences in terms of outcome measures?

Dr. Shaw: So I think this is a –

Dr. Sperling: You might have to speak – microphone.

Dr. Shaw: This is a topic that's under intense investigation right now. I know within ADNI, Cliff Jack and some of our colleagues have been pursuing this and I think it's probably accurate to say that, so far, is pretty strong correlation between low A beta and abnormal PIB, measuring plaque burden.

It's also probably accurate to say that we need a lot – and Cliff had alluded to this – we do want to see a lot more data though, a lot wider range of data and a lot larger number of subjects in order to more fully evaluate the question as to how tightly correlated these two measures are. But so far the data supports the concept under certain circumstances almost of interchangeability of these two; not saying that there aren't fine differences in certain individual

patients, because there could be. We do need a lot more study data. But I think we're on a track documenting this relationship.

Dr. Jack: So Reisa asked two questions, and I think the answer to the first question is yes. The answer to the second question is probably no, but we don't really know. So the first question is the one that Les just addressed and that is that, for purposes of identifying people with brain amyloidosis at a particular point in time for inclusion or exclusion from a clinical trial, I think the data is very strong that the answer to that is yes.

And, as a matter of fact, Les and I and some others in ADNI have a paper that – I don't know if it's been published yet, but it's accepted in Alzheimer's and dementia, where we've developed a mathematical formula for transforming one unit of measure, CSFA Beta 42, into standard units of PIB, and it seems to work pretty well.

So, for the first purpose, which is subject selection, the answer is yes. I mean the other point is that the natural rate of change of both of these measures over a fixed point in time that's you know reasonable for most of these measurements – six months, a year, etc. – is very slow. So it's not the case that you have rapidly changing rates that would you know confound interchangeability of the measures. They're both very slow. So wherever a person is on either one at a given point in time is pretty static.

The second point is – so the second point is: can they be used interchangeably as a measure of therapeutic efficacy. And again, my sense would be probably not, but no one really knows because obviously what you would be detecting with the two biomarkers would be quite different. You'd be detecting an actual reduction in the amount of fibrillary beta with one assay, and with another assay, you'd be detecting an increase presumably in sible A beta. And how interchangeable those are is anybody's guess, but my guess would be you'd have to do a lot of validation to convince yourself that you were measuring the same thing.

Dr. Sperling: Eric.

Dr. Siemers: So yeah, those are good questions. In terms of using either spinal fluid A beta or amyloid PET as an inclusion/exclusion selection criteria, certainly the one thing I probably should have mentioned before in terms of inclusion/exclusion criteria, I was kind of making the assumption that you'd use those same criteria for your Phase III. But, as has been brought up by others, you wouldn't

necessarily have to do that. So, in Phase II, you could have this enriched population.

Now the problem with that is then, when you go to Phase III, that's a risk because now you're in a different population of people. I think – personally I think it's just better practice of medicine to define your population better. And so I think, yeah, you do it in Phase II but then I would actually also want to see that in Phase III.

In terms of using those as outcomes, that actually brings up an interesting point that was alluded to also, and that's – and it's kind of the point of this meeting I think – is how can we do Phase II's that don't like Phase III's that don't take 18 months? So, if for CSFA beta, 1 to 42, as you all know, it's low in Alzheimer's patients and say, with secretase inhibitors, that gets complicated as an outcome because they would initially lower it, but you might normalize it after the drug washed out. So, and it's also complicated for antibodies. So, as an outcome measured, just because of the specifics, I think spinal fluid A beta is a little tricky to use in trials.

Now if you get to amyloid PET scanning, you know there's these very nice Bapineuzumab data that's showing an 18 month study. You really look like you had an effect on plaque load. The problem is that was an 18 months study and so you could look at those Zero Later time points and say, "Well, you know, are you gonna go with the trend that's not significant or the year?"

And then to get to your point about sort of throwing the kitchen sink in Phase II, that's I think where we are right now. So if you say that, okay, you have it in plasma and CSFA beta is good but we need to do a little bit better than that as an outcome in Phase II, then – and one of the problems of course is just the expense of these studies – but ideally, what you'd like to do is all of the above essentially. So you'd want to look at amyloid PET and you'd want to look at volumetric MRI in a Phase II, but the difficult part is gonna be do you make those shorter than 18 months. You would hope shorter than a year. If you could get those down to six months, you could do it, but then your sample size may go up so high that you know it takes forever to enroll.

So these are all things that need to be balanced against each other. But I think the general concept is there.

Dr. Sperling: So maybe it's time to see if you guys have some questions if you're not too hungry or your brain's not too full. Do you have

questions or things you would like think about biomarkers in particular and how we might utilize these? Yes Rachelle.

Audience (Doody): Just a question for Cliff. To what extent are the shapes of those S-curves that you showed influenced by the properties of the test? So detection levels and things like that. And, if they are influenced, then the overlaps will influence. So that's the first part of the question.

Dr. Jack: So I mean yeah, absolutely. Every test – I mean I'm speaking to the choir here. But I mean every test has, is you know like a filter. You have an underlying biological reality and then the test, whatever it is, has a certain fuzz around it or a certain penumbra of uncertainty and absolutely. And it may be the case that there, just like in any test, there are probably ceiling effects and floor effects.

So, in fact, you know it may be the case that, for MRI – you know why would the rate of brain shrinkage flatten out? Well, if you mathematically project a certain rate through the expected lifespan of someone who ultimately winds up in a nursing home, you'll wind up with a brain size that's smaller than the skull. It can't happen. So it has to level off at some point.

But, it is the case that the measurements have a certain degree of – there's a level of uncertainty and the level of uncertainty may change at different stages of the disease. So I'm not certain if I answered your question.

Dr. Sperling: Right, or a different location. So, for example, maybe tau goes up earlier than we think? But, because of levels of detection, you know we have it cast in a different spectrum of the pathophyisiology of the disease. These are just you know thoughts that lead to the second part of the question.

From the clinic based samples, largely coming out of Sweden and Netherlands and places like that, why in mild Alzheimer's disease or moderate Alzheimer's disease are there so many people who are biomarker negative? You know have we really validated that conceptualization of how the biomarkers change over time in naturalistic settings?

Dr. Jack: Yeah, so this is a good point; really good point. And I think the – so the reality of it; I don't know if this has ever been – I don't know if it was discussed earlier today, but if one looks at autopsy studies, you know the unpleasant truth is that most people who become – most elderly people who become demented don't just have Alzheimer's disease. They have Alzheimer's disease plus other stuff, and that other stuff is either vascular disease, the most

common pathology is micro infarcs, or alpha-synucleinopathy, or both. So all of which, in additional to Alzheimer's pathology, increase in prevalence with age. So, as the older a cohort gets, the less clear the relationship becomes between the presence and severity of Alzheimer's pathology by itself and dementia, because you have the additional influences of these other pathologies.

So, to answer your question, how could someone become demented or impaired but be negative on Alzheimer's biomarker studies? I think the simple answer is that the more population based and the older a cohort gets, the more you get into this regime where you have people with mixtures of pathology. And the cognitive status of any individual is a function of the additive sum of these pathologies. And so it's very easy to envision then, in this context, people with sort of marginally elevated AD biomarkers that don't quite cross the line to positivity, but who have other pathologies, and that are demented but with negative biomarkers.

Audience (Doody): But the flip side is that they may also have a protected factor, be that cognitive reserve or something else. And if we throw them out of our populations for clinical trials, we're throwing out people who suffer from the disease or who will suffer from the disease.

Dr. Jack: Right. So cognitive reserve; again, another, if you will, confounding factor. I mean it's very clear that cognitive reserve you know protects you from the, to some extent or for some period of time, against the clinical consequences of your pathological burden. And that varies a lot.

I have to say, in these biomarker studies that I showed the graphs of, it's easy to draw lines. But it's actually very difficult to really generate good data for the reasons that you've just mentioned, and that is that the real populations are dirty pathologically. In real populations, you have this effective cognitive reserve. In real populations, you have this fuzz around the reliability of any biomarker measurement. And that makes it very difficult.

The other thing is that the actual rate of change of anything, except in people who are really demented, but the actual rate of change is very small over a reasonable period of time over which you can make a measurement. So the actual slope of any measurement over a year or even two years, which is kind of getting to the outside of what a reasonable time measurement, the slope is very tiny in relation to the range of the intercepts in a population. So this makes it very difficult to I would say do these sorts of analyses in a reasonable way because you have to somehow adjust for inter subject variation in the intercepts without really having a great

handle on what the slope is. I think it's a very challenging thing. I don't know if I answered your questions.

Dr. Sperling: No, but I think you brought up one really important point, which I think we should talk about in terms of biomarkers in that the – I feel like they are mimicking the autopsy data and we shouldn't hold biomarkers to a higher standard than we hold autopsy data. So autopsy studies suggest that 15 to 20 percent of people, especially older populations, who die with probable AD clinically, don't have, don't need pathologic criteria for Alzheimer's disease or have mixed.

And similarly, again as you brought up, these 25 or 30 percent of people who die with normal cognition, who have evidence of Alzheimer's pathology, again in David Bennett's big studies, if you look carefully at those people, they have slightly lower episodic memory. So, in fact, I think that biomarker data thus far really suggests that we are only as good as the autopsy data and that some of the questions that remain from the autopsy data back in the '80's, which is how can you die with a head full of amyloid and not be demented, and how can you be demented with Alzheimer's disease and not have Alzheimer's pathology are still there. And the biomarkers now have take us to the point where we can see that *in vivo,* but we can't expect them I think at the moment to be any better than the autopsy data are.

Those questions and mysteries still remain and, hopefully, actually therapeutic intervention might help us solve some of those mysteries, which is if you take amyloid out of the brain, do you actually make it work better.

Audience (Doody): And the Phase II question is are all biomarkers equal for the purpose of deciding whether or not a drug works. And so we've all converged around amyloid for many reasons and it may work. But the worry is that we're gonna try to do that with other things.

Dr. Sperling: Right and I very much agree, or at least I was trying to make the argument, and perhaps other agreed or disagreed, that I think we might need multiple biomarkers at Phase II. And making decisions on the basis of one biomarker, I think we might be in the same phase of having Phase III failures. So I would argue it may be more cost efficient to do highly targeted biomarker intensive studies for shorter periods of time than picking one, maybe the wrong biomarker, and doing, as Rusty already said, long 18 month clinical trials only to find out you're underpowered to see a clinical effect and you don't have all the biomarkers.

Audience Male: I tried to look from the –

Dan Perry: Excuse me; please identify yourself for the records so that we have this captured.

Audience Male: My name is Ming Tong. I used to work at NIH doing Phase I/Phase II at a clinical building 10. And I was in the industry and then I went back to Harvard Medical. You know I tried to look from the clinical utility perspective. You know I was really impressed by Mark's comment, you know sometimes we look at biomarker from a scientific interest or versus the [inaudible] or are we helping the patient you know monitoring the treatment course. And I still frustrated. You know what are we measuring and what are we treating?

You know do we make a distinction between upstream and downstream? I really puzzled. The reason is that look into the recent paper published in 2010. You know one of the papers actually is dated back a couple of years and in a single statement said that dimethyl tropic glutamine increase the production of tau protein. I was shocked. You know I thought we also know that, you know the new toxicity, you know downstream, but it turned out to be very upstream.

But my point I try to drive it and the question I should ask the panel is that can we look from the different paradigm? Can you look into for me dynamic network, you know from the system biology perspective, and then we can measure not just a single point we measure the trajectory, for instance in metabolite, in one of the circulating biomarkers, that we might help to as we say enhance the treatment monitoring. Could it be possible?

Actually this is not a fantasy. In Peoria College in London you know they have a wonderful computational program to measure, just from the urine, you know metabolite, the trajectory and not one single point. But the pattern change that what we already talk about the progression rate.

Dr. Sperling: Does anyone want to take that on?

Dr. Walton: I think that systems biology models are conceptually really great because it tells you how to monitor for a variety of different effects to compare what you're expecting a drug to do, to what it's actually doing and whether or not it's meeting your expectations. The problem of course is that it's systems biology models are very complex and we need to be able to have enough knowledge of all the physiology and pathophysiology that are important in the

disease in order to build a model that we can rely on. I'm not sure that we're there yet.

Dr. Siemers: From a practical standpoint, I mean in the future that would be great if you could get it to work. It's really difficult to come up with a biomarker or a group in a peripheral compartment. It's watery urine that mimics what happens centrally. So, in the process of drug development, even if you had that, now you're giving your drug usually systemically unless it's intercranially. And so then you'd have to really understand your systems biology to know that this just wasn't some peripheral drug effect but it actually was doing in the brain what you wanted it to do.

Dr. Sperling: I think we're almost done, but I'll pose one last question to Rusty and to the panel, which is I didn't hear anyone fully tell me how we're gonna get from biomarkers to clinical. And I think Jeremy's talk on better ways to do the ADAS-cog is encouraging, but I'm still concerned that there's a lag between what we can measure with cognition and these changes in biomarkers. And I think this is a place where the field needs help, especially the long stage before we get to clear impairment of MCI and MCI about to tip over into dementia.

So I don't know if there's anyone in the room who has thoughts on this or, Rusty, how you would help us from a regulatory point of view to think, but I feel like this is the sticking point.

Dr. Katz: Well again, as I mentioned, it is possible. Again, we haven't been faced with it so we haven't had to make a decision. But it is possible that the patients who are asymptomatic and who have a biomarker identifying them, that maybe again a sophisticated change, you know a subtle change on a sophisticated, very sensitive cognitive measure, maybe that's good enough.

I don't think we're at the point where, if you have a change just on a – I'm talking now about a definitive trial. Again, for Phase II, it may just be good enough to show an effect on the biomarker. And again, I'm not sure it has to reach statistical significance, which is an issue I'm particularly interested in in Phase II.

But for a definitive trial, I don't think we're at the point where, if you show an effect on a biomarker or even an array of biomarkers by itself with nothing clinical, that that's – that we're ready yet to approve a drug on such a basis. We could; the law permits it. It says that, if the effect on the biomarker is reasonably likely to predict a clinical benefit, we could do it and others do it in other fields. But I don't think there's a consensus yet – you tell me if

I'm wrong – amongst the community of experts that, if you change the PIB PET scan and nothing else happens to the patient, that we know that that's – it's reasonably likely that that will be a useful drug clinically. So I don't think that we're there yet.

You know we've advocated sort of moving backwards, showing – it shows some correlation between the effect on the biomarker and the effect on a cognitive measure in MCI, the mild patients. So maybe people who are just – who we've identified as being pre-Alzheimer but who we think are just about to convert to some clinical state that we could detect. And then sort of validating it, the biomarker, in that sort of setting and then say, okay, now, in patients who are truly asymptomatic, in effect just on the biomarker will be sufficient. But that hasn't happened.

And I think, until we have, as Mark has been saying and others have been saying, until we have a better idea, a much better idea about the pathophysiology of Alzheimer's disease, and how a particular drug works in all its positive and negative effects, I just don't think we're there yet.

Dan Perry: Very good. Well Dr. Katz, we appreciate that final thought as an appropriate place to end this. Let's begin by giving this panel a big round of applause.

[Applause]

And my only closing thought is that we needed a particularly auspicious moment because precisely one month from today, one second after midnight, Kathleen Casey-Kirshling of Cherry Hill, New Jersey, will turn age 65. And within next year, our population will go from producing some 6,000 seniors each and every day will be past age 65; we'll go from 6,000 a day to 10,000 a day next year and we will stay at that level for the next 18 years. And if what we know of Alzheimer's disease is that a large fraction of them unfortunately perhaps already have the pathology underway.

So the fuse is burning on the senior group and on the Alzheimer's epidemic and there couldn't be a more important thing for us to be doing today than to listen to these experts, to have this back and forth with the regulatory officers and to begin to get our act in order for this oncoming avalanche.

So I want to thank, first of all, all of our speakers, all of our panelists, our superb panel moderators. I want to thank the Food and Drug Administration, Dr. Katz, Dr. Walton, all of your colleagues who are here today to take part in this. I want to thank

our co-host, the Alzheimer's Foundation of America, the Alzheimer's Drug Discovery Foundation, all of the member organizations of the ACT-AD Coalition, the sponsors who make all of this possible.

And finally, I want to thank and identify Cynthia Bens, who is in the back of the room.

[Applause]

Cynthia is the much valued Director of Public Policy at the Alliance for Aging Research and the principal staff of the ACT-AD Coalition.

I wish you all the best for the yearend holidays – they're starting very soon – and for a safe trip back home. We will be posting many of the documents from today's meeting and as much of the material that we can digest and get on our website. And that website is ACT-AD.org. ACT-AD.org. Thank you all very much. Appreciate it.

[Applause]

[End of Audio]