FOOD AND DRUG ADMINISTRATION

CENTER FOR DRUG EVALUATION AND RESEARCH

SIXTY-FOURTH MEETING

OF THE

ONCOLOGIC DRUGS ADVISORY COMMITTEE

8:05 a.m.

Tuesday, December 14, 1999

Versailles Ballrooms
Holiday Inn
8120 Wisconsin Avenue
Bethesda, Maryland

ATTENDEES

COMMITTEE MEMBERS:

DEREK RAGHAVAN, M.D., PH.D., Acting Chair
Associate Director
Head of Medical Oncology
University of Southern California
Norris Comprehensive Cancer Center
1441 Eastlake Avenue, Room 3450
Los Angeles, California 90033

KAREN M. TEMPLETON-SOMERS, PH.D., Executive Secretary
Advisors & Consultants Staff, HFD-21
Food and Drug Administration
5600 Fishers Lane
Rockville, Maryland 20857

KATHY S. ALBAIN, M.D.
Professor of Medicine
Division of Hematology/Oncology
Loyola University Medical Center
Cancer Center, Room 109
2160 South First Avenue
Maywood, Illinois 60153

DOUGLAS W. BLAYNEY, M.D.
Medical Director, Oncology Program
The Robert and Beverly Lewis Family
Cancer Care Center
Pomona Valley Hospital Medical Center
1910 Royalty Drive
Pomona, California 91767

DAVID H. JOHNSON, M.D.
Director, Division of Medical Oncology
Department of Medicine
Vanderbilt University Medical School
1956 The Vanderbilt Clinic
Nashville, Tennessee 37232

DAVID P. KELSEN, M.D.
Chief, Gastrointestinal Oncology Service
Memorial Sloan-Kettering Cancer Center

1275 York Avenue
New York, New York 10021

ATTENDEES (Continued)

COMMITTEE MEMBERS: (Continued)

KIM A. MARGOLIN, M.D.
Staff Physician
Department of Medical Oncology and
Therapeutics Research
City of Hope National Medical Center
1500 East Duarte Road
Duarte, California 91010

STACY R. NERENSTONE, M.D.
Associate Clinical Professor
Oncology Associates, P.C.
Helen & Harry Gray Cancer Center
Hartford Hospital
85 Retreat Avenue
Hartford, Connecticut 06106

JODY L. PELUSI, F.N.P., PH.D., Consumer Representative
Cancer Program Coordinator
Maryvale Hospital
102 W. Campbell Avenue
Phoenix, Arizona 85031

VICTOR M. SANTANA, M.D.
Associate Professor
Department of Hematology/Oncology
The University of Tennessee
332 North Lauderdale
Memphis, Tennessee 38101

RICHARD M. SIMON, D.Sc.
Chief, Biometric Research Branch
National Cancer Institute
Executive Plaza North, Room 739
Bethesda, Maryland 20892

ATTENDEES (Continued)

COMMITTEE MEMBERS: (Continued)

GEORGE W. SLEDGE, JR., M.D.
Professor
Departments of Medicine and Pathology
Indiana University School of Medicine
Indiana Cancer Pavilion
535 Barnhill Drive, Room 473
Indianapolis, Indiana 46202

COMMITTEE CONSULTANTS:

RANDALL BRAND, M.D.
Assistant Professor
Section of Gastroenterology and Hepatology
University of Nebraska Medical Center
982000 Nebraska Medical Center
Omaha, Nebraska 68198-2000

RUSSELL JACOBY, M.D.
Director, Colon Cancer Prevention Program
University of Wisconsin Medical School
600 Highland Avenue
Madison, Wisconsin 53792

JAMES LEWIS, M.D. (non-voting)
Professor of Medicine
Division of Gastroenterology
Georgetown University Medical Center
3800 Reservoir Road, N.W.
Washington, D.C. 20007

CHRISTINA A. SURAWICZ, M.D.
Member, Gastrointestinal Drugs Advisory Committee
Harborview Medical Center
325 9th Avenue
Box 359773
Seattle, Washington 98104

PATIENT REPRESENTATIVES:

SALLIE FORMAN
Washington, D.C.

ATTENDEES (Continued)

FOOD AND DRUG ADMINISTRATION STAFF:

MARK AVIGAN, M.D.
JULIE BEITZ, M.D.
JUDY CHIAO, M.D.
ROBERT JUSTICE, M.D.
RICHARD PAZDUR, M.D.
ROBERT TEMPLE, M.D.
GRANT WILLIAMS, M.D.

ON BEHALF OF G.D. SEARLE & CO.:

MONICA BERTAGNOLLI, M.D.
GARY B. GORDON, M.D., PH.D.
ERNEST T. HAWK, M.D., M.P.H.
DAVID JORDAN, PH.D.
GARY KELLOFF, M.D.
JAMES LEFKOWITH, M.D., PH.D.
BERNARD LEVIN, M.D.
HENRY LYNCH, M.D.
JAIME L. MASFERRER, PH.D.
PHILIP NEEDLEMAN, PH.D.
ROBIN PHILLIPS, M.D.
RICHARD SPIVEY, PHARM.D., PH.D.
GIDEON STEINBACH, M.D.
MARINA WALLACE, M.D.

ATTENDEES (Continued)

ALSO PRESENT:

CAROLYN ALDIGE
KEVIN LEWIS
ABBY MEYERS
BETH SCHREIBER
PAT WEIDNER

C O N T E N T S

C O N T E N T S (Continued)

NDA 21-156, CELEBREX (celecoxib)
Indicated for the Reduction and Regression of
Adenomatous Colorectal Polyps in
Familial Adenomatous Polyposis Patients

P R O C E E D I N G S

(8:05 a.m.)

DR. RAGHAVAN: Well, good morning. I'd like to call the meeting to order.

We're doing something a little different for the first couple of hours of the morning. We have two experts who are going to be taking us through a relatively free-form discussion that relates to updating some of our concepts on the design and analysis of clinical trials.

What I'd like to do is just have the members introduce themselves. Then we'll have the conflict of interest statement, and then we'll get right to the presentations because I'd like to keep us to time.

DR. JOHNSON: David Johnson, medical oncologist, Vanderbilt University.

DR. SANTANA: Victor Santana, pediatric oncologist, St. Jude's Children's Research Hospital, Memphis.

DR. SLEDGE: George Sledge, medical oncologist, Indiana University.

MS. FORMAN: Sallie Forman, Patient

Representative.

DR. NERENSTONE: Stacy Nerenstone, medical oncology, Hartford, Connecticut.

DR. BLAYNEY: Douglas Blayney, medical oncologist, Wilshire Oncology Medical Group, Pomona, California.

DR. KELSEN: Dave Kelsen, medical oncologist, Sloan-Kettering, New York.

DR. PELUSI: Jody Pelusi, oncology nurse practitioner, Arizona, and Consumer Rep.

DR. RAGHAVAN: Derek Raghavan, medical oncologist, University of Southern California.

DR. TEMPLETON-SOMERS: Karen Somers, Executive Secretary to the committee, FDA.

DR. MARGOLIN: Kim Margolin, medical oncology and hematology, City of Hope, Los Angeles.

DR. SIMON: Richard Simon, biostatistician, National Cancer Institute.

DR. ALBAIN: Kathy Albain, medical oncology, Loyola University, Chicago.

DR. WILLIAMS: Grant Williams, medical team leader, FDA.

DR. BEITZ: Julie Beitz, medical team leader, FDA.

DR. PAZDUR: Richard Pazdur, Division Director, FDA.

DR. TEMPLE: Bob Temple, Office Director.

DR. TEMPLETON-SOMERS: You're in luck. This is the shortest conflict of interest statement ever.

The following announcement addresses the issue of conflict of interest with regard to this portion of the meeting and is made a part of the record to preclude even the appearance of such at this meeting.

The purpose of this meeting is to have a general scientific discussion concerning the design and analysis of active controlled clinical trials. Since the discussions are exclusive of any particular products or companies, it has been determined that no conflict of interest or the appearance of a conflict exists. Therefore, all committee participants may partake in these discussions.

Thank you.

DR. RAGHAVAN: So, our first speaker this morning I think is well-known to this audience from the

Food and Drug Administration, Dr. Robert Temple.

DR. TEMPLE: Well, good morning. I get to talk to you this morning about one of my favorite subjects. I've actually been writing occasionally on this since 1982, which I can barely believe myself.

I want to talk briefly and generally about active control equivalence or non-inferiority trials and the difficulties that they can pose.

I also want to talk about two very different endpoints used in cancer trials and the different implications they have, namely response rates and other things like survival or time to progression.

Then finally, I want to consider a particular problem in the analysis of survival when hazard ratios are used to make the comparison.

These slides were obviously prepared for a different talk.

(Laughter.)

DR. TEMPLE: So, don't worry about it.

There are three major problems when one is using an equivalence or non-inferiority design. I'm not going to talk about the third, but I am going to talk

about the first two.

The first is that there's a historical assumption, that is, an assumption based on experience, that the trial has the ability to show anything, and I'll describe that in more detail.

The second is that there can be a lack of incentives to carry out an excellent study because, as a general matter, sloppiness obscures differences, and finding no difference is the goal of these trials.

The epistemology of showing effectiveness really comes in two different flavors. One is a trial that shows a difference between two treatments. As long as you can be sure that the control treatment is not worse than nothing, the superiority of the test drug to the control, whether that's placebo, active, or lower dose, shows drug effect, and you don't have to really think further about that. You then have to ask whether the benefits outweigh the risks.

In an equivalence or non-inferiority trial, what one tries to do is latch on to the known activity of the active control, find no important difference between the two, and therefore conclude that the new

drug works also.

Until moderately recently, these studies were called equivalence trials and in a certain naive sense, you'd run the trial, show no significant difference between the two treatments and say, aha, mine works too. For various reasons, the modern way of doing this is to use a non-inferiority design, which is to say and I know the effect --

(Pause due to audio interruption.)

DR. TEMPLE: For various reasons -- and many people have written about this, the problem with an equivalence trial is if you merely make it too small, you win. That's an undesirable attribute.

So, what one now does is define the effect of the control drug and show that the new drug is not inferior by the size of the effect of the control drug. I'll show that in more detail.

I've basically already said this. The naive version was equivalence and now in the modern era we use non-inferiority designs. Basically what you do there is you specify a null hypothesis --

(Pause due to audio interruption.)

DR. TEMPLE: In a non-inferiority design, one identifies a margin, M, which is the effect that the control drug is more or less guaranteed to have in the study. Then you look at the confidence interval for the difference between the test drug and the control drug, and as long as that difference is not more than M, you can be sure that the new drug has some effect, that is, more than no effect. It doesn't tell you how big the effect is.

In a lot of cases, people are not happy to lose all of the effect. People would want to know that the new drug has more than any effect. They might want to know that it has at least 50 percent of the effect. So, the margin to be tested might be smaller than the full effect of the drug like 50 percent of it. If it's a mortality trial, you don't want to lose all but a little bit of the mortality effect. You want to preserve most of it.

The advantage of this design is that if the confidence interval is very wide because the sample size is too small, the study will not mislead you and declare non-inferiority because the confidence interval will be

wide and you won't be able to rule out your null hypothesis.

Now, this design solves the size problem, but it doesn't solve the assay sensitivity problem. It's perfectly possible -- well, you have to be able to know for sure that in this study the control drug had an effect of at least this size, M.

Just to show how this is done -- this M1 should have a line going across it too, so imagine a line -- if M1 is the effect of the control drug that you're sure the control drug had in that study -- and I'll talk later about how you might know that -- then you can imagine a few different outcomes.

On this axis is the difference between the control drug and the test drug. So, going up means the control drug is better. And this is the point estimate for the difference and this is the 95 percent confidence interval. These data obviously are all made up.

In this case here, the 95 percent confidence interval permits the control drug to be superior by more than 2, and since 2 is the whole effect of the control drug, that means the new drug may be inferior by as much

as the whole effect of the drug. So, in this case you don't know that the new drug has any effect at all.

M2 is 50 percent of M1. If you didn't want to lose up to all of the effect of the control drug, but thought that you have to have at least 50 percent of it, you'd set your margin at M2.

In that case, this study here shows that the difference between the control drug and the new drug is less than 1, so you would declare that this drug is effective. If you could not be sure that in this study -- I'll talk a little bit later about why you might not be sure -- the control drug had an effect of at least M1 or at least M2, then the only time you can reach a conclusion is if the new drug is actually better than the control drug, in which case the difference between them has to be less than 0. In that case, only this study is informative because in this study the new drug is superior to the placebo. The 95 percent confidence interval is better than no difference.

This example here shows a study where the point estimate actually favors the new drug, but the confidence interval is so wide that you haven't ruled

out a loss of all of the effect. So, this study would be uninformative.

When you do an equivalence or a non-inferiority trial, there's always the question, did the active control drug have an effect of the size expected in the trial that was carried out? If it didn't, if for some reason it didn't have that effect, then equivalence or non-inferiority is completely meaningless because the equivalent or non-inferior drug could have no effect at all.

So, all of these trials are based on an assumption. It's what you could call a historical assumption, and that is that the active control was effective in the particular study in question. As I'll show you a few examples, that is just not always true, and one of the reasons that it's not always true is what you saw yesterday. You saw two trials of fairly similar design. One gave a fairly robust evidence that Taxotere was useful in non-small cell lung cancer. The other trial didn't show a statistically significant difference on the same measurement. So, if that second trial had been an active control trial and you saw no difference

between Taxotere and the new drug, you would have learned nothing at all because Taxotere in that study wasn't better than therapy we think has no effect.

Anyway, even if a drug is known to be effective, it doesn't show its effectiveness in every trial, and in an active control trial, you don't test the question of whether the control drug is effective. There isn't any placebo or no treatment control. So, you have to deduce it from some other information. What that means is that an active control study has some elements of a historically controlled study.

The ability of a particular trial to show a difference of a specified size between treatments has been called assay sensitivity. There's an international guideline under development on this question and that's the term that's used there.

The ability of a trial to show a difference is affected by a lot of things. It could be the ability of the population to respond. It could be whether -- well, this doesn't apply to cancer, but in some cases whether you've excluded placebo responders and included potential responders that may be important.

Sometimes the quality of the study or the precision of the measurement is important. For example, if you were looking at time to progression and measure infrequently, like every 2 months, you reduce the ability to distinguish between the activity of drugs because a wide range of progression times gets lumped in under the same category. And, of course, if the study is too small, it might not be able to detect a difference.

It's worth remembering again that in a trial showing a difference between treatments, the question of assay sensitivity takes care of itself. A successful trial had it. An unsuccessful may or may not have, but you don't reach the wrong conclusion.

In a non-inferiority trial, assay sensitivity isn't measured, and it has to be deduced. The way you deduce it is you look at historical experience showing that well done trials have a property called sensitivity to drug effects. That means they can tell the difference between active drugs and inactive drugs. And then one looks very closely at the present study, making sure that the patient population and other conditions

are similar to those in which the drug was able to show an effect.

So, the term "sensitivity to drug effects" is a historically based conclusion that appropriately designed, sized, and conducted trials in a particular disease with a specific active drug, or sometimes a group of related drugs, can reliably show an effect of at least some defined size on a particular endpoint.

Usually one shows this by showing that adequately powered, sized, and well-conducted trials in a specific population can regularly distinguish active drugs from placebo or from best available care or some other group for particular endpoints.

Just for the terminology, "sensitivity to drug effects," is an abstract conclusion about well-conducted trials. Assay sensitivity is a conclusion about a particular trial.

I don't have any oncology examples, although as I said, we sort of had one yesterday. But I just wanted to illustrate how drugs that we know are effective don't show their effectiveness all the time, and many of these examples are from psychotropic drugs

because they're rather difficult to show.

This is a slide showing six trials along here of an antidepressant. It's called "new" but it's really a drug called nomafensine, which was effective but causes hemolytic anemia and is no longer marketed. It was a comparison with imipramine, a tricyclic antidepressant whose effectiveness is not in doubt. And there was also a placebo group, but I'm not going to show you that till the next slide.

The measurement used was the Hamilton Depression Score at 4 weeks, a standard measurement for antidepressants.

These trials were analyzed by getting a common baseline for the two groups. That's not really important.

And this is the change in 4 weeks. You can see that the change is about 10, 13, 9. Those are reasonably sized changes in Hamilton Depression Score for an antidepressant. There's plainly no significant difference between these, but what I want to point out is that the HAMD scores at 4 weeks are virtually identical, 13.4, 12.8, 13, 13, 19.4, 20.3. There is

really no difference. If you're a believer in equivalence trials, this ought to be persuasive.

You will note that some of the trials are very tiny which limits their believability, but some of them are decent size, of the size typically used in depression trials.

Now you see the third group, the placebo group. What you can see is that five out of the six trials can't distinguish anything from anything. These numbers are essentially the same. Sometimes placebo is actually a little better. Only one trial, the smallest, too small to be remotely credible actually, showed that it had assay sensitivity, showed any ability to distinguish active from inactive drugs. This was true for both the new drug and imipramine. So, that's just an illustration.

I won't dwell on this, but I went back and looked at three years' worth of psychotropic drug experience, and about a third of all trials that seemed to be well designed and that are of drugs we know to be active, because many of their trials do show effectiveness, can't show it in any given study. I

think I won't go through each of those.

There's a number of settings, depression, anxiety, dementia, symptomatic heart failure, seasonal allergies, GERD -- GERD is very hard -- post-infarction beta blockade. Only 5 out of about 30 well-designed and pretty good size studies actually show the improvement in survival that we know exists. Even post-infarction aspirin, which we know is effective, isn't effective all the time. The largest trial ever conducted, the AMOS trial, actually leaned the wrong way. One could certainly add many oncology studies to this also.

A second, somewhat separate problem -- I don't know how much of a problem it is in oncology. I'd to think about it -- is that incentives to study excellence are not as strong in a trial intended to show no difference between treatments as in a trial intended to show a difference between treatments because as a general matter, sloppiness obscures differences, and the goal of a non-inferiority trial is not to see a difference. You don't have to be particularly cynical to think that is not a good incentive to give people, and I think I won't dwell on that too much more.

Now, there are situations in which active controls or non-inferiority trials are perfectly credible. Sometimes that's because the difference between no treatment and treatment is very large and very obvious. Infections, for the most part, don't cure themselves. The cure rate in an infectious disease trial can more or less be taken as the entire effect of the drug.

And it's important that response rate in oncology has similar properties. On the whole, tumors don't shrink by 50 percent by themselves. So, the response rate is pretty believable even without a control group. The typical phase II cancer study relies on that fact, and one believes the response rate.

Thrombolytics have been consistently superior to placebo. Deep vein thrombosis. It's fairly obvious. You can tell the difference between a highly responsive tumor whether you're looking at response rate or at survival.

Anesthetic agents. Most people stay awake until they're treated.

And things like beta agonists and

bronchospasm. They produce an immediate 15 or 20 or 30 percent improvement in FEV-1. It's fairly obvious that's not a response to placebo.

So, if one wanted to support an active control trial in, say, an oncologic setting, the first thing you'd have to do is review known placebo-controlled -- placebo not so usual, but no treatment controlled or add-on studies, or whatever you have to show a fairly regular ability to distinguish active drug from placebo.

Then one has to use pretty much the same population, same stage, et cetera to know that the data that you have in the past applies to that group.

One has to define the margin. You have to estimate a size of the survival advantage or whatever it is so that you can design the margin and then show that inferiority of the new drug by more than that is ruled out. And this can be very difficult to do, so it's worth spending a few moments to think about what one can do in a setting where it's very hard to say what the margin is, which is true in a lot of solid tumor situations.

Probably the best thing -- and we certainly urge this a lot in the meetings we have with sponsors -- is to add the new therapy to whatever the standard therapy is and produce a trial that shows a difference, not one that's intended to show similarity. That works pretty well if the drugs are different pharmacologically and have a different mechanism. This is now standard in anti-epileptic drug development and congestive heart failure. But it doesn't help you if what you're developing is a new dosage form or packaging the same drug in a liposome because there's no particular reason to think it will be superior, and you may or may not be able to do this.

Beating standard therapy is always good. I don't think that needs more discussion.

Sometimes a dose-response study can be informative, but it really has to show a positive slope. It's probably not acceptable to use a deliberately inadequate dose. That's just like using a placebo which in many oncologic cases isn't acceptable.

Sometimes you can study a population subset not known to benefit from standard therapy and use a

difference-showing trial, a placebo or no-treatment trial.

I wanted to spend a minute or two on a particular problem associated with at least some recent analyses of cancer trials.

Remember the first step in an equivalence trial or non-inferiority trial is to identify the margin, the size of the effect on survival or whatever the right measurement is. It's perfectly all right to use an absolute value like 4 months or to use the reduction in risk. Either of those is perfectly sensible.

For example, if the active control is known to increase survival by 4 months, one wants to make sure that the 95 percent confidence interval for the difference between the control and the test drug should exclude a difference of more than 4 months. Then you know you have some effect.

If losing all of the benefit of the control is unacceptable, you might say, well, we have to rely on a difference of more than 2 months or more than 1 month. That would preserve 75 percent of the effect of the

drug.

Now, suppose the control was known to give a 20 percent reduction in hazard rate. One of the problems with an absolute value is that you might not expect the difference to be the same from one population to another. So, you might try to choose the improvement in hazard rate. So, that's okay. Then the 95 percent confidence interval for the difference in hazard rate should exclude a difference of more than 20 percent or more than 10 percent if one wanted to preserve 50 percent of the effect of the drug.

The problem is that the hazard rates comparison needs to be based on the effect of the control drug, not on overall survival. If survival is 10 months, the drug is not responsible for all of that. It's only responsible for a certain part of it. So, assuring that there's no difference greater than 20 percent in survival is of no value if the control doesn't have a 20 percent effect. For example, if a drug gives a 2-month survival increase, say, 12 to 14 months, producing a hazard ratio of about .86 -- I'm sure that's not calculated properly. You've got to bear

with a nonstatistician -- then ruling out a loss of 20 percent doesn't rule out a loss of all of the drug effect. So, if hazard ratio is going to be used, it has to be focused on the actual effect the drug has on hazard ratio, not on overall survival. The same would be true for time to progression.

It's worth noting that this is not a problem when one is dealing with response rates because essentially all of the response rate can be attributable to the drug. So, you can just go about your business as usual.

This is mainly a problem when you're looking at survival effects, and given the relatively small effects in solid tumors, this is a really daunting challenge. It's very hard to rule that out unless the new agent is better. Then you can do it.

Now, it's extremely difficult in many cases to make the non-inferiority case credibly because the effect of the control is often very small. But we nonetheless tend to look at and accept non-inferiority designs and one question is why. Why do we find these believable? And I don't know the answer, but I'm going

to speculate. I'm not here to agree or disagree with this.

But I think we tend to believe in the response rate. That's measurable. That's credible. You can compare response rates without much of these anxieties because the response rate is pretty much entirely attributable to the drug. So that faced with equivalence studies that are not all that persuasive, we at least have the objective fact that these agents shrink tumors, and that perhaps gives us more confidence than the mortality results themselves do. That may be entirely reasonable.

I'm particularly concerned about therapies that work the same system, that have the same mechanism, that are the same drug packaged in a different way. It's probably not reasonable to think that they will be superior to the control agent. They may be better tolerated or easier to give. It may be that intuitive reliance on response rate is reasonable under the circumstances, but I think it's worth seeing if that's what we really think and acknowledging it and discussing it if that's really the basis for our confidence in many

of these therapies.

So, let me turn this over to Rich Simon who has written recently about how one can actually reach the conclusion that a therapy compared to an active control actually has some effect, an area that has had relatively little research up to now.

Questions now or later?

DR. RAGHAVAN: Just before Richard Simon speaks, are there any questions of fact or interpretation that the committee would like to address to Dr. Temple? Dr. Sledge?

DR. SLEDGE: Actually I have a question from a statistical size standpoint. If you are looking at a point estimate sort of thing, like either a response rate such as we saw yesterday, a 1-year survival rate, as compared to something like a log rank test, are there going to be major differences in sample size in an equivalence trial based on what endpoint you look at?

DR. TEMPLE: Rich, I am sure, will have a better answer.

I don't think there's any inherent difference in those two except that when you're looking at hazard

ratio, log rank, or survival, much of that survival can't be attributed to the drug. So, only a small part of the total survival is related to the drug. The difference between the two treatments is. In response rates, the entire effect pretty much is due to the drug. Also the rates tend to be higher. So, I'm sure you can get away with smaller sample sizes when you're trying to compare response rates, but Rich may want to answer that.

DR. SIMON: I think, for example, if you take the example yesterday, I think it would be problematic to deal with response rates because the response rates were so low that it looked like that the effect on survival was not mitigated through response rates. Therefore, we could say, well, we could look at stable disease plus response rates, but I think what it means is we don't really understand what the relationship in the example yesterday was between tumor shrinkage and effect on survival. So, I think we're sort of forced into dealing with survival.

But I think we need to distinguish between log rank tests, which is sort of a significance test --

and I think significance tests actually almost confuse the issue when you're dealing with equivalence more than they clarify it -- and the effect on survival.

I'll try to clarify that in my talk.

DR. RAGHAVAN: Bob, how do you bring into the discussion the issue of clinical relevance? For example, if you think about some of the data that have been presented in well-powered, randomized trials looking at the impact of combined androgen blockade versus monotherapy in prostate cancer, I think there's little doubt that there is a statistically significant difference between curves that have compared 500 versus 500 cases. In absolute terms, the difference in median survival is of the order of 4 to 6 weeks in a disease with a long natural history, and at 5 and 10 years, the percentage survival difference is about 2 to 3 percent.

So, how do you factor in, once you get beyond the statistics, in terms of looking at equivalence or lack of it? If you have a new, less toxic regimen compared to a standard toxic regimen, you identify that there is a small difference in success in terms people

alive, but in absolute terms, it's small. How do you factor that in to a logical discussion?

DR. TEMPLE: The difficulty there is to know what people really are valuing. I think sometimes people are content to see less toxicity, and they almost don't care what the effect is because it's so small. That's a funny basis for approving a new therapy: At least it doesn't make you sick. It's not really what the law says one is supposed to do. One is supposed to decide on whether it's actually effective or not.

So, the order of reasoning I think has to be what is the evidence that this agent is effective. And that raises all the problems that equivalence trials have. In many of those situations that you describe, the evidence of effectiveness is pretty modest anyway and the evidence that in a particular trial you could distinguish therapy from no therapy at all is very weak. That makes equivalence testing in that setting very difficult, and I'm very uncomfortable with that because I would like a way to decrease the toxicity of therapies.

I think the answer we actually intuitively

reach, even though what Rich says is absolutely true, is we look at response rates, stabilization rates, and we say, well, if those are similar, I'm probably okay even though I don't have a very good assessment of survival. I think the question that needs to be discussed is whether that's good enough because it's a dubious basis for concluding that there's a favorable, useful effect. Even if the therapy is better tolerated, if it wasn't doing anything, the best tolerated therapy is no therapy.

So, my intent is to raise these issues. I don't know the answers.

DR. RAGHAVAN: Dr. Nerenstone?

DR. NERENSTONE: I think it's even more complicated than that because if you look at the response rates that have been proposed in some of these equivalence trials, they're grossly overestimated, at least in a number of drugs. The one the comes to mind in particular was epirubicin where what they thought was going to be the response rate and looking at equivalence was something like double what the actual response rate was in the trial. I think the companies have -- when

you said they're sort of induced to be sloppy because they don't find a difference, there's sort of an inducement to overestimate it. So, your sample size is smaller to predict an equivalence study when in fact, the numbers are too small because the actual response rate is much lower.

DR. WILLIAMS: Stacy, you mean Evacet probably. Right?

DR. NERENSTONE: Sorry. Evacet, yes.

DR. TEMPLE: Well, there are potential protections against that. It certainly does seem true that phase II studies tend to show higher response rates than are seen ultimately. The reason for that certainly isn't obvious to me.

One way to do that is to have a control group not so much to compare the two agents, but to get some idea of what the population response is.

Of course, the best way to interpret the results is to make sure that the people who evaluate them are blind to therapy. We say that a lot, but it happens infrequently. It's still a good practice. It would be a good practice in oncologic trials, but it's

the devil to make it happen.

DR. RAGHAVAN: Dr. Blayney?

DR. BLAYNEY: You used the term what the law says is effective. In these low response rates and minimally effective solid tumor drugs that you allude to, is it more effective than no therapy, more effective than the standard therapy, or as effective as the standard therapy?

DR. TEMPLE: Well, in general, the law doesn't have a comparative efficacy requirement so that if the effect is clinically meaningful and is adequately described, that's probably okay.

Now, the exception to that is where the standard therapy is known to do something very important, like improve survival. You would not, without some good reason, want to have a therapy that's less effective, you know, in leukemia or something like that, than the standard therapy. So, the comparative effectiveness would be important.

Equivalent efficacy is fine. The question is whether the trials we call equivalence trials actually show that. I think in many cases in solid tumors it

would be hard to allege that they do if survival is the endpoint. They may well be pretty good at showing equivalent response rates, whatever that may mean.

DR. BLAYNEY: It seems to me, as the example Dr. Raghavan raised, the issue hinges upon what's a good reason, and that's what we end up spending a lot of time debating. Is less toxicity a good reason to fudge or blur the precision of the statistics that you're trying to tell us about?

DR. TEMPLE: Well, we would say that you have to be able to conclude that the drug is effective. Now, what effectiveness means, how you measure it, those are matters for expert opinion and judgment. Well, as you know, in refractory disease, we have said that we will accept response rate as a basis for accelerated approval. It's controversial. I would say the Europeans currently are not doing that, but we are with a requirement that subsequent studies be carried out to show that there is a really clinically meaningful benefit. Those are matters of judgment.

In the distant past, we accepted response rate all the time until a previous version of this

committee said, wait a minute, don't do that. You should have some clinically meaningful result. That has gone back and forth. Those are matters of judgment.

It needs to be debated, but the conclusion that a hormonal therapy should be judged according to its response rate, which actually is no so different from what we've said, is not crazy. That reflects the lesser toxicity of those treatments, and people could make a reasoned judgment that that's okay. They would have to know that that doesn't mean it necessarily improves survival, but if that's an acceptable endpoint and we conclude that it is, you can probably compare therapies for that response rate and then the fact that one is less toxic becomes interesting.

DR. BLAYNEY: And the struggle around quality of life and its definition I think is, as I understand it, a better way to quantify what a good reason is.

DR. TEMPLE: A documented improvement in quality of life, compared to a control, would generally be taken as evidence of value. It's important, however, to distinguish between lesser toxicity as a contribution to quality of life, which has nothing to do with

effectiveness, and tumor-specific findings, improved weight gain. Those things are evidence of effectiveness rarely seen in the trials we see, but we would certainly accept those. Drugs for prostate cancer have been approved because they reduce pain, need for analgesics. Those are, when you see them, relatively easy. Those seem like fairly obvious clinical benefits, and you weigh them against the toxicity.

It's when you see no difference between treatments and you really have very little assurance of what the active control did in this particular study, that you're up a creek.

DR. RAGHAVAN: The difficulty, though, with quality of life assays -- and that's I guess why you're having a workshop on this in the near future -- is that the robustness of measurement hasn't really stood the test of time. I've recently sat on an external advisory board, a data monitoring committee to a trial, in which all the objective measures of quality of life that physicians would deem important like performance status, and weight gain and things like that favored a particular combination chemotherapy regimen, but the

patient assessments of their own quality of life in linear analogue scales went the other way.

I mean, that's going to come up again and again as we try to understand in that area that's hard to quantify who's right. Intuitively one believes the patient must be right, but it may be that the instrument is wrong. So, as soon as the FDA takes aboard quality of life measurements, which I am sure they should do, it then opens up a whole new kind of area of potential imprecision as we're learning the methodology, where we know our statistical methodology.

DR. TEMPLE: We do accept quality of life instruments. Some of this discussion came up yesterday. The global quality of life, which measures psychiatric and social function, in addition to physical function, are extremely difficult to win on, and often the people who enter the trials are particularly impaired in those domains. Well, that's like studying a cancer drug in people who don't have cancer. No one would ever do that. But in quality of life places, nobody checks to see that there's an abnormality at baseline. So, the fact that they hardly ever win isn't too surprising.

We've urged for many years that some attempt be made to quantify the miseries of patients at baseline and then specifically test to see whether those improve. You could actually in some sense put them on the same visual analogue scale I think. You could have weight loss, appetite, pain over tumor sites, and a wide variety of things and see if treatments alter those. There has been very little attention to those kinds of things. If you don't do that, you're trying to show improvement in somebody who may not even have the condition, which is destined to fail.

DR. RAGHAVAN: Dr. Kelsen.

DR. KELSEN: We wrestle in groups and institutionally on the difference between phase II and phase III. We basically have had the feeling, or at least our groups have, that a phase II trial with a comparator arm -- and you can't really draw a lot of conclusions when you do randomized phase II's. We're facing this right now because of small sample size, et cetera. I'm getting the feeling that perhaps that paradigm might be a little bit changing. I'm not quite sure what you're saying. It's always good to have a

comparator. I agree.

So, I have two new treatments for a given disease. They're both experimental. If I have a third arm, like the octopus arm, or a fifth arm, I'm going to have really small groups of patients because we don't want to study 5,000 patients. We're going to have 50 to 100 patients in an arm. We sit down with our statistician. They say, well, you really can't draw any long, major conclusions about this. On the other hand, we're investing a lot of energy.

So, does it pay for us to rethink the way we look at phase II trials in the current circumstance?

DR. TEMPLE: Well, I really think so, but see, I'm not primarily in the oncology business. The only arena in which that practice continues I would say is oncology. I think we should leave that for Rich because he has actually written about that very thing.

Phase II trials in oncology have multiple purposes. They're not, strictly speaking, even when they're controlled, designed to show that one therapy is better than another. They're designed to help you plan your next study, and that is different. It does mean

the studies may not be very persuasive as evidence of effectiveness, but that's okay since that wasn't their purpose. But Rich has actually just put all that in nice article.

DR. RAGHAVAN: Last question, Dr. Blayney.

DR. BLAYNEY: The other question in effectiveness is in oncology we're stuck with legacy treatments, and I think we, in an equivalence trial, have the worry of setting a bar too low. I'm thinking specifically of the drugs looking at various comparators of 5-FU/leucovorin, which when given on a certain schedule, it's almost guaranteed to produce toxicity and many people have abandoned that daily times 5 schedule because of that toxicity, but yet that always appears in these comparator arms of the equivalence trial. Many of the drugs that have FDA approval, 10 or 15 years old, for various conditions now that are hardly ever used, I think when those are used as a comparator arm, the bar is set much too low and I don't know a way around that.

DR. TEMPLE: Well, the way to get around it is to conclude that you have to be better to be interpretable. Then it's okay if the bar is too low.

Then at least you know you're doing something.

The fact that a drug is FDA-approved or effective doesn't make it a suitable active control. Tricyclic antidepressants all work. They work just as well as the modern antidepressants, but an active control trial comparing a new drug with a tricyclic is totally uninformative or a new anxiolytic or a new drug for heart failure. We just have many, many examples of where drugs we know are effective -- the reason we know they're effective is they come up effective more than the predicted 1 in 20 or thereabouts. They come out effective sort of half the time, which is enough to show that they're effective. But half the time isn't good enough to make them a very good active control. So, our prior conclusion that something is effective, even if it's correct, may or may not mean that the drug is a suitable active control for a mortality study because the mortality effect might not be seen in that.

We recently reviewed fluorouracil results and the improved survival varies from half a month to 3 months or 4 months. What does that mean in any given trial? Was this one where the effect was half a month,

in which case the equivalence trial was uninformative, or was it 3 or 4 months, in which case the equivalence trial might be informative. And there isn't any way to know, unless Rich tells us how to know.

DR. RAGHAVAN: That's a good introduction.

DR. TEMPLE: That's what he's going to talk about.

DR. RAGHAVAN: Dr. Richard Simon who is an ODAC member, also from the National Cancer Institute.

DR. SIMON: Good morning. I'm going to basically discuss a paper I published about active control trials, therapeutic equivalence trials, but I'll preface it with some general remarks. Most of what I will say really will reinforce the things that Dr. Temple has said, although there are some new areas and some areas of minor differences.

I've encountered basically two kinds of therapeutic equivalence or active control trials, two somewhat different objectives.

There's a set of trials where you have a treatment that is very effective and you have some variant of that treatment which may represent shorter-

term delivery, shorter-length delivery of a drug, a more convenient type of administration of a drug, but where our active control -- there is a substantial body of evidence that it really is effective and you want to know whether the new regimen really is equivalent or close to equivalent to this active control.

Then there's the other situation where really what you want to do is establish that your drug is effective relative to, say, no treatment, but because there is some effective treatment for that disease, you feel like you can't do your randomized clinical trial with a no-treatment control group. So, what you're going to try to do is use some active control group, show that your drug is equivalent to that active control, and thereby indirectly claim that since the active control was better than nothing, that there was presumably some body of information that demonstrates that, that therefore, since your new drug is equivalent to this active control, your new drug must be better than no treatment.

These are, I think, two different situations.

For the most part, the first situation, where

you're dealing with something that's very effective, I think you can accomplish a successful therapeutic equivalence trial.

The second situation tends to be much more problematic because in many cases we're dealing with an active control which is not very effective and we'll see what sort of problems that gets us into.

In many cases in this latter situation, particularly if you're dealing with an active control which is not very effective or not consistently effective, everyone is much better off I think if you can wind up doing a superiority trial against a no-treatment control, such as what we saw yesterday with trial 17 in the Taxotere example.

The basic problem with equivalence trials is it's impossible to demonstrate equivalence. Generally in science we can show that if a body of data is not consistent with some hypothesis, then that's a basis for rejecting that hypothesis. The fact that a body of data is consistent with that hypothesis is not really evidence that that hypothesis is true.

At best really all we can do is that the

results are consistent with differences in efficacy between, say, a new treatment and a control treatment, consistent with differences in efficacy within specified limits. That puts us into the realm of what should those limits be. That's often very problematic.

One of the problems also with active control trials or therapeutic equivalence trials is the old saying, when your only tool is a hammer, everything looks like a nail. For evaluating clinical trials, very often significance testing is our hammer, and we try to put everything into a significance testing context. The problem with that is that it puts us into a binary way of thinking which leads us to believe that if we don't reject some hypothesis, then that hypothesis should be accepted. So, if we can't reject the null hypothesis that two treatments are equivalent, then that hypothesis must be true. And that's difficult because we may not have rejected that hypothesis because the sample size was too small or some other reason.

We're generally better off in thinking about active control trials thinking in terms of confidence limits, what sort of differences in efficacy are

consistent with the data at hand rather than testing hypotheses.

Another difficulty with equivalence trials is that large sample sizes are needed to establish that the differences in efficacy are within narrow limits. If our active control is not very effective compared to, say, no treatment, then we're going to have to establish that the difference between our new treatment and our active control -- we're going to have to establish that those things are equivalent within very narrow limits. That's going to really lead to very large trials, and that's often, in some cases, not doable.

Now, the limits to which the difference inefficacy should be bounded really depends on two things. One is the degree of effectiveness of the active control, and the second is the precision with which the effectiveness of the active control is estimated. Bob Temple gave an example in which he said, well, let's suppose that the active control adds 4 months of survival. If our point estimate is that it's 4 months, then there's some confidence interval around that, and if that active control was shown to be

significantly better than nothing, at a borderline p of .05, that means that confidence interval ranges from about 0 out to something in excess of 4.

If we really want to assure that we're not losing all of the effectiveness of our active control in our equivalence trial, we have to take into consideration the uncertainty with which not only what we think is our best estimate of the effectiveness of the active control, but also how it varies from trial to trial and our uncertainty in estimating what it is.

In general, I think therapeutic equivalence trials are not really feasible and they're not really interpretable. They're not really appropriate unless there's really a strong and quantifiable body of evidence for the effectiveness of the active control. So, if we say, well, we don't really have this body of evidence, well, if you don't have it, then you probably ought not to be doing a therapeutic equivalence trial because it means we're not really going to be able to interpret it very clearly.

One of the current criteria I think that's used for trying to demonstrate effectiveness is that the

confidence interval for the difference between the active control and the new regimen -- and I'm going to use a simple E for the experimental treatment and C for the active control -- is that confidence interval assure us that we lose, at most, 20 percent of the effectiveness on some sort of scale.

If we're dealing with a very effective control C and that if that effectiveness is consistently demonstratable, then this might work pretty well. But if we're dealing with a less effective C, then this may not work well. So, this is a rule of thumb which in itself doesn't really take into account how effective our active control C is or it doesn't take into account also the precision with which that estimate is determined.

Incidentally in your folders, I have actually a copy of all of these slides.

So, this was a reference to the paper that I published. In that paper, I tried to go through with some statistical rigor, in some generality, sort of the approach I'm going to try to present here in a more heuristic sort of way.

This is an application of Bayesian statistics. Sometimes I say this is an application of Ayesian tatistics because it's Bayesian statistics without the BS.

(Laughter.)

DR. SIMON: There's a lot of argument about Bayesian statistics on philosophical bases and a lot of controversy about Bayesian statistics because very often there's a subjective nature to it.

What I've tried to do here is to take a problem where I think by necessity we have to bring in external evidence in order to interpret an active control of clinical trials. So, I think it's a situation where, by quantifying what that evidence is, we're better off because we have to bring in external evidence one way or the other, and it's better to sort of have it all on the table.

So, we're going to be talking about two parameters. One parameter, which I'll call beta, sort of in a survival situation will represent the logarithm of the hazard ratio of our active control relative to either no treatment or placebo treatment or whatever was

the previous standard prior to the establishment of C as an effective treatment, what C is effective relative to. So, beta is that parameter.

There is some body of data, there are some clinical trials, presumably that were done, that established that C is better than P.

Then there's another parameter gamma that would sort of represent the log hazard ratio of our new experimental treatment E relative to P, but we're doing an active controlled trial in which what we're really going to estimate is the difference between beta and gamma because we're saying we cannot do the direct trial comparing our new treatment E to P.

So, we're going to obtain an estimate of this difference parameter, beta minus gamma, and we're going to use our previous data about the effectiveness of C relative to P in order to infer something about gamma in order to infer the effectiveness of E relative to P.

So, we're going to talk about a prior distribution for beta, but we're not going to pick this prior distribution out of our hat. We're going to use a normal distribution with some mean and some variance,

but it should be determined from the randomized clinical trials that were done comparing the active control C to P. If there were no such trials, then we really have to ask ourselves whether we should be doing a therapeutic equivalence trial.

In the cases where I've applied this approach to real clinical situations, I've used a meta-analysis of the randomized clinical trials that had been performed comparing C to P. In some cases there may be only a single clinical trial, and if there is more than a single clinical trial, then this is a step that has to be done very carefully in terms of determining what are the relevant clinical trials that are relevant with regard to patient population for the therapeutic equivalence trial we're going to do. For those clinical trials, we need to do a meta-analysis that tells us not only what is the average effect of C relative to P, but also how does that effect vary among trials.

If that does vary substantially among trials, then in the interpretation of our therapeutic equivalence trial, we have to represent that we have actually a very poor estimate, a very variable estimate

of the effectiveness of our active control. And that will play a substantial part of the analysis. It means that we will wind up with not a very good estimate of the effectiveness of our new treatment relative to P because it will mean we're in a situation where we really don't know very much about the effectiveness of our active control relative to P.

So, we do this meta-analysis and the results of the meta-analysis are a mean effectiveness of C relative to P and a measure of variation of that effectiveness across the trials.

We also have a prior distribution for gamma. In general, I have assumed that we have no real hard data. We have no clinical trials comparing our new treatment E to P. If we did have such clinical trials, we probably wouldn't be talking about doing a therapeutic equivalence trial. So, I've used a prior distribution with a variance of infinity, meaning I have no information really. All levels of effectiveness or ineffectiveness are a priori equally likely.

Then what we do is we do a therapeutic equivalence trial, and in this situation where we're

talking about survival as an endpoint, we obtain some maximum likelihood estimate of the log hazard ratio of E relative to C. I use a simple y to indicate that maximum likelihood estimate of that log hazard ratio, and that has some standard error, sigma.

Y divided by sigma is what we usually think of as sort of a z value. If this thing is greater than 2 or less than minus 2, since it has an approximate normal distribution, then we would be getting a statistically significant difference between E and C. And if y is 0, since it's a log hazard ratio, that means that the survival curves for E and C in this therapeutic equivalence trial are coming out right on top of each other.

I've defined hazard ratio here so that a negative y means that E did better than C. The survival of E was better than C in this therapeutic equivalence trial, and positive means that C did somewhat better than E.

From these two things, the prior distribution and the results of our therapeutic equivalence trial, we can calculate the posterior distribution of our

parameters beta and gamma. Under these conditions, the posterior distribution of beta is the same as the prior distribution because we have not really added any information about the effectiveness of C relative P. So, that distribution doesn't change.

But we can now infer something about the effectiveness of E relative to P. And we can summarize what we know about that in a normal posterior distribution. In this simple situation here, it turns out that the mean of that normal distribution is y plus mu. Intuitively it's the log hazard of E relative to C plus the log hazard from our prior distribution of C relative to P. And the variance of that posterior distribution is really the variance that we get from our therapeutic equivalence trial plus the variance that we had from our prior distribution. So, we can actually also calculate what the correlation in the posterior distribution is.

From this posterior distribution we can now calculate certain things. It doesn't really show up very well. But I don't really want to go through this in detail, but I want to say what the two things I'm

calculating are and what they depend on.

One of the things I'm calculating is the probability that gamma is less than 0. Now, gamma represents log hazard of E relative to P. So, this is the posterior probability that our new treatment E is more effective than, say, no treatment, or whatever P represents.

We can also calculate things like the posterior probability that beta is less than 0 which means that C is more effective than P and gamma is less than half of beta. So, this represents that E is at least 50 percent as effective as C. Things work in the negative direction here because we're dealing with sort of log hazard ratios in which we've measured it sort of in a direction so that negative represents sort of effectiveness relative to P.

So, basically we would be getting some data from our trial. The data is summarized in Y and sigma. From that data, combined with our prior distribution, we compute the probability that our new treatment is effective relative to P and that it's at least, in this case for example, 50 percent as effective as C, and that

C itself is effective.

Those quantities will depend roughly on three things.

One is the results of our equivalence trial, which I've summarized as y over sigma. If it's 0, it means that in our equivalence trial, C and E were sort of right on top of each other. If it's positive, it means that C came out better than E, and negative means that E came out better than C, although this is a ratio of one standard deviation in either direction.

The second thing it depends on is the strength of evidence that C was effective, from our prior distribution from the previous trials, relative to P. I've looked at two situations here, one where we sort of have borderline effectiveness, that the mean of C relative to P is two standard deviations away from 0, and then in a situation where I assume it's 3 where we have a much stronger body of data for the effectiveness of C.

And the third thing it depends on is the ratio of sample sizes or the sample size of our equivalence trial relative to the effective sample size

of this previous body of data. This is sort of a general way of showing how the results depend on these parameters. It's much easier to digest I guess if I had gone through one example. So, I won't go through this in detail right now.

But certainly if either the previous data sort of is borderline for the effectiveness of C or if our equivalence trial is too small or if our equivalence trial is relatively small and came out in the wrong direction, then we're not going to get very compelling evidence in these posterior distributions that either E is effective relative to P or certainly not that it's -- in terms of that it's at least 50 percent as effective.

The other thing one can do -- and I won't go through this in detail either -- is use this approach for planning sample size and basically you plan the sample size so that if E and C really were equivalent, you want a high probability of concluding that E is effective relative to P. And the calculation is made again assuming that C and E are equivalent and using the predictive distribution of the data y with regard to the prior. And you wind up with essentially a way of

calculating sample size, which again we won't go through it in detail, but it suggests, but it suggests to have definitive results. The equivalence trial needs to be of the order of magnitude of the body of evidence that demonstrated the effectiveness of C relative to P unless that evidence is very, very strong.

So, I'll just conclude by saying that I think that the therapeutic equivalence trials can't really be meaningfully interpreted without quantitative consideration one way or other, whether using this kind of methodology or doing it any other way, without somehow bringing into consideration the evidence that the control C is effective. And one really needs to consider both the strength of the evidence, the degree to which C is effective, and the degree to which effectiveness varies among trials. And the therapeutic equivalence trials really aren't practical or appropriate in situations where strong quantitative evidence for the effectiveness of C is not available.

DR. RAGHAVAN: Richard, I'm sure there are a number of people with questions. Maybe while they're thinking, I might ask you one, which I think is a

difficult one, and that is, if you think about, say, two different diseases that we treat commonly, one being testicular cancer and the other being bladder cancer, where there are now relatively accepted standards, particularly in testis cancer. As you know, in the early 1970's platinum was introduced, the platinum/vinblastine/bleomycin regimen came into play, and that seemed to be a breakthrough and has stood the test of time as a standard. It has been modified by the replacement of vinblastine by etoposide. But it's a context where patients are almost always treated with platinum, bleomycin, and something with curative intent. That regimen has never been tested in a placebo controlled trial, so that your mu/tao ratio can't be calculated unless you make intellectual simulations of what control would be expected.

More recently the MFAC regimen has come into play in bladder cancer, and that has been tested against single agent platinum, but has never been placebo controlled.

So, in the situation now, say, where in testis cancer or bladder cancer, new drugs come along

that you want to evaluate and you hope that they're, for argument's sake, less toxic, with the absence of placebo controlled data, how do you do the mathematical simulations that allow you to calculate sample size and try to get a sense of whether the new drugs are relevant or not?

DR. SIMON: I think in the testis situation, there was actually a strong body of evidence for the curative effect of the combination. Right? Now, maybe you're saying it didn't come from a randomized trial. But clearly that combination is highly effective, and I would think in that situation, we'd have to use that body of data, which clearly exists although in that situation I think it's not really from a randomized trial.

I think it would be much more difficult in the situation where we're talking about a regimen whose effectiveness is not so clear-cut as that -- and we're not talking about cures -- and then using sort of a non-randomized body of evidence to sort of estimate its effectiveness I think would be more problematic. I'm not too familiar with the bladder cancer situation.

But I think in that situation, as you move in the direction where the effectiveness is less and there weren't randomized trials demonstrating that effectiveness, then if you move too far in that direction, I think you have to ask should you really be doing an equivalence trial.

DR. RAGHAVAN: Dr. Johnson?

DR. JOHNSON: I just want to ask your thoughts about how we might deal with equivalence issues that seem to come up frequently in clinical trials, and that is the study that's designed not to show equivalence. It's designed to show therapeutic superiority but where none is identified and then equivalence is inferred as a result. Can one take these approaches that you've outlined and do posterior analyses to come up with a probability that those data in fact demonstrate equivalence?

DR. SIMON: I would say yes. But I would say the work part of it is putting together the evidence for the effectiveness of the control regimen. Right now that's sort of dealt with in a sort of offhand sort of manner, and it's not really dealt with quantitatively.

That's why in some cases it's so difficult to assess an equivalence trial. I think that step needs to be dealt with more carefully, but yes, I think one would have to do that. And the fact that the thing wasn't designed as an equivalence does not preclude you from doing that kind of analysis.

DR. JOHNSON: It's very important, I think, that we look at that type of approach because when one looks to define equivalence, typically the sample size is larger than what one would need in order to show superiority. Given the option every time, whether it's a cooperative group sponsored by the National Cancer Institute or a pharmaceutical company, always we go for the smaller size just for the practicality of trying to get the study done. I'm a bit troubled by that. It suggests that maybe doing equivalence studies in the way that maybe might be optimal isn't going to be -- there's no incentive to do that. Not only is there an incentive to be, quote, sloppy, there's not even an incentive to design it a priori. It's an incentive to do a superiority trial sloppy it seems to me.

DR. SIMON: A lot of things with randomized

clinical trials, sort of our approaches to designing them and interpreting them are pretty effective, and we can reach pretty reliable conclusions and we sort of know how to do that. We know sort of what the signposts are for questionable areas.

But I think with the equivalence trials, we're sort of still at a level where we have not established good criteria, good practice, and I think there's a lot of misleading conclusions being made.

DR. RAGHAVAN: Dr. Margolin?

DR. MARGOLIN: My question has to do with one of your last statements about selecting or calculating an adequate sample size for the equivalence trial based on the body of evidence that exists for the log of C versus P, which you've said one would base at least your assumption of its efficacy on pooled data from meta-analysis. I guess your ultimate sample size selection for an equivalence trial is going to depend on what your assumptions are, but would you use this entire body of evidence from the pooled trials or would you pick only the ones that are positive? How would you recommend selecting that number?

DR. SIMON: I certainly don't think you would pick the ones just that are positive. I think what you have to do is decide -- in the stage of planning your trial, you would have to look at -- for example, suppose we're looking at a situation where there are randomized clinical trials. You'd have to look at those situations. I think we haven't worked out all of the ways you would do this, but if it's clear, for example, that if you deal with heavily pretreated patients and not going to find improvement, if you don't, then you will find an improvement, and therefore you're going to focus this trial on the situations where you're dealing with non-previously treated or non-heavily previously treated patients. I think you could bring into bear that and that would limit the studies that would go into the meta-analysis.

But if it's really the situation that you have a wide range of outcomes for the trials that have been done and it's not at all clear what patient factors are involved with that, then I think you have to use essentially more like that entire body of randomized clinical trials.

DR. RAGHAVAN: Dr. Albain?

DR. ALBAIN: I think you alluded also to one of the problems we were having yesterday with the two Taxotere trials. You really wanted a second trial of the treatment versus best supportive care, but what you had instead, because the practicing community would not accept a best supportive care arm, two drugs that probably, quote/unquote, weren't effective when, in fact, you had never had a comparison of those drugs versus nothing. I think we're seeing a growing number of trials being designed that way where you pick a comparator that hasn't been studied against placebo or against nothing.

DR. SIMON: But I think that's difficult. I think when we can get a trial like the 17 trial, it makes things so clear compared to getting a trial like the 20 where we really don't know whether those drugs were active for survival or not.

DR. RAGHAVAN: Dr. Williams?

DR. WILLIAMS: Rich, I'm interested in your use of the prior and how conservative that is. Another way of approaching this would be to take the meta-

analysis and look at the 95 percent confidence intervals of your estimate and to say, okay, there's this much activity, either the point estimate, which has, I guess, probably been the tradition, or the lower bound of the confidence interval. That would be much more conservative.

Do you have a sense of how your analysis fits between those two extremes?

DR. SIMON: Well, first of all, I think my approach is less conservative than assuming that the lower bound of the 95 percent confidence interval of the effectiveness of C relative to P is the true. So, this approach would be less conservative than that.

So, if you have a situation actually in which you have C represents a very effective regimen and you have one or more previous trials that demonstrate that effectiveness, then to do an equivalence trial, if you limited it to the objective of demonstrating the effectiveness of E relative to P, might not give a huge sample size. However, in that kind of situation you're probably going to want more. If C is highly effective, you're probably going to want to know that your new

regimen doesn't lose too much of that effectiveness.

DR. RAGHAVAN: Dr. Nerenstone.

DR. NERENSTONE: I guess I wanted to pick up what Dr. Albain was saying. I think the no-treatment controls are statistically more pure, but I think all of us who treat patients -- you can't always go for purity because you have people sitting there in the office who are not going to agree to a no-treatment control for a variety of reasons. I submit that maybe what we need is the statisticians to help us come to a new trial design, something like treatment now versus treatment later, which I think is much more acceptable to patients than treatment versus non-treatment. Am I being too statistically naive to think that perhaps looking at this in a new way might be acceptable to both sides?

DR. SIMON: Well, I think what's important is to look at the body of evidence for the effectiveness of your active control. If it indicates that that body of evidence doesn't exist or that the degree of effectiveness is very small or that it occurs very inconsistently, then I think there's a question as to whether you're doing your patients any favor by giving

them a toxic treatment whose effectiveness is highly questionable.

DR. NERENSTONE: And that's always the argument we use, but it still doesn't sell well.

DR. SIMON: For example, then in terms of an early versus late, I think it depends on what your endpoint is and whether you think late treatment will -- in other words, if you're using an endpoint of survival, then the question is -- if late treatment is going to have that impact on survival, then you're not really going to be able to evaluate the survival effect.

DR. RAGHAVAN: Richard, life is going to become much tougher for you because Dr. Temple, among others at the table, introduced the concept of new parametric measurements. So, I can guarantee you that within your life on this committee, or maybe the person who follows you as the resident captive statistician, there will be a tension between survival and alleged measures of quality of life.

I think my bladder example is a reasonable surrogate for that discussion. You have an established, toxic, multi-drug treatment that has some efficacy but,

as you said, is usually not curative. There are a ton of new drugs that are out there that presumably will come to this committee at some point that, whether they're effective or not, are less toxic. So, there's going to be a balance between proven efficacy predicated on survival but an imperfect standard and new drugs that are less toxic which will be acceptable to patients but which may or may not give equivalent survival.

How are you going to model that mathematically? What are going to be the parameters that you put in that allow you to attribute weighting to quality of life versus duration?

DR. SIMON: I started writing on therapeutic equivalence trials I think even earlier than Bob did, and my first introduction to it was a randomized clinical trial of mastectomy versus lumpectomy. Although it didn't involve sort of a multi-dimensional quality of life assessment, it was essentially the kind of trial you're talking about.

I think we do need to have some assessment, in the imperfect world that we live in, of what the effect on survival is of what we're holding up as the

standard. If we don't have that from randomized trials, then we don't, but we have to try to put together what we believe it is and how uncertain our uncertainties are and then use that as some kind of a yardstick for evaluating the therapeutic equivalence trials that are done.

DR. RAGHAVAN: Dr. Temple.

DR. TEMPLE: Rich, we've generally said that if the control treatment more than occasionally fails to beat placebo or whatever its control treatment was, it's not a suitable setting for an equivalence trial.

If you took the data we saw yesterday and combined it, you might well conclude that there's some evidence of an effect on survival. Yet, it's also true that of two studies, one, the larger, didn't show that effect. So, presumably had there been an equivalence trial in that very setting, you would have been misled to conclude that the equivalent drug was effective because the control group wasn't effective in that setting.

How do the quantitative aspects of this work? If you reach the conclusion that you've looked at all

the placebo controlled or whatever controlled data and on the whole it shows an effect, any given trial, it seems to me, has a considerable chance of encountering one of those settings where it wasn't showable.

So, is the answer to this to do more than one or more than two? How do you get reasonable assurance that your conclusion is true?

DR. SIMON: Well, quantitatively this approach would support your conclusion that that's not a good situation, that that type of a control is not an appropriate control for an equivalence trial because, remember, I said you would do a random effects meta-analysis. That means that your -- I think I called it -- tao squared, the variance, represents the effectiveness of your control C relative to P and how that varies among trials.

So, in your situation, the example you're giving, where your active control is not consistently effective even in large trials, that means that tao squared is large. Then if you wind up going through these calculations, you will find that in that situation, your sample size for the active control goes

to infinity essentially for the equivalence trial because you really can't establish effectiveness of E very well because you have such an uncertainty about the effectiveness of C in your particular trial.

DR. TEMPLE: So, that makes it all statistical and quantitative.

One of the worries we have is that there are conditions associated with particular trials. I don't know if this is true in oncology, but maybe, that there are conditions associated with particular trials that make them poor assays, not because the study is too small or because the effect is a little variable, but because somehow that population just couldn't show anything. Will this take that into account? Not perfectly I think, but some.

DR. SIMON: I guess I don't like that notion of poor assays. Maybe that's a population of patients for whom you should be doing superiority trials. Those are the patients who need an effective treatment, and that's where the trials should be done, not in the others.

DR. TEMPLE: But we don't usually know how to

identify them. You deduce they're present because the trials give such different results, without usually being able to figure out why. Sometimes. Maybe we can get better at that. That would be good.

DR. RAGHAVAN: Well, it looks like we've exhausted the discussion on this. I personally am quite disappointed that Dr. Johnson didn't favor us with a Southern homily.

(Laughter.)

DR. RAGHAVAN: But maybe in the next session, we could --

DR. JOHNSON: No. Actually I learned something beyond statistics today, and that is that statisticians have humor.

(Laughter.)

DR. JOHNSON: Ayesian tatistics and BS. I love it.

(Laughter.)

DR. JOHNSON: It's going to go into a future homily.

DR. RAGHAVAN: Anyway, we've finished a little ahead of time. I think we should reconvene at 10

after 10:00 so that we can get off to a crisp start for the next session.

(Recess.)

DR. RAGHAVAN: Good morning. I'd like to call the meeting to order. This will be a discussion of NDA 21-156 of Celebrex. As we're reopening the session, I'd like to start with an introduction of the committee members, starting from the left-hand side of the room.

DR. SURAWICZ: I'm Chris Surawicz. I'm from the University of Washington. I'm a gastroenterologist, so I'm a GI, I guess, equivalent. Analog?

(Laughter.)

DR. SURAWICZ: A GI representative to this committee for this event only.

DR. JOHNSON: I'm David Johnson, medical oncologist, Vanderbilt University.

DR. SANTANA: Victor Santana, pediatric oncologist, St. Jude's Children's Research Hospital, Memphis.

DR. SLEDGE: George Sledge, medical oncologist, Indiana University.

MS. FORMAN: Sallie Forman, Patient

Representative.

DR. NERENSTONE: Stacy Nerenstone, medical oncology, Hartford, Connecticut.

DR. BRAND: Randall Brand, gastroenterologist from the University of Nebraska.

DR. BLAYNEY: Douglas Blayney, medical oncologist, Wilshire Oncology Medical Group, Pomona, California.

DR. PELUSI: Jody Pelusi, oncology nurse practitioner, Consumer Rep.

DR. RAGHAVAN: Derek Raghavan, medical oncologist, University of Southern California.

DR. KELSEN: David Kelsen, medical oncologist, Sloan-Kettering, New York.

DR. TEMPLETON-SOMERS: Karen Somers, executive secretary to the committee, FDA.

DR. MARGOLIN: Kim Margolin, medical oncology and hematology, City of Hope, Los Angeles.

DR. JACOBY: Russell Jacoby, gastroenterologist, University of Wisconsin.

DR. SIMON: Richard Simon, biostatistics, National Cancer Institute.

DR. ALBAIN: Kathy Albain, medical oncology, Loyola University, Chicago.

DR. LEWIS: James Lewis, a gastroenterologist at Georgetown University, and I'm a consultant to the reviewing division.

DR. AVIGAN: Mark Avigan, gastroenterologist at the FDA in the Division of Gastroenterology and Coagulation Drug Products.

DR. CHIAO: Judy Chiao, medical reviewer, FDA.

DR. BEITZ: Julie Beitz, medical team leader, FDA.

DR. JUSTICE: Bob Justice, Deputy Director, Division of Oncology Drug Products, FDA.

DR. PAZDUR: Richard Pazdur, Division Director, FDA.

DR. TEMPLE: Bob Temple, Director, ODE I.

DR. RAGHAVAN: Dr. Somers will now read the conflict of interest statement.

DR. TEMPLETON-SOMERS: The following announcement addresses the issue of conflict of interest with regard to this meeting and is made a part of the

record to preclude even the appearance of such at this meeting.

Based on the submitted agenda for the meeting and all financial interests reported by the participants, it has been determined that all interests in firms regulated by the Center for Drug Evaluation and Research, which have been reported by the participants, present no potential for a conflict of interest at this meeting with the following exceptions.

In accordance with 18 U.S.C. 208, full waivers have been granted to Sallie Forman, Dr. Russell Jacoby, and Dr. Derek Raghavan. A copy of these waiver statements may be obtained by submitting a written request to the agency's Freedom of Information Office, room 12-A30 of the Parklawn Building.

In addition, we would like to disclose that Dr. Scott Lippman is excluded from participating in the discussions and vote concerning Celebrex.

In the event that the discussions involve any other products or firms not already on the agenda for which an FDA participant has a financial interest, the participants are aware of the need to exclude themselves

from such involvement and their exclusion will be noted for the record.

With respect to all other participants, we ask in the interest of fairness that they address any current or previous involvement with any firm whose products they may wish to comment upon.

Thank you.

DR. RAGHAVAN: The next scheduled component is our open public hearing in which we encourage people to express personal opinions. Before each speaker presents, I'd be obliged if they could let us know whether they were given an inducement or remuneration from the sponsor to appear and would ask you all to keep to the allocated time.

We have one additional speaker who has asked if he could speak first because he has another commitment, and that's Mr. Kevin Lewis. Is Mr. Kevin Lewis here?

MR. LEWIS: And how would you like me to -- I've received no payment or remuneration for being here.

Thank you for letting me speak. Ordinarily I'm an Internet entrepreneur, but I also have and carry

a colon cancer gene. At 35 years old I know what is most likely to kill me. As the Vice Chairman of the Colon Cancer Alliance, a national patient support organization, we work to increase the information that's provided to patients so we have tools and that we can make informed decisions about the difficult choices that we have in our lives and in our family's lives.

I come from a family of colon cancer survivors -- thank God -- and have had experiences that most people my age don't have to deal with. I am a carrier of a gene called MSH-2 which causes colon cancer disease or a syndrome called HNPCC, a little bit technical, but it means that I very much have to watch out for what I do and work for ways of managing the quality of life that I have and the choices about treatments that I must take.

I was a participant in one of the Celebrex studies, so the Colon Cancer Alliance asked me to be our speaker here. At a personal level, I found that the drug was relatively innocuous. In fact, I found the worst part of the study was no longer taking the drug and having to deal with the aches and pains that I

ordinarily would have without some of the aspects of Celebrex.

Our organization has many people my age dealing with difficult choices about how they will either prevent or recover from their initial diagnosis of either colon cancer or something as advanced as myself who is actually carrying a gene. We ask that you provide us with all of the tools that we possibly can have to both prevent colon cancer from completely damaging our lives and giving us the tools to make informed decisions about the treatment options that we have and the various quality of life issues we face.

Thank you.

DR. RAGHAVAN: The next speaker is Ms. Abby Meyers.

MS. MEYERS: Well, I'm Abby Meyers of the National Organization for Rare Disorders. We're the group that works for orphan drugs and orphan diseases. We're here today about FAP because it's a very rare disease. It's only 1 in 10,000 people, and we're trying to start a support group for this particular type of hereditary colon cancer.

I want to say two things. Number one, when you look at the data -- and I haven't seen the data, so I can't make any judgment -- be aware that the drug cannot be studied on thousands of people, and even though the company has not asked for an orphan drug designation, understand that all of these types of drugs have to be studied in very small numbers of people. So, the data is not going to be voluminous, and you must keep that consideration.

Number two, our biggest concern is if this drug works -- and that's your decision, whether it works or not -- if it's effective, it's very, very important that the labeling specifically says that it's approved for FAP. And the reason is the reimbursement problem. Health insurance is not going to pay for an off-label use. This is an arthritis drug, and a person who's taking it for cancer is going to have a very, very hard time getting reimbursement.

So, those two things, if you can keep them in mind, when you make your decision, are the most critical because it doesn't do any good to have a drug out there that works on a disease if your insurance won't pay for

it.

Thank you.

DR. RAGHAVAN: Thank you, Ms. Meyers. Did you receive any support from the sponsor for your appearance today?

MS. MEYERS: No, and I don't own any stocks either.

(Laughter.)

DR. RAGHAVAN: Thank you.

The next speaker is Jean Marie Baxter. She's not here unfortunately.

How about Beth Schreiber?

MS. SCHREIBER: Good morning. My name is Beth Schreiber. I'm a wife and mother of two children and am the Executive Director of the Hereditary Colon Cancer Association. I am here as a part of NORD. NORD has helped us pull together a few families with FAP to make this association so that we can be represented since we are a rare disease.

I have had FAP since I was 4 years old. My father died of colon cancer at the age of 27. My grandfather died at the age of 22. My family has a very

aggressive form of this disease. My son was born with a cancerous liver tumor, which has a higher in people with FAP, called the hepatoblastoma. He had a precancerous colon at the age of 7 and is the only documented case that I can find of a 7-year-old who has hyperplastic and adenomatous polyps in his stomach. At the age of 10, he has had a total of eight operations. Two desmoid tumors have been removed.

This disease occurs 1 in every 10,000 people in the population and is considered a rare disease. This chronic disease is one that affects the whole body by accelerated cell growth in our body and lack of cell death in the mucosal binding in our bodies, mostly the colon and stomach. Left untreated, these cells produce cancer and tumors and can attack any areas of our body.

Prevention and follow-up are the most effective way to treat this disease, and with modern technology our life expectancy has improved remarkably. The new millennium will bring with it scientific breakthroughs. Among them, we hope to find one for our disease. I don't know the effectiveness of this drug for our disease, but we need to come up with effective

drugs to treat this disease. These drugs need to show acceptable levels of toxicity for us because we need to live on them the rest of our lives to minimize the size, number, and risks of cancer caused by the adenomatous polyps in our bodies.

We also need to keep the children in mind, as this disease is usually present in children before the age of 10. One-third of the FAP population are children and teenagers. This means that we need to study medications for the long-term effects, not just in a 6-month trial, and their effectiveness in the FAP population. Because it is a rare disease, we cannot test as many people in new clinical trials as with other diseases, but I feel that safety of the dosages studied in FAP trials, the overall effectiveness, and toxicities have to be a known factor before we can distribute these medications to the FAP population.

Things that I do in terms of my disease is I'm doing this association for my son, and it is a life-threatening reality that we have to live with every single day when we pass this chromosome on to our children. I never knew how life-threatening and deadly

this disease could be, seeing as though my father died before I was old enough to remember him, but I'm doing this for the children.

Thank you.

DR. RAGHAVAN: The next speaker is Pat Weidner from Youngstown, Ohio.

MS. WEIDNER: Good morning. My name is Patricia Weidner, and I have lived with the fear of FAP for most of my life. At the age of 7, I was introduce to this disease by my father having surgery to remove a cancerous growth. He always felt responsible for making a life-threatening disease a part of our legacy. However, myself, I was glad at that time the treatment was a colostomy because I would have grown up fatherless without it.

Many years later when I became a mother of two children, I better understood his feelings of impending doom because I had passed the disease on to my children. All three of us have had most of our colons removed as a preventative measure to extend our lives and hold the cancer at bay.

The only known fight against this illness is

to have screening scopes and removal of the polyps that are found. This still leaves a large margin for the tumors to grow and my family's and other families' numbers to decrease.

Three years ago I was asked to become a member of a drug study group to investigate the effect on FAP. I was honored to be a part of the study because it offered hope to me and others like me. For the past 3 years, I have been taking the medication without any side effects and with a notable decrease in the number of polyps. I have gone from 30 polyps to 5 polyps during the course of the study.

This drug can only be tested on FAP patients and should extend over a longer period of time because this is a chronic condition extending over a lifetime.

I ask you to realize the importance of a drug becoming a part of the treatment of FAP. It may not cure the patient of the disease, but clearly would make a very strong improvement in their lives and the generations to follow. Some day I'd like to be able to tell my grandchildren that you have to take a medication for the rest of your life. You don't have to face

surgeries.

Thank you.

DR. RAGHAVAN: Ms. Weidner, did you get any support?

MS. WEIDNER: No.

DR. RAGHAVAN: Thank you.

We also have two short letters that Dr. Somers will read.

DR. TEMPLETON-SOMERS: The first letter is from Dr. Leon Wang, Ph.D., of Suffern, New York. I'm reading these today because they are very brief.

"Many FAP patients have more polyps in their colons than in their rectums. It has been established that colon polyps have fewer COX-2 receptors than their counterparts in the rectum. Thus, my educated guess is that colon polyps will respond less to COX-2 inhibitors than rectal polyps. G.D. Searle's Celebrex clinical trial did not include the efficacy of this drug on colon polyps, which represent a major and possibly less responsive population of polyps in the FAP patients.

"I am deeply concerned about the risk of approving a drug based on its efficacy on a minor and

easier to treat population of polyps."

And the second letter is from Ronald Fuller of Dallas, Texas.

"I am representing myself and would like to make the following comment at the December 14th ODAC meeting on Monsanto/Searle's FAP application.

"In treating the precancerous disease of FAP, one must be aware of the warning signs in case the treatment has failed and FAP has progressed to colon cancer. One of the warnings signs for FAP is occasional crampy abdominal pain. In general, pain is the most common symptom of cancer. Although Celebrex was not able to obtain the acute pain labeling due to placebo response on the short-term pain studies, Celebrex was shown in three dental pain studies to be significantly effective in managing acute pain using 200 and 400 milligrams per day doses. In the case of the FAP study, the results indicate that 800 milligrams per day needs to be used. This means that the analgesic efficiency is equal to or much greater than the OA and RA dosing levels. Therefore, the panel may want to consider the safety consequences of Celebrex's analgesic effect in

delaying the early detection of cancer progression."

Thank you, and both these letters are available for viewing in the notebook at the registration desk. Thank you.

DR. RAGHAVAN: Thank you, Dr. Somers.

Now we're going to hear from Dr. Richard Spivey, who's the Vice President of Worldwide Regulatory Affairs with Searle, and his team. Dr. Spivey.

DR. SPIVEY: Thank you, Mr. Chairman. Advisory committee members, representatives of FDA, and members of the audience, as was indicated, my name is Richard Spivey and I am Vice President of Worldwide Regulatory Affairs at Searle.

Today we are here to discuss a supplemental new drug application for Celebrex. This application was submitted under subpart H of 21 C.F.R. 314, meaning that the application is requesting accelerated approval based upon an effect on a surrogate endpoint that is reasonably likely, based upon epidemiologic, therapeutic, pathophysiologic, or other evidence, to predict clinical benefit.

One requirement of subpart H is that a

follow-up trial be conducted to verify and quantify clinical benefit. Later today we will describe the proposed study to meet this requirement.

I would also mention at this time that the application was granted priority review.

The specific indication being sought is as follows. Celebrex is indicated for the reduction and regression of adenomatous colorectal polyps in familial adenomatous polyposis patients.

The issues discussed here today are important. There is an unmet medical need for pharmacologic intervention as an adjunct to the treatment modalities used in FAP patients.

This trial is the first of its kind to be presented to ODAC and represents a collaboration of the National Cancer Institute and Searle. Representatives from both organizations will be participating today, as well as experts in the fields of gastroenterology and FAP treatment.

Our agenda is projected here on the screen. We will spend about 20 minutes outlining the very strong rationale and evidence supporting pharmacologic

intervention in FAP, including the data from animal studies. We sill spend the remainder of the allotted time describing the results of the clinical trial, including the observed safety profile. Finally, we will review an outline of the proposed follow-up study.

We intend, through a review of the data, to demonstrate the following: that celecoxib at a dose of 400 milligrams twice a day is safe and effective for the reduction and regression of adenomatous polyps in patients with FAP in conjunction with usual care; two, that celecoxib shows a consistent benefit throughout the GI tract; and three, that celecoxib is well tolerated in FAP patients.

We would like to request that questions during the presentation be limited to those of clarifications. However, at the end of our formal remarks, we'll answer any questions that the committee might have.

It's my pleasure now to introduce Dr. Philip Needleman, who is co-President of Searle and Chief Scientist at Monsanto. Dr. Needleman.

DR. NEEDLEMAN: Good morning.

In 1990 we discovered the existence of a second enzyme involved in the synthesis of prostaglandin from arachidonic acid and made the proposal that there were two pathways for its production. We named the housekeeping pathway COX-1 as a constitutive pathway which was involved in the physiological protection of the mucosa of the stomach and in platelet function and is always turned on and active. We found the second form not normally expressed, in fact, suppressed by steroids, is inducible in models of inflammation and tissue injury and we subsequently found also induced in chemical and genetic oncogenesis.

Now, at the time in the 1990s, nonsteroid anti-inflammatories were used for the treatment of inflammation and pain of osteoarthritis and rheumatoid arthritis. The difficulty was these were non-selective and equally as effective in inhibiting COX-1 as COX-2. So, inherently you had the limitation of mechanism based side effects because of the suppression of the GI response.

So, we set out with the notion that we would do mechanism-based targeting, hypothesizing that a drug,

in this case Celebrex, which could be preferential in its inhibition of COX-1 could give advantageous therapeutic responses without inducing the side effects.

Indeed, a year ago, we presented data which showed its efficacy in arthritis versus its safety profile. You heard the desire to have a safe agent.

On the left is one of many parameters in our phase III trial, looking here at the ordinate at swollen doses versus a dose-response curve from a negative placebo control versus a fully active traditional NSAID that was COX-1 and COX-2 non-specific. What you see is the determination of a full and plateaued maximum response with Celebrex, and note that you'll be studying a lot the doses, 100 and 400 milligrams at the plateau, and while we achieved the full efficacy, if we looked at markers of endoscopic ulcers, you see that the response through the 400 milligram b.i.d. is similar to placebo, and they were all statistically different than the profound ulceration that's produced endoscopically with the traditional naproxen.

So, we put forth the hypothesis, now validated in biochemical experiments, in animal models,

in vitro, and then in humans. We initiated the arthritis trials in 1995.

Also by 1995, based on our own science and a lot of work in academia, we began the collaboration with the National Cancer Institute and initiated a trial in experimental animals, which we can make genotypic of FAP. Those trials initiated in 1995, and then we initiated the clinical trial that we'll be talking about today.

In December of last year, this was approved both for osteoarthritis and rheumatoid arthritis, and today we're here to talk about that Celebrex will be used safely and effectively to reduce and regress the polyps in the FAP patients.

I would like to point out that we view FAP as proof of concept of the COX-2 relationship of the gene and the progression of the disease. And you'll see we're committed to the long-term follow-up. In addition with the NCI and a number of academic collaborators, we've identified other cancer events driven by COX-2, and we have initiated or are initiating trials in the sporadic adenomatous polyposis, actinic keratosis,

Barrett's esophagus, and the superficial bladder cancer. So, we are committed to really understanding and attacking these diseases with a unique, safe agent.

Now, in today's presentations, we capture the wealth of understanding of the genetics and the relationship of the genes and the genetics and the progression of the disease with the over-expression of COX-2 in all phases of colon cancer. There's already considerable epidemiology capturing over a million patient-years of experience with NSAIDs, suggesting its risk factor reduction, but clearly unmasking the safety concerns that we heard about.

Indeed, by working with St. Mark's and M.D. Anderson, we've really been able to access extensive registries to attack the unmet pharmacological need of a unique agent. That's built on our genetics and chemistry and now our safety base in which we've accumulated greater than a million years of patient experience time on the safety of Celebrex, a drug that's now used in over 6 million people.

So, while the population is very limited, in the past the FAP trials have been limited to 20

patients, 83 patients at least gives you a chance for placebo control and dose ranging and is so far the largest trial in FAP.

In embracing the accelerated review, that was with our commitment to do the long-term outcome trial, and we're working with the agency and with you to see if the design is sufficient.

We see this as adjunctive therapy with the standard of care, which is the surgical treatment of the patient.

I'd now like to introduce Gary Kelloff, who's Chief of the Chemopreventive Agent Unit of the NCI.

DR. KELLOFF: Thank you, Dr. Needleman.

For over a decade, the NCI has been actively supporting and evaluating the field of the development of nonsteroidal anti-inflammatory drugs for the prevention of cancer. The weight of the efficacy data from human epidemiology, intervention research, the preclinical animal data, and the mechanistic data has become progressively compelling over the last decade. And this has led NCI to focus on the safety side of this equation for this intervention. Searle's substantial

safety data developed for its drug celecoxib for an arthritis indication led NCI to an agreement with Dr. Needleman and Searle in May of 1996 to evaluate this drug in several chemoprevention settings.

As Dr. Needleman and Dr. Spivey described, you will hear the data today from a recently completed collaborative study of this drug in patients with FAP. First, however, I wanted to briefly summarize our strategies for colorectal cancer prevention and the historical NSAID data that led to this study.

The adenoma carcinoma sequence first described by Muto and Morson in 1975, after two decades of intense study on over 3,000 patients, many of whom were FAP patients, and the more recently described genetic progression model of Bert Vogelstein has provided us with the conceptual framework to develop strategies for colorectal cancer prevention. The germline lesion found in FAP patients is present as an acquired genetic lesion in 85 percent of sporadic colorectal cancers.

Intervention strategies derived from the adenoma carcinoma model led to the National Polyp Study

from which we know that surgical intervention by excision of adenomatous polyps has led to a reduction of 90 percent of expected colorectal cancer in two reference populations not having polypectomy and led to a 75 percent reduction of expected colorectal cancer incidence in our gold standard reference data base from SEER which is NCI's reference base that allows us to keep score as to how we're doing in cancer incidence and mortality. This database involves about 10 percent of the people in the U.S. and, therefore, over 25 million subjects.

From these data and the compelling NSAID epidemiology and intervention data, which I will summarize, we know that adenomatous polyps are risk markers and disease markers of colorectal carcinogenesis and, as part of the neoplastic process itself, our near obligate precancer lesions that will likely provide acceptable surrogates for colorectal cancer incidence and mortality.

We know that intervention with NSAIDs in prevention of colorectal cancer has provided a consistent effect in animal studies, human observation,

and clinical intervention studies. We've seen reduced incidence, multiplicity and size of tumors in animal models, and this is over 50 quality reference publications involving intervention at all different times of the carcinogenic process. We know from 24 completed human studies, involving over a million subjects and several million subject-years, that the NSAIDs have consistently reduced the incidence of adenomas, colon cancer, and colon cancer deaths. Time only permits me to show this compelling data in summary form on the next three slides.

First, NSAID use and colorectal cancer incidence. You see the seven studies' compelling agreement of point estimate relative risks from all seven studies in reduction of adenoma incidence.

NSAID use and colorectal cancer incidence. Again, 23 studies, the vast majority of which are statistically significant for NSAID use and colorectal cancer incidence reduction.

The gold standard of endpoints of colon cancer mortality and NSAID use remarkably consistent, four fairly sizable studies showing reduction in

relative risks.

In summary then, the NSAID epidemiology in humans has consistently shown adenoma incidence reduction, carcinoma incidence reduction, cancer-associated mortality reduction. This activity importantly has been observed across wide cohorts of general population and at-risk subjects, and by inference, many of these subjects started these studies with prevalent adenomatous polyps, indicating that NSAIDs offer promise in late intervention as well. This activity has been observed in men and women, middle and older ages, left and right-sided lesions, wide-ranging geographies. And there's extraordinarily consistent results.

Really the efficacy of NSAIDs for prevention of colorectal cancer is not a question in our mind. It's really a question of safety tolerance. The limitations of NSAIDs use for colorectal cancer prevention really go to the safety issues. These safety issues and limitations have been well described and are well known to all of us from many publications.

So, in summary and conclusion then, NCI's

interest in celecoxib derives from the NSAIDs' impact on colorectal carcinogenesis from mechanistic, preclinical, clinical research.

The safety of the non-selective NSAIDs somewhat limits their use.

The preclinical efficacy of celecoxib is at least comparable to the NSAIDs, as you'll hear from Dr. Masferrer in a moment who will summarize this data.

I would say that NCI has data that celecoxib has promise for other target organs as well.

Fourth, there's a significantly better safety profile for celecoxib than the NSAIDs.

And finally, we see a substantial potential of celecoxib for patient benefit that we believe, with completion of the several ongoing trials, will extend to many people in addition to patients with FAP.

Thank you. With that, Dr. Masferrer will summarize the preclinical pharmacology and efficacy data.

DR. MASFERRER: Thank you, Dr. Kelloff.

The evidence for the use of a COX-2 inhibitor in cancer is based on the overwhelming data that you've

seen on the epidemiology with NSAIDs, the over-expression of the COX-2 enzyme in human tumors, and the efficacy by this compound shown in animal pharmacology.

Analysis of the literature and our own data shown in this slide, using a specific inhibitor and a specific antibody against the COX-2 shows that COX-2 is expressed in all stages of human colon carcinogenesis. This can be observed here with the red color. This is a sample of polyps from FAP patients. COX-2 is also expressed on colon cancer here in the red, and in colon cancer cells that we can observe in metastasis like in this example in the liver.

In contrast, if one looks at normal colonic cells, the top panel on the left, COX-2 is not normally expressed with a few exceptions of a few cells that express low intensity of the COX-2.

The expression of this enzyme in the polyps, as well as in the cancer, makes a clear target as a cancer preventive agent, and we tested that pharmacology in several animal models.

The first one was azoxymethane-induced colon carcinogenesis. Chemically is given to the animal

azoxymethane, and a year later we can observe the presence of adenomas in about 10 percent of the animals as well as non-invasive and invasive adenocarcinoma. About 90 percent of the animals will have a form of tumor.

Celecoxib was given to these animals at the dose of 1,500 milligram per kilo in the diet for 1 year, and we observed a remarkable inhibition on the numbers of adenomas, as well as in adenocarcinomas, non-invasive and invasive, with an overall reduction of about 86 percent.

Celecoxib in this study also affected tumor burden by about 75 percent inhibition.

We were very fortunate that animal models for the FAP have been developed. These animal models have a mutated Apc gene, and these animals developed large numbers of polyps, very similar to the FAP conditions. The rule of COX-2 in this formation of polyps was assessed by eliminating the COX-2 genes from the Apc animals. The result of that experiment was an 86 percent inhibition in the formation of polyps seen here in the animals who did not have the COX-2 gene. So,

this is the first direct genetic evidence of a key role for this enzyme in the formation of polyps.

We did the same genetic experiment but now pharmacologically inhibiting the COX-2 enzyme, and that was done in the other animal model called the MIN mouse that also carries an Apc mutation and they also developed several tumors.

Celecoxib was treated here in two ways. We went early after the weaning of the animal at day 30, a preventing modality, and also we went on day 55 when all the animals already have tumors. So, that will be a more therapeutic ordinary regression modality. I've just seen here celecoxib dose-dependent inhibit tumor multiplicity when go early in a preventive modality as well as their regression of the tumors when we go late.

Celecoxib in this experiment was specific to the COX-2 enzyme. It did not inhibit the COX-1. We measured the thromboxane level coming from the platelets of these animals, and there was no inhibition.

Also, in all the experiments that we've shown here, celecoxib was very well tolerated and there was no sign of any toxicity on these animals.

When we compare, as we typically do in these experiments with an NSAID like piroxicam shown here, we observed similar efficacy with the drug. The difference is that these NSAIDs will have, together with the efficacy, the inhibition of the COX-1 measured again as the thromboxane in the platelet, and we can observe sometimes signs of gastrointestinal toxicities that are not observed with celecoxib.

Now, if I can have the next slide. So, in summary, the COX-2 enzyme is over-expressed in all different stages of colon oncogenesis. The very clear experiment showing the genetic deletion of COX-2 inhibits polyp development, key on the role of this enzyme in tumor formation.

Then the pharmacological effect that we see with celecoxib and COX-2 specific inhibitor reducing colon adenomas and cancer development either if we go early or in the late administration in the two models that I'm showing you here.

Finally, celecoxib is effective and well tolerated in the animal models of cancer prevention.

It's a real pleasure to introduce to you Dr.

Bernard Levin, the Vice President for Cancer Prevention and a professor of medicine from the University of Texas, M.D. Anderson Cancer Center.

Thank you.

DR. LEVIN: Dr. Raghavan, members of ODAC, members of the FDA, ladies and gentlemen.

As you've heard so movingly, familial polyposis is an uncommon but devastating disease. It is inherited as an autosomal-dominant due to germline mutations in the Apc gene at this locus, and the clinical severity depends on the phenotype. It affects approximately 1 in 10,000 individuals, and an understanding of FAP provides fundamental understanding into the biology of adenomas and the development of colorectal cancer.

Depicted here is the gross morphology from a surgical resection of the numerous adenomas that carpet the colon and rectum. Juxtaposed is the endoscopic view with a millimeter rule set against several of these adenomas in an assessment of size.

The adenomas appear by about 10 years of age in about 15 percent of people, and by 30 years of age,

90 percent of people have numerous adenomatous polyps.

Characteristics of these adenomas are that they are indistinguishable from those that occur in sporadic populations of adenomas, and the distribution of polyps in established familial polyposis includes complete distribution throughout the colon and rectum in 100 percent of people and virtually 93 percent cover the duodenum when this disease is established.

The natural history includes rectal bleeding, changes in bowel habit, and abdominal pain. At the time of symptomatic diagnosis at an average age of 36 years, 70 percent of these individuals have a colorectal malignancy, and over the lifetime of this illness, there's 100 percent cancer risk, typically in the fourth and fifth decades, with an average age of death of 42 years. The extracolonic manifestations include duodenal adenomas leading to dysplasia and cancer, as well as desmoid and other tumors.

Screening includes, in usual clinical management, flexible sigmoidoscopy for all first degree relatives, with the initial screening beginning at age 10 or 12 and then annual videorectoscopy to minimize

invasiveness until age 18 to 20. Colonoscopy to examine the entire colon with dye spraying to bring out the smaller lesions at age 18 to 20 and then every 5 years thereafter if surgery has not taken place. The initial upper endoscopy usually around age 20 to 25, and then surveillance for those with mild duodenal polyposis being monitored every 2 to 3 years and with significant involvement every 6 to 12 months. Of course, genetic counseling with appropriate genetic testing will be an important, necessary adjunct to screening in these individuals.

With respect to colon cancer in familial polyposis, a few more details. Untreated, the mean age of diagnosis is 39 years, but 87 percent of these individuals untreated develop cancer by age 45 and over 90 percent by age 50. Unfortunately, the life expectancy after the diagnosis of cancer is only 2.6 years.

The polyp number and age correlate with cancer risk such that for each 10-year age group, there is a twofold increase in risk, and today with screening and surveillance and greater knowledge of the disease,

still 25 percent of patients have cancer at the time of diagnosis of this syndrome.

The primary management of FAP is surgical prophylaxis. For the colon and rectum, this includes ileorectal anastomosis after colectomy or, depending on the clinical situation and the extent of rectal involvement, proctocolectomy with ileal-pouch-anal anastomosis and subsequent surveillance of the rectum or the pouch depending on the surgical procedure. For the duodenum, there is no standard approach, and additional secondary surgery is often needed as clinically indicated.

However, while the primary management is surgical, this is not optimal and despite standard screening, prophylactic surgery and endoscopic surveillance, the relative risk of death for these individuals is over three-fold from data from St. Mark's Hospital. The causes of mortality include duodenal cancer, desmoids, rectal cancer, and perioperative complications, and other extracolonic manifestations. The impact of surgery on the quality of life includes such disturbing issues such as nighttime fecal

incontinence and sexual dysfunction, failure of ejaculation, failure of erection. There is no approved pharmacologic agent available for these individuals.

Nonsteroidals, as you've heard, have been studied. The database includes about 100 patients in uncontrolled studies. There have been three controlled studies. The largest number is 24 in these studies.

The findings include a reduction and regression of polyps, but no consistent effect on duodenal neoplasia, and unfortunately, these studies are not comparable to each other due to methodological differences.

The overwhelming concern for long-term therapy is that of the NSAID side effects, which you've already heard about.

So, the possible clinical benefits of this drug in FAP management could include the reduction and regression of polyps, thereby facilitating endoscopic surveillance, a delay or prevention of secondary FAP-related GI surgical procedures, a reduction or delay in duodenal neoplasia, a delay or prevention of emergence of the phenotype in adolescents, and the long-term

overall favorable safety profile for its administration over long periods of time.

I'd now like to introduce Dr. Ernest Hawk from the National Cancer Institute where he is Chief of the GI Cancer Research Group.

DR. HAWK: Good morning. Thank you, Bernard.

This morning I have the pleasure of presenting the results of our placebo-controlled, randomized trial of celecoxib in FAP subjects. What I'll cover over the next 20 minutes is, first of all, a discussion in a single slide of the rationale that went into the trial. You've heard a great deal about that already.

Next I'll cover the trial design, the methods that were used, particularly for outcome assessment, since that's a critical feature of your evaluation.

Next I'll cover the results in terms of the demographics, the colorectal results and the duodenal results, and finally a single-slide conclusion.

You've seen this slide before. I use it merely to underline the fact that a great deal of scientific background underlies this trial, both in

terms of the disease, its mechanisms, and the steps within the disease process that could be used as meaningful surrogates, the efficacy of nonsteroidals more generally in more than a million subjects and several million subject-years, the understanding of FAP patients and the strength that this trial was offered by the participation of M.D. Anderson with its long history of innovations in prevention research, and St. Mark's which is arguably the world's premier institution for the registration, management, and care of FAP patients, and finally the innovations that Searle brought to the table in terms of Celebrex, the safety database that they have accumulated, as well as the preclinical data that was brought to bear on the issue.

Now, covering a bit about the trial design, we conducted a double-blind, placebo-controlled, parallel group study of celecoxib in persons with familial adenomatous polyposis. There were three treatment arms to the study: first of all, placebo; the other two, celecoxib at 100 or 400 milligrams po b.i.d. For the remainder of the presentation, I'll refer to these as celecoxib 100 milligram dose group and 400 for

convenience sake.

We planned to accrue 81 subjects randomized in a 1 to 2 to 2 manner, placebo to the two active arms. In reality we accrued 83 subjects because there were 2 replacement patients replacing patients that dropped out in the 100 milligram dose group for other than toxicity reasons.

I'll also point out, even at this early point, that the point was we had 75 patients that were enrolled with regard to the colorectal endpoint. Because these patients suffer from duodenal disease as well, we thought it was important to the n because they undergo serial surveillance of that target organ as well. We felt it was an opportunity to study the activity of this compound in both the upper and lower GI tract. So, we also evaluated patients with duodenal disease in the upper tract as well. You'll see some of that difference coming out later.

The important point here is we accrued 75 patients to the colorectal endpoint per se and we allowed 6 patients on trial with duodenal disease only, randomized them in a 1 to 1 to 1 manner to try to get a

bit more data in duodenal effects.

We administered the drug over a 6-month period. As I've already alluded to, the two sites were M.D. Anderson and St. Mark's.

Here I represent the rationale that went into dose selection. At the time this study was conceived, there was already preclinical data on the efficacy of this compound in animal models that you've heard about. However, the NDA was obviously not yet approved for arthritis, and we built the dosing on the basis of phase II data arising out of osteoarthritis and rheumatoid arthritis, anti-inflammatory efficacy, as well as safety, and selected what we thought would be a minimal but effective dose, as well as a higher dose to afford, perhaps, more activity.

The study duration of 6 months was premised trying to balance a couple of issues. First of all, we reviewed the world's literature for intervention studies in FAP subjects, particularly related to nonsteroidals. There were three. Those studies were in durations of 4 months, 6 months, or 9 months. Activity was seen as early as 4 months, but no major difference after 9

months. Therefore, we concluded that a 6-month trial was reasonable.

It's also important to point out that this patient population has tremendous clinical needs, as you've heard earlier, and we were cognizant of the fact while we wanted to get convincing data, we didn't want to run a trial unnecessarily long because these patients typically are involved in other important trials as well.

The eligibility criteria are outlined here.

First of all, in terms of inclusion, we obviously required a diagnosis of FAP.

We required that subjects have a retained colorectal segment. You'll see that was either rectum or complete colorectum.

We required that they have 5 polyps greater than or equal to 2 millimeters in size, that is, evaluable, in a focal colorectal segment. This could have been anywhere in the GI tract, but it was required to be in an endoscopically focal area that we could get very tight data on in terms of polyp counts.

We also required obviously they be abstinent

from frequent nonsteroidal use over the previous 6 months to try to avoid confounding effects.

We excluded patients on the basis of gastric ulcers or erosions based upon, obviously, a sacrifice of the target organ of primary interest within 8 months of randomization.

We also excluded patients on the basis of previous colectomy within the prior 12 months because of some old literature that suggested that there was, indeed, adenoma regression associated with surgical removal of part of the colon. So, we made sure these patients were remote from that effect as well.

And obviously we excluded patients with prior metastatic cancer.

The study endpoints that were chosen for the trial were several. The investigative team felt it was important to assess the burden of this disease -- and the literature would bear that out -- in several different complementary ways.

First and perhaps most important was adenoma number, but to correlate with that was adenoma size. We also felt that perhaps a small area in doing these

counts wouldn't be representative of the potential overall benefit within the target organ. So, we felt it was important to look at the complete remaining target organ in both the case of the colorectum, as well as the duodenum.

Now, trying to prioritize those in a regulatory manner was very challenging for us. We selected, on the basis of our best estimation, as a primary efficacy measure the colorectal result, that being a percent change in the number of colorectal adenomas greater or equal to 2 millimeters at 6 months compared to baseline.

In terms of the secondary efficacy measure, we wanted to base that in the duodenum, and we thought the complementary approach of a percent change in here the area -- not the number, but the area -- of plaque-like involvement in the duodenum at 6 months compared to baseline would be a reasonable single measure, although we felt all were important.

We also assessed both safety and tolerability.

This then includes what got relegated to

supportive analyses, but again they reflect important parameters of the disease burden.

The first is really just an ordinal look at the primary outcome data. That was not an a priori hypothesis.

The following were a priori hypotheses, and I want to stress that. We wanted to look at residual polyp size to see if there was a reduction in polyp burden, which was the sum of the polyp diameters. This is somewhat different than these two measures, although they're both capturing size data. Here we were able to assess polyps that might have had a 100 percent reduction in size, complete resolution.

Then to complement these focal assessments again, we thought it was important to do endoscopic videotaping to look for changes in both the colorectum and duodenum.

Moving on to the methods then used in the outcome assessment, as I've described, we did focal assessments which were based upon endoscopic photographs of various GI tract segments, designated specifically so we could return to them again by other anatomic markings

such as the ileocecal valve or the appendiceal orifice or by tattoos placed at baseline.

To complement that again, we did a global assessment, and in order to try to capture the information in those, we had that reviewed by a panel of experts.

This is an example of how the photographs were taken and assembled. First of all, the anatomic landmark or tattoo in the center of the first photograph. Then we took photographs with that tattoo or anatomical landmark at the periphery to broaden the area under evaluation and assure that the counts we were taking here were as accurate as possible.

We then took these slides, and Dr. Marina Wallace put them together and assessed very carefully the size and number of adenomas within that focal area.

This is a real-life example of what this looked like. The prior mock-ups show five photographs. However, in reality, the gastroenterologist sometimes felt it was helpful and important to take additional photographs. What you see here, first of all, is the tattoo in the center, and then outlined subsequently in

the yellow circles on each of the complementary photographs. You can also see we placed either a graduated measuring stick or an open biopsy forceps, both of which have constant size, right next to polyps to try to gauge their size accurately.

This is an example. Here is the tattoo. We're looking just proximal to that anatomically in the colon. You can see this adenoma just to the right. I'll follow it in subsequent photographs here so that you can get an idea of how these photographs complement one another here, here, there, and there. Again, it may not project well, but the gastroenterologists were able to see that.

So, that was the colorectum.

Now, in terms of focal assessments in the duodenum, these were based similarly on discrete photographs. However, we felt a different approach was necessary because of the plaque-like disease in this target organ. Therefore, we took what we felt were the best representative photographs of a high density area of disease and then a low density area. We took those two photographs. We averaged them and came up with an

overall percent involvement for that subject in the duodenum. So, you can see there's a bit more variation here than there was in the very tight assessments in the colorectum. Also, I'll point out that there's very little literature in this area in order to guide these approaches also.

So, here's a real-life example again of the duodenum. You can see, even to a non-endoscopist such as myself, obvious plaque-like involvement here, and there's a very small plaque right there. So, a high density example and a low density.

In terms of video assessments of global involvement, we took videotapes and had them reviewed in a blinded fashion by five experienced endoscopists. I'll point out that these were not merely gastroenterologists. There were two gastroenterologists from the University of Texas M.D. Anderson, two colorectal surgeons from St. Mark's Hospital, and one surgeon from Roswell Park, which was a nonparticipating center but has clear expertise, registry, et cetera in polyposis patients.

The videotapes were played on a single

monitor. Each of the endoscopists viewed that and performed an independent assessment of what they were seeing in terms of disease response on the screen to individual scores: a negative 1 for a clear worsening, 0 for no change, or plus 1 for clear improvement. I'll stress the fact that this had to be clear improvement or clear worsening. We weren't looking for marginal changes. These scores were then compiled into a mean physician assessment score.

I'll point out also this was a blinded review, and in the videotapes these endoscopists were blinded not only to treatment assignment but also to whether they were looking at a before or after videotape and also to patient identifiers.

The results then. On this slide is depicted the baseline demographics of the cohort that was accrued. You can see the randomization worked well in distributing variables in terms of race or ethnic origin and gender, a well-balanced distribution here, 1 to 2 to 2 remember. However, it failed to distribute age in a balanced manner. The placebo group had a mean of 41 years old, 39 years old for the 100 milligram dose, and

33 for the 400. We assessed this was statistically significant. We assessed the importance of this -- tried to, at least -- in three ways.

First of all, were different patients entering each arm? Well, the age ranges are clearly the same in each, so the same sorts of patients entered both arms in our estimation, although perhaps in different quantities.

The next thing we did is perform adjusted analyses of the significant outcomes that we noticed in the trial. None of those age adjustments decreased the magnitude of the effect we saw or the statistical significance.

Third, we supposed that age, if it was going to affect the outcome, should have an impact on the disease either in adenoma number or in size. So, we did comparisons of those at baseline as well. While the celecoxib 400 group had slightly fewer adenomas than the placebo group, and the 100 milligram dose group fewer yet, these differences are not statistically significant and there's absolutely no difference in mean polyp size.

Now, by enrollment by center, approximately

50 percent of the cohort was accrued at the University of Texas M.D. Anderson and 50 percent at St. Mark's with an excellent balance across treatment arms.

Now in terms of baseline surgical status, this was important to one of the speakers early on -- or the letter, I guess. Approximately 30 percent of the cohort had intact colons. This was well distributed across arms. These were individuals who had not yet come to their prophylactic colectomy.

About 60 percent of the cohort had a prior colectomy. These are patients that would have retained rectal segments, ileorectal anastomoses in some cases. 6 patients had a portion of the sigmoid remaining as well. Again, well balanced across arms.

Then finally, these are 5 patients with total proctocolectomies. These would have been patients evaluable for the duodenum that were accrued on the basis of their duodenal disease only. And 1 of these patients similarly didn't have colorectal burden in terms of polyps, and therefore was a duodenal only patient.

Now, how do we arrive at the analytic cohort?

Intent-to-treat concepts were followed throughout. We accrued a total of 83 subjects here for the focal assessment, first of all. We, as I mentioned early on, accrued 77 subjects to the study for the purpose of a colorectal assessment. We accrued 52 in total for evaluation of the duodenum, that being composed of 46 subjects from the colorectal group that also had duodenal plaque-like disease at baseline, therefore could be evaluated for regression, as well as the 6 additional patients with duodenal only disease.

In terms of the global assessment, we chose to just take as many patients as we could have for both target organs that had both before and after videotapes. So, we accrued 83 subjects. 73 of them had both before and after colorectal videotapes. 78 had both before and after duodenal videotapes. That wasn't because we selectively videotaped individuals, but rather due to technical difficulties, we didn't have both studies in all subjects.

Now, moving on to the results in the colorectum. This shows the data from the primary efficacy analysis of the colorectum, the percent change

in the number of colorectal polyps. You can see a 4.5 percent reduction in the placebo arm over the 6-month interval, a larger 11.9 percent reduction in the 100 milligram dose group, and a larger still 28 percent reduction in the 400 milligram dose group, which was statistically significant compared to placebo at a p value of .003.

This is an example of a what a patient that responded to the therapy looked like before and after. On the left, you'll see there's the tattoo and the polyp burden by one of the usual cloverleaf photographs. On the right, there's the tattoo again, and I think even again to me, a non-endoscopist, I can see a vast difference in the amount of disease here. By strict count, it was 41 polyps here and 21 there, leading to a 48.8 percent reduction, which was not our best response.

On this slide, we depict the individual patient responses, as well as the median response. You'll see that by median in the placebo arm, there was no effect over the 6-month interval, as opposed by the mean where there was a 4.5 percent reduction. That 4.5 percent reduction is largely driven by a single patient

who had clearly an unexplained reduction over the 6-month interval of approximately 50 percent.

Also clearly obvious on this slide is a great heterogeneity in responses or disease change in the 100 milligram dose group. However, the whole distribution of patients has shifted toward a reduction in the 400 milligram dose group. That reduction by median is 32 percent, again statistically significant. I think it's important to point out that that's a median response of 32 percent, but clearly many patients benefitted far more, even up to 80 percent reduction in their polyp burden over a 6-month period.

Also, I'll point out, in terms of progressive disease, here 3 patients. That would have represented 6 compared to here, and here we have 2 patients that progressed on therapy.

So, that was adenoma number, the primary endpoint. Now we'll move to one of the supportive analyses, the percent change in the colorectal polyp burden.

This is an assessment of size considering polyps that could have regressed completely. We have a

4.9 percent reduction in the placebo arm; 14.6, again intermediate between the higher dose and the placebo with the 100 milligram dose; and a 30 percent reduction in the 400 milligram, which was again statistically significant.

Putting all of the supportive analyses as well as the primary analysis in the colorectum by focal assessment together, we have the reduction in number of adenomas in the 400 milligram dose group that's statistically significant compared to placebo. We have, obviously, the ordinal response, and here we have 53 percent of patients having at least a 25 percent reduction over 6 months in the 400 milligram dose group compared to placebo, 6.7 percent.

We have a reduction in residual polyp size that did not achieve statistical significance at a 4.9 percent reduction, but in terms of the adenoma burden, which considers adenomas that could have regressed completely, probably a more complete picture of change in size, we have a 31 percent reduction, statistically significant.

Now, to complement that focal assessment --

let's make sure that that focal assessment was accurate in terms of the overall global burden in the colorectum -- we have data here from the videotape assessment. To orient you, this is the physician's score that I described earlier, worsening going up, increase in adenoma burden, improvement going down.

We have colorectal segments, the cecum to the ascending colon, the transverse to the sigmoid. Remember, these two are smaller than the overall cohort because not all patients had complete colorectums.

And then the rectum.

What we see is the placebo group had worsening over the 6-month interval throughout the colorectum by our expert opinion. The 100 milligram dose group had a bit more heterogeneity, but I think a trend toward response. But clearly the 400 milligram dose group has a profound reduction in the number of adenomas globally throughout the colorectum and that was statistically significant, as is shown here in the p values for each independent segment.

This shows the consistency among the video reviews conducted by the expert panel. In the placebo

arm, we had one reviewer who felt things got better, but four who thought it got worse. So, a bit of noise, you might say, around that measurement. With the intermediate dose, we had improvement, but statistically significant improvement, remember, here as well as great consistency among scorers in the 400 milligram dose group.

So, in summary, in the colorectum by focal assessment the primary endpoint, we saw a reduction in polyp number. By supportive focal assessment data, we saw several of those individuals have significant responses. And we saw the overall polyp burden measurement of size reduced. Then to complement those focal assessments, they were, indeed, representative of what was going on in the entire colorectum for there was improvement across all regions by all five scorers.

So, all of the analyses confirm a consistent, substantial, and statistically significant improvement in the colorectal polyposis in the 400 milligram dose group.

Now, in the duodenum, this is an area where, remember, these patients have no current therapeutic

options. In terms of the primary outcome in the duodenum, the percent change in the area of duodenal plaque-like polyps, we saw a trivial 1.4 percent reduction in the placebo and a 14 percent reduction in the 400 milligram dose group, with again an intermediate response here in the 100 milligram dose group.

I will point out that 2 patients in this group are not included on this slide. They had no disease at baseline and had disease at the 6-month point. So, if they would have been included in this percent change from baseline, this bar would be up here. So, clearly there's not an overwhelming activity here in this arm. And this response was not statistically significant. Nevertheless, we're hopeful that that shows a trend. I think the other data that I'll show you may substantiate that hope.

Here's an example of what a patient who did respond looks like, here again with the focal plaque-like polyp at baseline from a high density region, and then looking for disease that we're again going to try to call high density here at 6 months, and we really couldn't identify significant disease. So, this is a

100 percent responder from baseline to 6 months.

So, again, the same rationale. We've got the focal assessment, a non-statistically significant reduction in duodenal plaque-like disease, but a suggestion of benefit.

What happened globally by expert opinion? Well, there was no change in the placebo group, no change in the 100 milligram dose group, but again a statistically significant improvement in video endoscopic scoring from baseline to 6 months, which achieved statistical significance at a p value of .033.

Once again for consistency sake, let me show you how these observers agreed or failed to agree on what they were seeing. In the placebo arm, a real mix, some heterogeneity, again with a single reviewer feeling that things got better there, two clearly no change, and two improvements. So, a bit of heterogeneity there, as you might expect in a placebo arm.

The same sort of heterogeneity in the 100 milligram dose group, but clearly consistent results among all five endoscopists in the 400 milligram dose group.

So, in the duodenum, I want to point out again we saw a non-significant reduction in the 400 milligram dose group, but at least a trend. In the global assessment by the expert panel, we saw significant improvement in the 400 milligram dose group. And our conclusion then are these findings are suggestive of a beneficial effect in duodenum where no current therapy currently exists.

Our conclusion then, celecoxib 400 milligram b.i.d. results in a focal reduction and regression of colorectal polyps by very careful measurement using standardized methods and techniques, and globally there's improvement in the endoscopic appearance of both the colorectum and the duodenum by a panel of five experienced endoscopists who care for FAP patients daily.

I'd now like to turn over the podium to my colleague from Searle, Gary Gordon, who is the Director of Cancer Prevention and Treatment in Clinical Research.

DR. GORDON: Thank you, Ernie. Good morning.

What I would like to do is to briefly discuss with you some background surrounding celecoxib, the

methods used in this study, patient disposition, adverse events, and conclusion, and then turn to a discussion of the follow-up study.

As you know, celecoxib was approved 1 year ago for use in osteoarthritis and rheumatoid arthritis. At that time for the submission, there was data on 9,400 individuals who had received celecoxib. Since that time, as Dr. Needleman has mentioned, over 1 million patients have received this drug -- or 5 million, and we've accumulated 1 million years of patient experience. The incidence of adverse events has been low, both in the original filing and in the follow-up studies, and the drug has a similar short and long-term safety profile. We've noticed no dose-related increase in adverse events over the dose ranges studied.

As you've heard, celecoxib has efficacy that's comparable to the NSAIDs both in preclinical models and in clinical settings. Turning to some of the safety findings that separate or distinguish celecoxib from the NSAIDs, if you focus on those side effects that are related to COX-1, dyspepsia, abdominal pain, and nausea, you can see that celecoxib in a large database

of 4,100 individuals studied at doses of 200 milligrams twice a day or less or 615 individuals studied at 400 milligram twice a day is substantially lower than the NSAIDs, and the other events are at rates nearly comparable to placebo.

This is shown more clearly on the next slide where this is shown graphically where we look at adverse events, percent of individuals, and look at placebo, a dose of 50 milligrams twice a day up to 400 milligrams twice a day. You can see for any event, for headache, for dyspepsia, upper respiratory tract infection, diarrhea, sinusitis, abdominal pain, and nausea, there's really no evidence of any sort of dose response.

This differentiates us from the NSAIDs in the sense that there's a reduction in endoscopic ulcers, as Dr. Needleman showed earlier, and ulcer-related complications compared to NSAIDs, and that in fact it's quite similar to what's seen with placebo. There's a reduction in upper GI symptoms compared to NSAIDs, and in terms of hemostasis, there's no effect on platelet aggregation at doses up to 1,200 milligrams twice a day which is three times the dose being used in this study.

In terms of the FAP study, we collected information on adverse events in three ways. We had unsolicited reports. We had a standardized patient questionnaire that was administered once a month throughout the study, and we had clinical laboratory tests that were collected at baseline, 1 month, 3 months, and 6 months. The NCI common toxicity criteria were used to grade all adverse events.

In terms of patient disposition, we had 17 patients who enrolled in the placebo group, 34 in the 100 group, and 32 in the 400 milligram group, for a total of 83. Approximately 95 percent of the individuals completed the study. 1 patient was lost to noncompliance. 1 was lost to follow-up, and 1 discontinued the study due to a serious adverse in the 100 milligram group. And there were 2 individuals who did not complete the study in the 400 milligram group, 1 due to an adverse event, and 1 due to a serious adverse event. So, again, as shown, 94-95 percent of the individuals completed the study.

The serious adverse events observed in this study are listed here.

One occurred in an individual taking 100 milligrams of celecoxib twice a day on study day 104. This individual committed suicide and this individual had a history of previous suicide attempts and a complicated psychiatric history. And the event was judged unrelated by the investigator, and the individual did not complete the study.

There was a second individual in the 400 milligram b.i.d. group on study day 94 who had an episode of an acute allergic reaction that was characterized by urticaria and minimal respiratory distress. This individual was treated in an emergency room and discharged. He did have a prior history of urticaria. This was judged by the investigator at the time as being probably related to drug, and this individual did not complete the study.

The last serious adverse event was also in the 100 milligram b.i.d. group, and this was an individual who on study day 20 was admitted to hospital for an elective resection of a pre-existing angiofibroma. This was not felt to be related to the study drug, and the individual did complete the study.

On the next slide, we show grade 2 adverse events and greater, focusing on the gastrointestinal events. You can see, if you look at the grade 2 events across all the treatment groups, that they're roughly comparable between placebo and the two celecoxib groups. And if you focus on the GI events, again it's fairly comparable for all these events overall the patients. And the grade 3 events again were not different than placebo.

For laboratory testing, there were no differences observed between the celecoxib and the placebo groups in terms of hematology, including hemoglobin, hematocrit, white blood cell count, platelet count. In terms of clinical chemistries looking at BUNs, creatinine and liver function tests, there were no differences, and there were no differences observed in the urine analysis.

So, the safety summary is we feel there are no differences between the celecoxib and the placebo groups, that celecoxib was well-tolerated in this setting, and that this is consistent with the experience in the larger osteoarthritis and rheumatoid arthritis

databases and post-marketing experience.

The overall conclusion is that we believe celecoxib 400 milligrams twice a day in patients with FAP is safe and effective treatment for the reduction and regression of colorectal polyps, and as Dr. Hawk has pointed out, we believe through the variety of measures that we've looked at, that there's substantial indication of possible benefit in the duodenum.

What I'd like to do now is turn to a discussion of the follow-up study and briefly mention what the objectives of that would be, what study populations we've considered, the endpoints, design options, and then put up a proposed study and sample size assumptions. I will preface this all by saying that what we will be showing you is an outcome of many discussions that have been held internally. We have had some advice from the FDA, and this is really a start-off place for a discussion about what the follow-up trial should look like.

So, our objectives are to fulfill NCI's and Searle's commitment to patients with FAP and to this field of research. Also our objective is to fulfill the

subpart H requirements, as we've discussed them with the FDA.

In terms of potential study populations that we could look at, one possible group of individuals that we could look at would be individuals who have established disease similar to the population used in this study. Endpoints in that sort of study could consist of a composite endpoint which would look at a variety of FAP-related outcomes, such as FAP-related death, FAP-related cancers, secondary surgeries, development of high grade dysplasias, or other measures of progressive disease.

One could also have a more tightly focused study that would just look at secondary surgeries, for instance, the loss of a retained colorectal segment.

And duodenal disease, as you've heard several times, is an important problem in this patient population, and we could incorporate an endpoint that looked at duodenal disease in either of these designs.

As you've also heard, there is a great challenge for the FAP population in terms of how to manage this disease and how the medical community can

address this disease. Addressing this disease pre-phenotypically may also be a possibility, with an endpoint being time to recommendation for primary prophylactic surgery, the goal being to have young patients complete adolescence so that they can have physical growth and emotional maturity and complete psychological development prior to surgical intervention, if we can show maintenance of phenotype suppression.

If we were to do this type of study, one would be targeting adolescents 12 to 19 years of age who have a genetic diagnosis of FAP but do not have any phenotypic expression of the disease and obviously would have no history of prior colorectal surgery.

The primary endpoint for this study -- again, I point out that this is for discussion -- would be the proportion of patients that reach age 21 prior to having their colorectal surgery, and 21 is based on the Leeds Castle International Polyposis Group Guidelines.

Supportive endpoints could include time to phenotypic expression, time to recommendation for surgery, extent of disease that develops over this

period of time, both gross disease and on the more histopathologic level, and other measures that could look at health care resource utilization, quality of life, and other measures of progression of the disease.

In terms of the study and apropos to some of the issues discussed this morning, one could consider a single-arm study or a two-arm study.

In terms of a single-arm study, one would be relying on a comparison to historical data, and would also be relying on an internal control that would show some sort of reduction of the disease development or polyp development in that cohort.

A two-arm study would raise the issue of what should the comparator be, whether it should be a placebo-controlled trial or potentially even two doses of celecoxib.

The issues that arise around the use of the placebo control I think are fairly obvious and include the question of patient and family acceptance, given the initial results of the study that we presented today. The other is the willingness of individuals to continue on the study even if there's minimal evidence of disease

progression, and of course, one needs to consider the acceptability of the design and randomization strategy to physicians.

The use of celecoxib at two dose levels would allow all patients to receive an active agent in this study.

So, the design that we're putting on the table is a double-blind, controlled trial of celecoxib in individuals with genetically diagnosed FAP and no phenotypic expression of the disease. It would be two treatment groups, celecoxib at 100 milligrams b.i.d. and celecoxib at 400 milligrams twice a day. We estimate that this trial would require roughly 322 individuals randomized 1 to 1. The age inclusion would be 12 to 19. This would be a stratification variable. The duration of therapy would be until the need for prophylactic colorectal surgery, and the endpoint would be a measure of the proportion of individuals developing a phenotype or requiring surgery prior to age 21.

Just to give you a sense of the assumptions underlying the study design, we've been able to determine from the literature that roughly 80 percent of

individuals with FAP have their initial prophylactic surgery by age 21. We're estimating celecoxib, based on the data we have from today's study, would have roughly a 10 percent on this incident rate, so we'd see a reduction from 80 percent to 72 percent in the 100 milligram bi.d. group, so a 10 percent drug effect, a 30 percent drug effect in the 400 milligram group, so lowering the rate to 55 percent. We've allowed a dropout rate of 15 percent, and using a desired power of 80 percent and a p value of .05 with a two-tailed test, we came up with the estimate of requiring 322 patients. We would obviously have an independent data safety monitoring board that would be responsible for monitoring the study and for making sure the trial assumptions were correct.

So, to wrap up, we hope that the design that we presented would serve as a basis of discussion to meet NCI and Searle's commitment to the FAP cohort, to the individuals with FAP, and to fulfill subpart H requirements.

What I'd like to do now is to return the podium to Dr. Needleman.

DR. NEEDLEMAN: May I just have the last slide please?

Our concluding remark is based on the aggregate of the clear history of FAP and the progression from the adenoma to the cancer, the long NSAID epidemiology, the recognition of the COX-2 involvement, and this clinical data, we believe we have a sufficiently persuasive case that would warrant your consideration of our proposed target.

So, thank you for your attention, and we're prepared to answer any questions you might like to raise.

DR. RAGHAVAN: Thanks, Dr. Needleman.

Perhaps while the committee is thinking up their questions, I have one that I think is seminal to the whole presentation and that I've been struggling with since I read the data that were presented.

I understand fully that FAP is relatively uncommon in the community, and you've cited figures of 1 in 10,000 to less common. But even making those calculations, it comes out in my statistics as being about as common as testicular cancer where big trials

have been done internationally for many years. You've taken two of the finest institutions internationally, St. Mark's and M.D. Anderson, and I understand the quality of the research that would be done in those two places. But you do have some real issues that relate to statistical power.

Why did you content yourselves with doing just a two-center trial when you could have recruited four or five other centers, doubled or trebled the numbers, and increased the statistical power of your observations?

DR. NEEDLEMAN: I think what I'll do is call on people from St. Mark's and talk about what it takes to accrue a trial of an adequate size to have the number. Understand that this already was four times bigger than any trial that has been performed in the past, but the reality is, let me call on Dr. Robin Poole from St. Mark's Hospital. Sorry. Robin Phillips.

DR. PHILLIPS: Thank you very much. I think one of the issues you're addressing is the difference between incidence, which is 1 in 10,000, and the prevalence, which is 3 in a million. So, if you

actually look in the United States with a population, say, of 255 million, you're only going to have 3 in every million of those. That's not very many, and it's far, far, far less than testicular cancer.

DR. RAGHAVAN: I accept that. On the other hand, you do have the, I guess unattractive feature, that there's such tight clustering that you don't have to go looking for the cases. As I said, you've got St. Mark's as one pivotal center in the United Kingdom. Between the United Kingdom and the USA, your millions of population start to increase.

I do fully understand and I'm sympathetic to the problem that previous studies haven't addressed big numbers. I'm not critical of the difficulty you've encountered. I guess my question perhaps is not unreasonable. While I understand the illustriousness of St. Mark's in Britain and of Anderson in the USA, you've already cited Roswell Park has a program. I don't know it to be a fact, but my guess is Dr. Kelsen's team have something doing at M.D. Anderson. There are many centers nationally that do accrue patients with this problem.

And my question stands. Why did you choose to do a two-center study keeping the numbers down when you could have gone to a five or six center study and potentially doubled your numbers?

DR. PHILLIPS: Just from a St. Mark's perspective, there is obviously a variability of the delivery of care and the quality of these videotapes and the quality which has been a tremendous issue that you have addressed here in actually being able to get numbers associated with it.

I've been involved with six randomized trials in polyposis to date. One is the CAPP 1 study, which has involved a trans-European study, and that is doing exactly what you have said. And it has been a complete disaster in terms of the quality of data that is coming in. You have different endoscopes. You have different experience. You really find that it is virtually impossible to determine the endpoint that you're after.

I'd make one other point. In health care systems which are government orientated, registries have been built up so that it is fairly easy to go to some of those and access patients with large numbers, large

numbers of patients. When you're in a private health care system, if you go tracking individuals, you may well find individuals who are not covered by health care insurance, and you are left with a very difficult problem. Because of that, the development of registries in the United States, in particular, has not been as strong as in some areas where there is government-funded health care.

DR. NEEDLEMAN: I'd point out two things. First of all, that is captured in the whole concept of the follow-up trial. I'll remind you, indeed, I think that was the thinking of the agency for priority review. We're looking for an unmet medical need, and this is the proof of concept population. We zeroed in on a dose to establish efficacy and its limits of side effects, and we committed to the larger trial which we now have the reasonable ability to project the incidence, the dose, and to continue. So, the context really of a priority review is the commitment to the much bigger trial in many centers, and here we can now calculate the appropriate population of patients and doses.

DR. RAGHAVAN: Dr. Margolin?

DR. MARGOLIN: I have a set of three related questions. Then I have a separate question. The set of three related ones is technical, which is as follows.

DR. NEEDLEMAN: I'm sorry. I can't hear you.

DR. MARGOLIN: Three related technical questions and one about the follow-up study design.

The technical questions are as follows. What fraction of the colon surface is supposedly represented by the area involved by this tattoo cloverleaf and how that was selected, how far up the colorectum it is, and whether the endoscopist who did the pre and post-treatment assessments -- I assume it's the same endoscopist because of the tattoos -- is blinded to the treatment assignment? That's the first cluster of questions.

DR. NEEDLEMAN: I'll keep catching.

DR. MARGOLIN: The other one is completely separate, so I'll wait.

DR. NEEDLEMAN: Let's call on two people. Let's call on Ernie Marks, and then I would also like to call on Marina Wallace who has really been engaged in the visualization. First, Ernie, about the location of

the tattoo and then Marina in terms of evaluation.

DR. MARGOLIN: And whether it was chosen based, in any way, on the number of polyps in that area or the appearance or the ease of following this.

DR. HAWK: I think the gastroenterologist will speak specifically to this because they were ones doing the actual work.

But the tattoos, as I think I've pointed out I think in my eligibility slide, were based upon having a number of polyps, at least 5 polyps, within a focal area that could be visually assessed repeatedly over time so that you would have some degree of change, possible at least, from baseline. Those were placed in the rectum in all patients with a colorectal endpoint.

Marina, do you have more to add?

DR. WALLACE: Answering your question in two ways, firstly, the gastroenterologists doing the endoscopies had no idea which treatment group the patients were on, as neither did any of the people involved in the study itself.

If I can have slide 204.

DR. RAGHAVAN: Could you identify yourself,

please?

DR. WALLACE: Sorry. Marina Wallace from St. Mark's Hospital.

This is a picture you've seen before. As you can see, at the baseline picture, we chose an area which had a cluster of polyps that could be reassessed at 6 months. So, the tattoo was placed in the center and the photographs taken around to represent the area of that tattoo. This distance view is approximately 2.5 centimeters.

At 6 months, the endoscopist went back to the tattoo and took the photographs in the same quadrant way, north, south, east, and west. You can see this shown by the polyp here, which is the same as the polyp here at baseline.

DR. RAGHAVAN: Dr. Kelsen?

DR. KELSEN: The ultimate goal of this is to decrease the incidence of cancer. Once the study was completed and the patients were unblinded, do you have any data on how many remained on Celebrex and what the incidence of cancer was in this population?

DR. NEEDLEMAN: The question really was how

many patients remained on Celebrex after the 6-month trial. The reality is FAP patients are engaged in very many other trials, and we haven't been tracking those patients because they moved into other test agents. Nor do we know which then patients then started to use Celebrex because it's readily available to them.

DR. KELSEN: I would imagine someone who had a photo like you showed with this decrease who was unblinded to Celebrex may have stayed on it.

Do you have any data on the incidence of cancer in these patients at this point?

DR. NEEDLEMAN: I don't think we do in this time period. Again, that's a primary objective in the continuing, long-term follow-up study.

DR. RAGHAVAN: Dr. Nerenstone?

DR. NERENSTONE: Is this the appropriate time to ask about the follow-up trial?

DR. RAGHAVAN: Sure.

DR. NERENSTONE: Do you have any safety toxicity data on children perhaps with JRA who have been on Celebrex for a long time, and could you just address that briefly?

Then I have a question. Why would you use an inferior dose? You're talking about a randomized trial, and it appeared to me, if you believe the data, that the 400 b.i.d. dose is clearly superior to your 100 b.i.d. dose. So, why would you take that into a subsequent trial?

And also with children, obviously their meter squared, their body mass, is less, so you would be getting more medication although there are other pharmacokinetic considerations perhaps in the way children metabolize medication. Are you looking at that in terms of trying to equate the dose of 400 b.i.d. in adults to children?

Could you just comment on those issues?

DR. NEEDLEMAN: You asked two questions.

Let's go back, first of all, to the consideration of why the dose was selected. Let me bring up slide 92 again please, and in that slide we looked at the dose-response curve. But this in the phase III trial and we established the dose-response curve. This was just one of the parameters of rheumatoid arthritis. Indeed, the 400 and the 200 were

at the plateau. Understand that the FAP trial was done in the midst of the phase II data when we didn't have this complete data set and have the safety data.

Let me show you slide 94, please. In fact, if we look at the aggregate of an index known as the American College of Rheumatology Composite Arthritis Index, which takes into account six different parameters in the efficacy -- as a matter of fact, Celebrex was the first drug ever approved that fulfilled this index. If you look at the dose-response curve, you see we find, while 100 is clearly in the dose-response curve, we didn't achieve the maximum responses until we hit 200 or 400. If we had the aggregate of the data, it might well be true that the 200 was there.

So, in order to design the trial to maximize the opportunity to see the patients, knowing that there's a limited number of patients, we bracketed the high dose and the low dose, which indeed captured the regular dose. So, otherwise we would have had to wait another year or so to continue the trial.

Let me turn to the safety --

DR. NERENSTONE: But my question is, why will

you take that into the second trial? I understand why you put it in your first trial, but now that you have the results and it appears that the 400 is better, why would you take that inferior dose to a second trial?

DR. NEEDLEMAN: You mean the 100.

DR. NERENSTONE: Right.

DR. NEEDLEMAN: We'll come back to that point.

Briefly, the question is, do you want a placebo and the fully active dose, or are you going to put the patients at risk at placebo or use the low dose and look for a statistical difference between the two? And indeed, that's what we're prepared to talk about here with you and with the agency, what should be in the arms of the trial.

Let's turn to the safety data, and I'll call Dr. Jim Lefkowith from Searle, who has reviewed our safety data and could track what we know about children as well.

DR. LEFKOWITH: Can I have slide 212, please?

We have actually substantial data in the NDA showing that we can safely dose patients down to age 18.

At the time of submission of the NDA, we had sufficient preclinical toxicity to support dosing down to age 12. By the end of this month, we'll have additional studies complete in order to support dosing down to age 2.

Now, although we didn't have patients younger than age 18 in the trial, we nonetheless had a lot of individuals of small weight, elderly females. We did a safety analysis looking at a cut point of 55 kilograms, specifically for the dose relevant to this trial, celecoxib 400 b.i.d. The incidence of adverse events, segregated out by weight, is the same in both populations at this dose. So, there's no indication that dose adjustments need to be made for individuals probably relevant to the population of the study.

DR. RAGHAVAN: Dr. Surawicz.

DR. SURAWICZ: I have a couple of questions. The first one should be fairly easy. These are young people. What about the safety data for pregnancy if women should become pregnant on this drug?

DR. NEEDLEMAN: Actually in our safety data, there were really a limited number of patients in pregnancy. As with NSAIDs, in the label, it's

contraindicated in pregnancy. So, in the actual trial or in the FAP trial, I think there were no pregnant patients and there is no isolated population that's been so studied.

DR. SURAWICZ: And is that your recommendation then in the rheumatoids, in your arthritis patients?

DR. NEEDLEMAN: Yes, that's right.

DR. SURAWICZ: That they stop if they become pregnant?

We know that for spontaneous polyps, that there are some influences of things like exercise, diet, vitamins. Were any of these things controlled for in this study?

DR. NEEDLEMAN: I'll again have to turn to the experienced group. Maybe we should hear from Giddeon Steinbach about the trial, his experience, and the inclusions or those details in the M.D. Anderson trial.

DR. STEINBACH: To answer your question, those particulars were not used as exclusion criteria. Only NSAID use was used and steroids were used as

exclusion criteria.

It's unlikely that these would have had a strong influence that is particular in the 400 versus placebo group, but that is not determined.

DR. NEEDLEMAN: Thank you. I'd like to also --

DR. SURAWICZ: Actually my question -- can I clarify my question? It was more did you look at these factors in all of the patients to make sure that this wasn't an additional variable that might have influenced your results. For instance, the exercise or vegetarian diet or anti-oxidants in their vitamins. There are a lot of things that may play into this as well as the celecoxib. In this study, did you look at this, and if not, are you planning to in future studies?

DR. STEINBACH: Well, the data was collected and it is in the database and is collected in the Willett Frequency Questionnaire. That hasn't been analyzed in particular to the various groups.

DR. SURAWICZ: Can I have a final question?

DR. NEEDLEMAN: We have an additional comment. Dr. Monica Bertagnolli from Women's and

Brigham's.

DR. BERTAGNOLLI: It's a very interesting question. There have been trials done in FAP patients. The one I'm most familiar with is the one from Cornell that was done by Jerome DeCosse several years ago where he looked at anti-oxidant vitamins. He looked at vitamin A and vitamin C, and he also looked at very, very high doses of fiber in FAP patients. That was actually one of the largest studies that's ever been done in FAP individuals with a retained rectal segment. Unfortunately, even with an extremely rigorous application of both the anti-oxidants and the very high fiber, they were unable to show -- and I believe the total study duration was over 3 years -- a statistically significant difference between the treatment and the study group, although there was a trend.

So, given that kind of pressure, that kind of fairly well done study, I think it's probably unlikely that there would have been a significant difference among the groups that could be related to those. I can't speak about activity or the other known variables involved in sporadic colorectal cancer.

DR. SURAWICZ: Then my last question. Was there no biopsies in this study? This was just purely what the polyps looked like grossly and how big they looked?

DR. NEEDLEMAN: Ernie Hawks, comments about biopsies in the patients, please.

DR. HAWK: Indeed, biopsy samples were taken of the normal appearing flat mucosa as well as some of adenomas. They were used, however, or are being used, rather, for assessments of various biomarkers that might correlate with mechanisms of activity and provide us further insight. They were not in any way systematically used to support the clinical data per se. Rather, we're exploring a number of other things.

DR. SURAWICZ: Well, I'll tell you why I ask. My concern would be if the polyps got smaller but the adenomatous tissue is still there and the cancer risk is still there, then it's possible that when cancers develop, they could metastasize even quicker because they'd be more similar to the flat lesions that you see in other polyposis syndromes. So, I would encourage the use of histology, as well as your biomarkers, in future

studies to evaluate where polyps were and the areas adjacent to the polyps that now look smaller to make sure that the risks isn't being transferred.

DR. NEEDLEMAN: I think that's certainly reasonable. I would, though, still turn back to the wealth of epidemiology with NSAIDs which saw a successful application in all stages from the pre-cancerous even to the cancerous state as a predictor. Probably that was part of its COX-2 inhibition, but following those polyps, especially when we have more biomarkers, would be very powerful.

DR. RAGHAVAN: There are a number of people that are identified as having questions. So, I'll get to you all. Dr. Santana is next.

DR. SANTANA: I have two questions about biology and then one comment about your proposed study. I thought I heard a comment earlier in the day that the COX-2 expression may be different in rectal tissue versus other parts of the colon versus extracolonic lesions. Could somebody clarify that biology for me?

And then a related question. In your MIN mouse model, you've very nicely demonstrated statistical

reduction in the number of adenomas. Can you tell us about the survival of those mice that were treated?

And then I'll come back to the study.

DR. NEEDLEMAN: Let me call up about our experience with the immunohistochemistry from the samples we had in terms of is it rectal or colon. The fact of the matter is we find good evidence of immunohistochemical presence of COX-2 throughout the intestinal tract.

The slide that you saw before, number 704, indeed shows that -- I think we did archive 25 immunohistochemical analyses of FAP patients. In 704, you'll see, in the upper quadrant, all of those were positive for COX-2. No COX-1 present, and traditional immunohistochemistry that's blocked out by pre-absorption.

While we don't have many rectal segments, we always find it throughout the intestinal tract. In fact, the amazing thing, indeed, is you also see it even in the metastatic tissue. So, we think, in fact, that's not a warranted response.

About that concern, let me bring up slide 22

also because another way to look at that is actually not immunohistochemistry, but the actual experiments that you see in patients. You'll notice on the right side -- well, you could see, if you just follow the 400 milligram b.i.d., when you look by cecum to ascending colon through the sigmoid and the rectum, comparable responses, indeed, to the agent. So, there is evidence of immunohistochemistry, but the real proof is in the responsiveness as you look at polyps across the --

DR. SANTANA: Do you have any data about extracolonic lesions like the desmoid lesions, et cetera, in the expression of COX-2?

DR. NEEDLEMAN: In these patients? No.

DR. SANTANA: Or any patient biologically. Does anybody have any data?

DR. NEEDLEMAN: I don't know if we have desmoids.

You know, we have mapped hundreds and thousands of tissues looking for the presence of COX-2 gene. That's pretty interesting. The places where you don't see it so far in our experience is neuronal tissue, neural blastoma, glioma. The richest incidence

of expression in cell types are around colon cancer, but the reaches are variable in other places, for example, prostate and breast. So, in every epithelial tissue we studied, there is COX-2 early. You could see it in cancer and you could see it when it's metastasized. It's particularly pronounced in the targets that we're going with the NCI. So, the SAP, the Barrett's esophagus is rich. Actinic keratosis is almost crystalline prostaglandin and COX-2. The same thing with superficial bladder. On the other hand, the case in HNPCC is quite a bit less than the SAP patients.

Now, then when we make decisions about where to take COX-2 inhibitors, we don't just look at the immunohistochemistry, we look at the completeness of expression in the cells, and we look for is there any epidemiology to make the choice and is it a discrete population that we could really have enough of an incidence to follow it at trial.

You asked about the MIN mouse. I guess I don't know the answer about the survival. Jaime Masferrer, would you come up please?

DR. MASFERRER: In that experiment, as I

mentioned, celecoxib was very well tolerated by all the animals. There was no effect or any toxicity in the animals. The experiment was terminated in the two modalities at day 80, so we could count the tumor sizes.

DR. NEEDLEMAN: The striking thing in the MIN animals, which is really quite dramatic, is any time in the course from pre-exposure to polyps virtually to death, once you start COX-2 treatment with Celebrex, you start to turn around and have regression in the number and in the volume of the polyps.

DR. SANTANA: But you don't know if when you stop the medicine, it comes back and if it impacts on survival. So, it's unfortunate you didn't follow the mice for longer periods of time.

DR. NEEDLEMAN: I agree. Again, the NSAID epidemiology, while it didn't have all the controls and had the safety problems, really has a risk factor reductions of .2 to .5 both in the cancer and in the pre-cancerous state.

DR. SANTANA: Then one last comment and it's just a comment on your proposed follow-up study. I think it's laudable that you're going to study an

earlier age population which clearly, if this intervention is going to have an impact, will be very important for them.

But I'm concerned that if 50 percent of the patients develop adenomas or polyps at the age of 15, which is what we all accept as probably correct, and you're starting at age 12, you're already going to throw out a lot of patients if you're excluding patients who have phenotypic disease as a criterion to be enrolled on the trial. So, you're going to start with less and less numbers just because of the issue that the majority of patients by age 15, or at least 50 percent of them, may already have adenomas.

Then the second comment is you have to be very careful because anytime you introduce an intervention, the surgeons and the people who are going to consult on these patients are going to pull back. If your endpoint is when is the surgery going to occur, when is the surgery indicated, that everybody have very clear guidelines of what are the criteria to submit patients to surgery and, therefore, having considered them a failure on the study. Because I fear that with a

new intervention, people are going to say, these are the two options. I'd rather stay on the medicine for a long period of time rather than going to surgery unless the criteria are very well defined.

DR. NEEDLEMAN: There are two parts to your question, and I want to come to the second. We are really open to -- and we've worked hard with the agency -- advice in the final selection of the patient population and how to do it in the follow-up study. So, here we've picked a point that we think is reasonable. I think this time there's really advice that's usable.

The other question really -- it could even balloon -- is I think there were concerns that if you have a pharmacological agent, that you'll interfere with the standard of care and that you'll change it. I think that's important to comment about. Let me first call on Monica Bertagnolli, and then I would like to call again on Robin Phillips.

DR. BERTAGNOLLI: I just want to introduce myself. I'm a GI surgical oncologist presently at Brigham and Women's Hospital who's had a long interest in the management of FAP patients and in their surgery.

You raise exactly the issues that we've really been struggling with as consultants to this group in how to design a follow-up study. As Dr. Needleman said, we'd love to have advice.

One of the things that I'm sure you all took away from the patients with FAP who were describing what their lives are like, one of the things we struggle with as clinicians is we have very young individuals who are 14, 15, 16 years of age, and are facing a life-altering surgery to try to prevent them from having cancer. We're always as clinicians trying to balance the risk of letting them get a little older, a little more mature through their high school football season, to the next summer, with the risk of are they going to develop an advanced lesion in the meantime.

So, if there's a way to design a study that will allow us to meet that need and to give them a little margin of safety, to me that would be a clinically very important thing. Again, how to design the trial to achieve that and to know that we have safety and know that we have continued to follow them very, very carefully is what we're all struggling with.

DR. NEEDLEMAN: As I call Dr. Phillips, I would ask him to talk about his decision tree about the patients, but I'm reminded to mention that the patient who had a concern in the letter, while I'm glad he acknowledged the analgesic effectiveness of these drugs, maybe Dr. Phillips should also talk about would that preclude a clear decision about when the surgical intervention is appropriate.

DR. PHILLIPS: I'm sorry. I didn't really introduce myself before. I'm a consultant surgeon at St. Mark's Hospital in London, and I'm the Director of the polyposis registry in London and also the Honorary Administrative Director of the worldwide polyposis registry, which is called the Leeds Castle Polyposis Group.

Operating in a child, the first thing is you have to wait for the phenotype to be present. So, I don't know any surgeon who would operate on a polyposis patient in the absence of the phenotype. So, if there is a delay to the presentation of phenotype, there will be a delay to the start of surgery.

As far as once the phenotype is discovered,

then under the age of 20, certainly in the United Kingdom and perhaps under the age of 18 in the United States, the decision about when to operate is more to do with family convenience than it is to do with the fact you've diagnosed the phenotype. So, in practice I might diagnose a young child at the age of 12 with polyposis. In the absence of high density, it may be more convenient for them to have their operation at age 15. This is normal standard practice around the world.

The risk of cancer in this age group is infinitesimally small. These children are not necessarily psychologically or physically prepared for the sort of surgery that we wish to do. So, the standard level of care at the moment is that we wait, but if we could wait a bit, added with the added safety of polyp reduction and the knowledge that the parents involved would not be so anxious because of polyp reduction, that would be well worthwhile.

DR. RAGHAVAN: Dr. Lewis?

DR. LEWIS: Thank you. I have two questions. One deals with the dose that was used and the other was whether we can get some information on the phenotypes of

these patients and how many actually had attenuated Apc because of the older age group that was studied, the small number of polyps in many of these patients.

But my first question is, do we really have an adequate dose-ranging study to determine what the effect of the drug will do? The animal studies we showed, if I wrote it down right, the dose was 500 milligrams to 1,500 milligrams per kilogram. If you divide 800 milligrams by a human body weight, we've got between 10 and 20 milligrams per kilogram. What do we know about the higher doses in the animals, and do we have any information on higher doses in this group? Would we melt away more polyps in these patients?

DR. NEEDLEMAN: So, two questions. Let's first deal with the dose.

When we compare animal data to human data, we analyze area under the curve exposure and then mechanistically work out what are the concentrations that you need to suppress the COX-2 activity while we analyze COX-1 activity. We do the same thing in the patients.

Let me talk about the adequacy of the dose.

There's always also the interesting question, if you've got a response and you have so much safety in the drug, why not just keep pushing and pushing and maybe the efficacy is due to something else than the mechanism.

Our whole intention of the trial is to identify the dose that's necessary to achieve inhibition of COX-2 from our analyses both in arthritis and the patients without inhibition of COX-1. If you want to push to enormous kinds of levels, you'll take a burden of side effects and you'll really not know the mechanism.

So, let me first provide you with some data again. Remember I showed -- maybe I should reshow -- the dose-response curve with a threshold response at 100 and it flattens out at 200 and 400 milligrams per kilogram.

There are two ways we assessed the adequacy of the inhibition of the enzyme in the patient population. 67, please.

Now, this was studies not in these patients, but we also ought to talk about the dose in the patients because some of them will be lacking a colonic segment

and you might have a change.

Here what you see is the -- what we did is we pulled the blood samples from the patients, and you could stimulate these patients after -- here's single dose therapy. You could stimulate the blood with endotoxin which turns on COX-2 in the monocytes, and I'll compare prostaglandin E which is the COX-2 product versus thromboxane. Here's placebo. Here's the 100 milligram, and here's the 800 milligram dose of either the 400 or the 800 milligram causing comparable suppression to what you see with ibuprofen.

Slide 68 is work from the University of Pennsylvania and the laboratory of Garrett Fitzgerald where they looked in normal urine at the urinary excretion of the prostacyclin analog, the PGI2 which reflects activated COX-2 probably along the vascular bed. You see, indeed, we have this 80 percent suppression of the prostaglandin in the urinary marker that you look in patients.

On the other hand, if you look at slide 97 and look for the COX-1 equivalent, here you see measuring thromboxane in the serum. The white is the

placebo. No effect at the various times of treatment. Here is the standard non-selective naproxen, COX-1, COX-2. Here is 400 milligram b.i.d. No effect at all on serum.

If you look at slide 99 also, you then see the absolute separation in the patients with no effects on platelet aggregation. Platelets are COX-1 only. They don't have an inducible enzyme because they don't have a nucleus. Here up to 1,200 milligrams b.i.d. of Celebrex, no effect on platelet aggregation on a standard inhibitor.

So, our trials are designed totally to be mechanism-based effect. We have already arranged that we think it's consistent with the arthritis dose and achieves it, and the 1,500 milligram that you saw in the diet is equivalent to the upper arthritis doses in terms of area under the curve. So, that was the basis of that.

I'll have to ask Ernie or someone else, what do we know about the phenotype of the patients in the enrollment, about the Apc? Gary Gordon.

DR. GORDON: Can I have slide 303?

We actually have some information about the genotype of the individuals and were able to look at attenuated FAP genotype versus non-attenuated mutations and also those individuals in whom the mutation was not found. Although we have a limited number of individuals with the attenuated phenotype, you can see they fall within the much broader response curve that's seen by the overall population in the study.

DR. RAGHAVAN: Ms. Forman?

MS. FORMAN: From your studies, you certainly can claim that Celebrex reduces the size and sometimes, I guess, numbers of polyps. Are you also claiming it reduces the incidence of cancer, and how would you treat that in your labeling?

DR. NEEDLEMAN: We are only claiming about the reduction and the regression of polyps in the FAP patient. No claims about cancer or its outcome, no claims about any other colon cancers. That will be the subject of the follow-up trial in FAP, and then its role in cancer will be the subject of the ongoing sporadic adenomatous polyposis.

We're very, very careful just to ask for what

the data says, and we're committed to this being an adjunct to the armamentarium of the surgical approach to the agent. So, no claims about cancer, no claims about anybody but FAP, but within the context of its big safety record, it then has the potential for the longer trial.

DR. RAGHAVAN: Dr. Sledge.

DR. SLEDGE: I think I'm wrestling with what I suspect several people around the table are wrestling with here which is the issue of what represents clinical benefit.

As I look at clinical benefit endpoints with NSAIDs and with this agent, the possible ones -- you've already listed them, but the ones that seem fairly obvious are you could eliminate surgery potentially. You could delay surgery. You could eliminate screening or potentially reduce the frequency of screening. You could reduce colon cancer risk, and you could reduce colon cancer mortality. Now, the last two of those are based largely, presumably, on population, epidemiology studies with NSAIDs.

I guess a question that I have for your

clinicians who are here today is, based on the 30 percent reduction in polyp burden that was seen in this particular study, can you tell me if there's any individual patients who you think would meet any of the clinical benefit endpoints that are listed, eliminating or delaying surgery, reducing the frequency of screening? Would it make any real difference to any individual patient?

DR. NEEDLEMAN: Yes, I think that's the heart of the question.

DR. SLEDGE: I'd like to ask the clinicians.

DR. NEEDLEMAN: Right, and so I'd like to sequentially first call Dr. Hawk and then Dr. Bertagnolli and then Dr. Phillips to answer the question.

DR. HAWK: Slide 220, please.

Very briefly I'd like to put the results in context one more time. We've seen this slide before in the primary presentation. In fact, most of the slides you see from me will be those.

It's important in our view not to focus merely on the 28 percent mean reduction or 32 percent

median reduction in adenoma burden that we saw in the 400 milligram dose group in isolation because, as you correctly point out, we saw reductions in number, depicted here, and several individuals with very substantial reductions in number.

222, please.

To complement that number finding, we saw reductions in adenoma size that I reviewed for you.

224.

We've also shown you consistent effects, benefits in the colorectum, here in the rectum alone, although the benefits were similar in the other segments of the colon, a very substantial and consistent effect.

And then 226.

Consistent benefit by one of the measures by video endoscopic evaluation in the duodenum. I'll point out again that while I'm not a practitioner caring for FAP patients -- you'll hear from those in a moment -- there is no approach really to duodenal disease in these patients. So, we feel this is a significant benefit, but I'll leave it to the others to elaborate on.

DR. BERTAGNOLLI: I think one clear benefit

would be the 16-year-old who was genotypically positive who could have complete phenotype suppression. If they don't show up with an adenoma, I don't think we would recommend that they have surgery.

I also echo the concerns of the group. I certainly would not mean that would reduce surveillance. I'm thinking that reduction in surveillance is the furthest on the list there. I wouldn't think we could reduce surveillance. I would also add random biopsy because there is also a concern that maybe they're not polyps, but there might be something else there. So, that very much would need to be studied. But clear phenotypic suppression would be a benefit to young patients.

I think the patients, of whom there are many throughout the world who have a rectal segment remaining who develop a burden of adenomas in that rectal segment, an indication for surgery is an increasing polyp burden there. If we see a marked reduction for those patients along with continued surveillance, I think that's a clear clinical benefit because you have 50-year-old patients who then need to undergo a proctectomy, and

that's a significant life-altering event as well.

There are two other places where there could be benefit. Obviously duodenal disease, as Dr. Hawk has brought up.

But then there's a fourth one. There are a lot of patients now who have had the pouch procedure. They've had the entire colon and rectum removed. I had a patient this year turn up with cancer in a rectal pouch. Now that we're seeing these people 20, 30 years after their pouches, we're starting to see adenoma formation in the remaining pouch. It has been reported before in ileostomies, but we might even see an increased frequency in the pouch. That's another situation where we really don't have anything to do other than give these people a permanent ileostomy. So, those four very specific incidents, I think we could see a benefit.

And I'll just reiterate again I wouldn't reduce surveillance in anybody.

DR. SLEDGE: Perhaps I didn't make myself clear because I think this is real important. I agree that those are potential benefits, and I think the

potential benefits are obvious to everyone around the table.

Given the data that you have here, if we handed you this drug tomorrow and said it was approved, how would it change your clinical practice?

DR. BERTAGNOLLI: I would, given its safety profile, use it for phenotypic suppression or attempt to use it for phenotypic suppression, again if the follow-up study showed that that was successful. I would certainly use it on my patients right now who have rectal segments that are remaining that have adenomas present in them. And I would use it in patients who have pouches who also seem to develop adenomas there, and I would use it in any patient with a burden of duodenal disease.

Right now I'd have to do it, again, without the follow-up study data, which is important. But I think that the immunohistochemistry, the NSAID epidemiology, all the lines of evidence are proceeding in the same direction, and I think that plus the safety profile are compelling enough that I would use the drug right away.

DR. NEEDLEMAN: Dr. Phillips, anything to add, please?

DR. PHILLIPS: From the United Kingdom perspective, I would agree entirely with Monica. I would use it now in 100 percent of my polyposis patients because the major cause of death is upper gastrointestinal cancer. We have no treatment for this disease at the moment, and because of that, any treatment would be a worthwhile treatment in these patients.

All the other points made about the difficulty of managing the rectal segment in some youngster who might have to face an operation to remove their rectum with a risk to erection and ejaculation, the difficulty sometimes encountered with desmoid disease that may make it impossible to do that operation means that we need to be able to be forced into these operations rather than just choose to do them. So, I need other things in my armamentarium.

DR. RAGHAVAN: Dr. Needleman, before the next question, I think we're starting to run quite late, and I think we would take each of your speakers as beyond

reproach. Therefore, we'll accept one person for the team answering a question rather than three answers that say similar things.

Dr. Brand is next, followed by Dr. Albain.

DR. BRAND: Because of the concerns of duodenal polyps, when you examined -- this is my first question -- did you use a side-viewing endoscope? Because most of these tumors are ampullary, and you can't always see the ampulla well when you use a forward-viewing endoscope. So, were side-viewing endoscopes used with particular attention to the ampulla?

My second question is regarding the number of polypectomies performed. I thought at some point if a polyp was found that was greater than a centimeter, that during the procedure a polypectomy would be performed. Was there any differences in the groups about a need for polypectomies?

DR. NEEDLEMAN: Two questions. I understand I was admonished.

(Laughter.)

DR. NEEDLEMAN: But we do kind of like to

answer things with data. First I'd like to call on Marina Wallace to talk about the side-view endoscopy.

DR. WALLACE: Two answers to the question. The simple one is, yes, we used a side-viewing endoscope in all patients, which is our current minimum clinical care for these patients. They all get side-viewing endoscope. In fact, to improve both the quality and to make the discomfort less for the patients, we actually obtained a diagnostic side-viewing endoscope, which is much thinner. That's what's currently used in ERCPs. So, we used a thin, side-viewing endoscope to obtain the best pictures possible.

DR. NEEDLEMAN: Ernie Hawk, you talked about the number of polypectomies in your slide. Do we have any more evidence than what you've seen?

DR. HAWK: That would really be a question for one of the actual study participants.

DR. NEEDLEMAN: Marina, you count as one answerer. This is good.

(Laughter.)

DR. WALLACE: Polypectomies were performed as clinically indicated before any videos were taken at

baseline. So, if a patient had a polyp that looked suspicious, this was removed endoscopically and sent for histology. Then the photographs and video were taken.

At 6 months, if large polyps were seen, these were included in the final videotapes and photographs, obviously because they were important to the final study data. They were then removed and sent for histology.

There were no cancers in the group at 6 months.

DR. BRAND: But was there any difference between the groups in terms of developing? Because 6 months later, now you have a polyp that would be increased in size. Was there any difference in the groups?

DR. WALLACE: I can't answer that question.

DR. NEEDLEMAN: Gary Gordon, do you have an answer to that?

DR. GORDON: All the patients had the 1 centimeter or greater polyps removed at baseline. This was a relatively short period of time, and there was a relative paucity of polyps removed. I can't give you an exact count by group, but overall there were very few

polyps larger than 1 centimeter removed at the end of the 6 months.

DR. RAGHAVAN: Dr. Albain, followed by Dr. Jacoby, followed by Dr. Avigan.

DR. ALBAIN: Yesterday morning we were struggling a bit with the concept of sentinel lesion or indicator lesion and what goes on with that lesion versus what goes on in the rest of the body for the particular indication yesterday morning.

Could you comment on the correlation between your primary endpoint target area and what went on in the rest of the bowel mucosa studied? In particular, were there any circumstances where you had a response or a decrease in number in your target area for your primary endpoint, yet there were more polyps elsewhere?

One of the open hearing statements or one of the letters we received also was concerned about the representativeness of the indicator area.

DR. NEEDLEMAN: I'm not sure, so I want to recap. Are you first concerned of the choice that the polyp is the surrogate marker?

DR. ALBAIN: No, no. I'm talking about the

area you chose to study for your primary endpoint, your cloverleaf area. How representative of what was going on in the rest of the mucosa --

DR. NEEDLEMAN: Got it.

DR. ALBAIN: In other words, did polyps decrease in that area, yet perhaps increase elsewhere? Did you do any correlative studies of that with one of your secondary endpoints, the video studies?

DR. NEEDLEMAN: Good. Dr. Wallace, please.

DR. WALLACE: As the data showed in the presentation, we took two endpoints. The primary endpoint was the reduction in polyp size, and as you say, this was a focal point. However, I think the global video assessment takes into account your concern that you may have reduction in one area but severe polyposis developing in the other area. I think the fact that the global video response matched the focal response removes that worry, removes that concern.

DR. ALBAIN: Patient for patient it matched? I realize your percentages were similar, but in a given patient, could something different have been going on elsewhere than in your primary endpoint?

DR. WALLACE: No.

DR. RAGHAVAN: Dr. Jacoby.

DR. WALLACE: Gary Gordon I think wants to answer.

DR. GORDON: Can we show slide 21 first?

So, again, to address the points that have been brought up, there was substantial benefit observed in the 400 milligram group in terms of the percent change from baseline in the number of polyps.

If you go to slide 22, which is the overall assessment by videos, again addressing the question that you brought up of is there a difference in the different parts of the colon, which I believe was refereed to in one of the letters, again as Dr. Needleman and Dr. Hawk have indicated, we show responses in all areas of the colon as well as the rectum.

And then your precise question I think was, is there a correlation between the focal assessments and the overall video assessments on a one-to-one basis? If you show slide 50 I believe, what we can show here is that given the vagaries of these measurements, in fact there is a significant relationship between, on a one-

to-one basis, the video assessments, as measured here with worse, better, and the percent change in rectal polyp number by increase/decrease with a better result being up in this box. There's a substantial agreement at a .05 level between the two measures.

DR. JACOBY: Dr. Needleman, I have two questions. The first is I know you attempted to exclude NSAID users, but many patients are unaware of over-the-counter preparations containing aspirin or NSAIDs. The placebo patient, the single patient who had a 50 percent reduction in polyp number, I'm wondering if there was any indication of NSAID use, and how rigorously did you monitor the NSAID use? I'm more concerned about the celecoxib treatment groups. Did you look at thromboxane B2 levels, platelet function, drug levels to specifically exclude NSAID use in those groups?

DR. NEEDLEMAN: Dr. Gordon?

DR. GORDON: As part of the entry criteria into the protocol, we did ask people to exclude NSAID use and did provide acetaminophen for them. We had different washout periods as well, depending on the extent of prior NSAID use. We did have a few

individuals during the course of the study who did identify that they were taking these medications usually for short periods of time.

If I can have slide 344. What we attempt to do here is to show those individuals who used either NSAIDs or corticosteroids, generally inhaled corticosteroids, and show how they were distributed in spite of the fact that they weren't supposed to be taking the drug. You can see again they fit within the broad responses seen within the overall study population.

DR. NEEDLEMAN: Do we know anything about the 50 percent reduction in the placebo?

DR. GORDON: The individual who had the dramatic response in the placebo group -- we do not have any clear-cut explanation for why this patient had this change.

DR. JACOBY: My second question is based on my experience with the Apc mutant mouse model, I have a concern that we do no harm. One concern I have is if the polyps are partially regressed in a manner where they're flattened, it may be more difficult to visualize

them. We're saying that we'll continue surveillance, but will the surveillance be as good if the lesions are flattened? What method would you recommend, particularly for your follow-up study, to look at this? Magnifying endoscopies, spraying of dye can be done, but it's very time consuming if there's a large area to examine. I'm wondering what your thoughts are on this.

DR. NEEDLEMAN: Dr. Phillips?

It is a different kind of a problem. Now you're worrying that it might be too good and it would mask it.

DR. PHILLIPS: Endoscoping a patient with dense polyposis, it can be very difficult to see a small carcinoma. If you can drop down the number of polyps in the rectum and thin them out, it makes it very much easier to see the carcinoma that is there.

Perhaps hidden in that question is also the question about carcinoma development in patients who have taken Sulindac in the past. I think that the question is whether they really did develop cancer in that way or whether a cancer was already there and it was missed, and when the Sulindac was taken, the polyps

went down and then eventually you could see the cancer.

At entry into this trial, we excluded one patient who was exactly in this category, who had a flexible sigmoidoscopy with dense polyps in the rectum, and we identified a very small carcinoma, fortunately, because it might have been thought to be one that developed in this mechanism.

DR. RAGHAVAN: Dr. Avigan?

DR. AVIGAN: You alluded at the beginning of the presentation to the rationale that there's a stochastic relationship between polyp numbers and colon cancer, and you showed evidence with a number of NSAID studies to show that there's a kind of reduction that's complementary.

Are those studies not in patients with sporadic polyps and sporadic cancer? And are there any studies like that in this particular disease?

What I'm getting at, doesn't this disease in fact represent a different kind of stochastic problem which is not population based, but rather where you have multiple polyps in specific individuals where the concept of risk reduction by polyp reduction might be

different?

The follow-up question has to do with the two groups that were compared, the placebo group and the treatment group. There was a substantial age difference of almost 10 years. Do you think that that age difference might have played into the result?

DR. NEEDLEMAN: Let me see if I really captured your first question. Were you asking in the epidemiology with NSAIDs is there a subset that was in the SAP-specific group? Gary Kelloff, do you know the answer to that in those 24 trials?

DR. KELLOFF: The wealth of the data on the observational epidemiology is in sporadics. That's true. The data in FAP comes from the small intervention studies primarily.

DR. NEEDLEMAN: I harken back to the target, though. There is the clear evidence of a COX-2 over-expression early and throughout, coupled to epidemiology, and the whole point about a mechanism target agent is we could really test, in the ongoing SAP trial, the consistency with our level to inhibit COX-2 and to go after the histological marker that we have now

of the lesion.

DR. AVIGAN: But just reading between the lines, the problem or the concept here is a modest reduction of polyp numbers in individuals who have many polyps. So, in a way, the concept or the consideration of such a partial suppressive effect might be a little different given that there is more than one polyp to consider the issue of colon cancer risk, and that was what I was getting at.

DR. NEEDLEMAN: I certainly accept that the population of the polyps could be heterogeneous.

Actually, we're struck with the fact that the polyps regress and some disappear. The implication is that's a dynamic, pathological event that can have an intervention. Where we're able to do it in the MIN mouse, we could see reversal, and in fact in the patients. We don't know. There is no reason to think that COX-2's inhibition will stop with time. So, here you have a safe agent and you could chronically look. Indeed, there's a reversal of an early marker.

Now, the discussion about the age. Is there more to add than what we had before? Gary Gordon.

DR. GORDON: I believe it's slide 33.

So, the question that was asked is there was a difference in age between the groups and did that make a difference in response. As Dr. Hawk pointed out in the beginning, A, we did a statistical approach to look at this question and didn't see a difference. The age range of individuals involved in the study per group was not different.

Furthermore, when one looked at the underlying number of polyps in the group or the size of the polyps in those individuals, it was not different.

Lastly, if one takes an age cut in these different groups of age 35, one can see that the distribution of responders and nonresponders doesn't substantially differ by age.

DR. RAGHAVAN: Dr. Bertagnolli?

That was a two-part question.

DR. BERTAGNOLLI: I think the part of the question that's very, very interesting is what's the biology of FAP and how do the genetic events that occur in polyp progression to cancer in FAP relate to sporadic disease. At a very basic level, we have some data from

the sporadic population saying that an adenoma that is less than 5 millimeters has about a 1 in 100 chance of being cancerous or even less, and an adenoma that's 1 centimeter perhaps has about a 10 percent incidence of cancer dwelling in it.

Overlying that is now a little better understanding that mismatch repair mutations produce an accelerated course of the carcinogenesis pathway, and that's something that you would have great concern about missing or hiding the progression of a polyp to a tumor.

What we know about FAP that I think is interesting is that the progression along the adenoma-carcinoma sequence seems to be much less than that that we see in HNPCC. It makes sense because you don't have mismatch repair complicating things and allowing you to pile up mutations -- and that it is less than sporadic disease, which also makes sense because sporadic disease is probably contaminated by those people around carrying undiagnosed defects in mismatch repair. So, I think all of that we know about the natural history of FAP is that it seems to follow the adenoma-carcinoma sequence, and that each one of those polyps is an individual event

that follows that pathway.

Did that answer maybe some of what you were getting at?

DR. AVIGAN: To some extent, the issue really has to do with the notion of numbers not so much genetic pathways. A linear relationship is easy to develop on a one-to-one basis, but if you have many polyps and you have a partial suppressive effect, then the question becomes what, in fact, would you predict the risk reduction to be given the number phenomenon.

DR. BERTAGNOLLI: Right, and I think that has been done because if you look at the risk of -- and you have the tiny, tiny polyps and from what certainly imperfect studies have been done and that the risk of cancer in a very, very small polyp is 1 in 1,000. If you have 1,000 of those, they're independent events and you might think with 1,000 polyps you get one cancer. The only data I know that exists is based on size and that backs it out.

DR. RAGHAVAN: Dr. Blayney.

DR. BLAYNEY: Elegant presentation and proof of principle.

I have three questions. One, do you think the effect of Celebrex is due to a systemic effect? In other words, does it require absorption or just bathing the intestinal mucosa?

Secondly, in your trial that you presented, how did you monitor compliance with the medicines that the study subjects took?

And thirdly, if you think that your drug has a mechanistic target of inhibiting COX-2, in your follow-up study why not compare it against an effective dose of Celebrex, a COX-2 inhibitor, and another COX-2 inhibitor rather than a relatively ineffective dose, 100 milligrams, of Celebrex?

DR. NEEDLEMAN: Let's talk about what we can about the mechanistic component and was it systemic. There is a wealth now of studies with human tissues in in vitro and in samples with animals -- and while the clear observation is there, it is certainly clear that tumor tissue has lost apoptosis. They've lost programmed cell death. That could be put back in cycle by COX-2 inhibitors and reconstituted by the addition of prostaglandin E, and there are in vivo markers of that.

The second thing is that it is an impressive inhibitor of angiogenesis. Now, while you can't really say what is the mechanism, it could have both a direct effect on the tumor and an effect on supplying the blood supply to a tumor for its progression. By the way, the anti-angiogenic effects have clearly been shown to be due only to COX-2 inhibition and not COX-1 inhibition.

So, the circumstantial evidence is there. Mechanistically if you take those human tumors, put them in nude mice, those that have human COX-2 develop the tumor and that's suppressible by a COX-2 inhibitor. So, there's that kind of evidence.

The other thing about Celebrex, it really penetrates all barriers and has excellent distribution. We didn't show -- for example, there was a worry that absorption may be a problem in people missing some segments of their colon. We could present data -- but I won't call someone up to do it -- that the pharmacokinetics, area under the curve, and the drug levels are all consistent within the populations.

The answer about compliance monitoring. Is there something to add, Dr. Gordon?

DR. GORDON: I believe the question was what measures did we use to examine patient compliance during the course of the study.

DR. BLAYNEY: And what are the results of that?

DR. GORDON: So, we used two methods to examine compliance. One was patient diaries. The second was pill count. What we know is that 90 percent of the individuals had at least 80 percent compliance during the course of the study.

DR. NEEDLEMAN: As to what the comparator should be, I think we have to design trials for agents that are approved and can't use agents that are not approved as the comparator. Here we either work against placebo or find a dose-response curve. I think ending up with equivalent COX-2 inhibition doesn't particularly give you an insight if, indeed, your data show it's the level of inhibition of the enzyme and the metabolites and its selectivity is what you demonstrated.

DR. BLAYNEY: I think you're being inconsistent there.

DR. NEEDLEMAN: Try me again.

DR. BLAYNEY: If you're postulating that inhibition of COX-2 is the final common pathway for prevention of these endpoints --

DR. NEEDLEMAN: Right.

DR. BLAYNEY: -- why not use any COX-2 inhibitor as a --

DR. NEEDLEMAN: Besides Celebrex.

DR. BLAYNEY: Besides Celebrex.

DR. NEEDLEMAN: Well, I don't know anything about their biology in tumor tissues or anything else. We have other COX-2 inhibitors in phase III, and that's a consideration.

I think the question is, what is the long-term efficacy in the chronic progression for the patients going to cancer? We know in the FAP follow-up trial that enrollment will be a heroic effort. Indeed, adding arms to the trial would predict that it would be a long time before you could get a decisive answer, that is, do you have a safe COX-2 inhibitor that indeed is changing the progression of the disease.

DR. RAGHAVAN: Dr. Simon.

DR. SIMON: I had three questions.

One, could you clarify your graphic on page 75 for me? It's labeled Percent Change in Area of Duodenal Plaque-like Polyps.

Then there are 12 patients in the placebo group in that graph, 21 in the 100 milligram group, and 17 in the 400 milligram group. What do those numbers, 12, 21, and 17 represent?

DR. NEEDLEMAN: I might give my colleagues a chance to find the page and look at the patient numbers. Do you have the briefing books also? Can you tell us the page number?

DR. SIMON: 75.

DR. NEEDLEMAN: Page 75?

DR. RAGHAVAN: It's your slide 75.

DR. NEEDLEMAN: Oh, I'm sorry, of the slide set.

DR. HAWK: Slide 327 please.

What these numbers refer to are the number of patients that were assessed for duodenal disease. Recall that in the primary presentation I pointed out we took all patients with colorectal disease that had duodenal disease because this was a regression endpoint.

So, it was only a subset of patients with colorectal phenotype that expressed the phenotype in the duodenum as well. Then we had the 6 additional patients that had duodenal disease only. So, the numbers are meant to reflect those patients that had either duodenal disease at baseline or end of study.

DR. SIMON: Well, then my question is there were, I think, 15 or 17 placebo patients. Right? And you're saying 12 of them had duodenal disease. There were 32 or 34 patients in the 400 milligram group, and you're saying, however, only half of them had duodenal disease. Doesn't that represent an important maldistribution of those groups?

DR. HAWK: Well, we didn't randomize on the basis of duodenal disease. We randomized on the basis of colorectal disease.

DR. SIMON: Could you tell me the details of how you did the randomization?

DR. HAWK: I can't.

DR. NEEDLEMAN: Dr. David Jordan.

DR. SIMON: Because you have imbalances in a number of baseline characteristics, age and now duodenal

disease.

DR. NEEDLEMAN: And Dr. Robin Phillips as well.

DR. JORDAN: David Jordan with Searle.

The randomization was done within each center, so that for the St. Mark's for example, for the patients who had colorectal disease, it was 1 placebo, 2 on each of the active arms.

DR. SIMON: But the details of how, logistically. Were these sealed envelopes? Were these done in the pharmacy?

DR. JORDAN: Yes. The packaging of the supplies. I'm going to have to turn that to another person, someone from St. Mark's or M.D. Anderson for packaging.

DR. NEEDLEMAN: Dr. Wallace?

DR. WALLACE: The central code I think was held at Searle. We had absolutely no idea who was on drug or placebo, as I said. So, the patients were randomized, sort of chosen, came 1, 2, 3, 4, 5, as they arrived at the hospital and took their numbers to the pharmacy that held the boxes that were randomized by

Searle. So, no one at St. Mark's or M.D. Anderson had any idea what was in the boxes.

DR. SIMON: What was in the boxes? Sealed envelopes?

(Laughter.)

DR. WALLACE: Yes, it was sealed. We took a case report form, so the case report form came off the shelf with a number, and that number was taken down to the pharmacist, and then the pharmacist gave the drug out. So, at no point was the patient linked to the drug. We didn't know who was getting what.

DR. SIMON: And the randomization lists themselves were prepared --

DR. WALLACE: Held elsewhere. We didn't have access.

DR. SIMON: They were prepared centrally or --

DR. WALLACE: At Searle, yes.

DR. SIMON: At Searle.

DR. WALLACE: Yes. So, we had no access to that at all.

DR. SIMON: One other question, a procedural

question.

Well, did you look at whether duodenal disease influenced what happened in the colon in terms of percent reduction in number of polyps, since there seems to be a substantial imbalance?

DR. PHILLIPS: We've previously done a lot of genotype/phenotype correlation in polyposis patients in this trial and in other trials. There is no correlation between duodenal disease and colonic disease in all of our previous studies that we have done. For example, the severe 1309 mutation in the colon does not give a severe phenotype in the duodenum. We don't understand what causes the severity of disease in the duodenum, but there are abnormalities in the bile in polyposis patients which may be responsible. That is a separate issue.

DR. SIMON: But you didn't look at it in this clinical trial. Is that right?

DR. PHILLIPS: I'm not aware that we looked at it separately because we had previously looked at it --

DR. SIMON: Well, I'm saying there's a

substantial imbalance among the treatment groups in the presence of it. So, you would think that would be reason enough to look at it.

My third question is, also in terms of the procedures of the trial, when was the blind broken and was any interim analysis done?

DR. NEEDLEMAN: There was no interim analysis.

When was the blind broken? David?

DR. JORDAN: The blind was broken after the database was closed and after agreement upon the final statistical analysis plan with the FDA.

DR. RAGHAVAN: Dr. Temple, final question.

DR. TEMPLE: Wow.

Just first an observation and then a question. I think as Dr. Nerenstone was suggesting, it seems very unlikely that you're going to be able to randomize people to 100 after these data are available. I'm skeptical of your ability to do it even though I understand the desire to be able to show a difference between treatments, heaven knows.

The other question I have is you and a number

of your consultants have made the point that the goal is to achieve better management without doing harm. But your proposed follow-up study doesn't even mention the occurrence of local or disseminated cancer as an endpoint, and you obviously haven't calculated sample sizes to be able to see whether you're making an adverse difference on that outcome. Now, I have no idea what the likelihood of local or disseminated cancer is at the time people carry out these procedures, but presumably you and your consultants do.

Don't you need to take that into account in the follow-up study? That is the point, after all.

DR. NEEDLEMAN: Dr. Phillips?

DR. PHILLIPS: The problem is that death from cancer is now not that early an event in patients with polyposis. If you treat a young polyposis patient at the age of 15, we have extended their life by 30 years simply by doing colectomy and ileorectal anastomosis. So, if their age of death untreated is 38, their age of death now is 68. Therefore, death is 10 years short of that of the general population. So, you're going to have to do a lifetime study if you're going to enter

patients at 15 and put them on celecoxib in order to be able to show death differences at that sort of time frame.

DR. TEMPLE: Well, survival isn't the only possible endpoint. Local cancer, even if it's not disseminated, might be. But in some way, it seems to me you need to evaluate what the consequence of delay is. If it were obvious that delay is good per se and that nobody ever gets in trouble, that would make the whole situation very easy, but that is the question, after all. Right?

I mean, your follow-up study would be to see whether you can intervene later, and that would be determined by the observations you make on endoscopy and otherwise. The big question is, does that cost you anything? If it were free, everybody would think that's great. But how do you make sure that you haven't done harm? I thought Dr. Bertagnolli was addressing that in part by saying she certainly wouldn't stop the observation at all. Yet, if you delay colectomy by 3 years, or whatever it is, how do you figure out whether that's a good thing or a bad thing?

DR. BERTAGNOLLI: Two ways to answer your question. I think, first of all, I've said before this is a surgical disease and the way that it should be treated is with removal of all the target tissue possible, and for all the reasons that I don't need to go into that everyone has described today, surgery is inadequate. I think that because it is an aggressively treated disease surgically, we have dropped the mortality from colorectal cancer in these patients to a very low level.

The people who die from colorectal cancer with familial polyposis are generally people who have had an ileorectal anastomosis and somehow slipped through their surveillance and came up with an advanced cancer, the very, very, very rare individual who comes up with an advanced cancer during surveillance, or the approximately 10 to 15 percent of people who present with a new mutation in a primary diagnosis, and 70 percent of those present with their original cancer.

We would all love to be able to design a study that would really use cancer as an endpoint because it would be proof in principle for a lot of

things that we're going forward with, clinical trials both for sporadic and for familial polyposis. But I think reasonably that is why we've defined clinical benefit in the narrow definition that we've given you today because I think to do anything else because we have dropped the rate of those adverse events so strongly with surgery, we just can't do it.

DR. TEMPLE: But you're planning to introduce a change in your surgery. That's the whole point, isn't it?

DR. PHILLIPS: Could I have slide 140 please?

There are other ways that we can look at this. What you're asking really is the reduction of surgical need, whether this drug might allow us to reduce surgical need, whether there is a follow-on study that potentially could address this question. We have, to some extent, tried to take you down the arm of delay in phenotype, and you have questioned that because, of course, the other arm is whether we can devise a study that would lead to acceptance of reduction in the extent of surgery that would allow primary rectal preservation or secondary rectal preservation. I can walk you down

this one and then walk you down that one.

In terms of delay in phenotype, you have the genotype positive/phenotype negative individual randomized either to celecoxib or placebo or to celecoxib low or high, or 200 and 400, however it gets worked out, and looking at time to phenotype expression.

With age 12, annual rectoscopy, dye spray and colonoscopy at age 20, a biopsy proven single adenoma in the rectum would be the endpoint to say that the phenotype in the rectum is expressed. That gives you something measurable.

But I think what you're really getting at is what has evolved in the United Kingdom in recent times, which are the pragmatic trials based on physician uncertainty where we're not trying to change in any way an individual physician's practice. We're saying at a certain part in any of your practice, a physician becomes uncertain about what to do. When that physician becomes uncertain, they toss a coin, and it's at that point that you randomize the patient. You may be aware of very successful United Kingdom trials, AXIS and QASAR, which are trials of adjuvant chemotherapy in

colon cancer, that use this technique.

So, you can have people with a phenotype expressed who can enter either because they're genopositive who've now become phenotype positive or new patients who are phenotype positive. They come to the treating physician who is simply asked a question, is maximal large bowel therapy indicated in this patient? Different physicians will have a different view of this.

For example, there is one polyp in the rectum. I happen to know that the Cleveland, Toronto, St. Mark's would say no, and you would take the patient down the ileorectal anastomosis arm and they would be randomized between celecoxib, high dose or low dose, or placebo and celecoxib, whichever way you want to call it, and the endpoint is whether the use of this drug actually is going to preserve rectums.

I happen to know that at Mayo even one polyp in the rectum, they would say yes to that question, and that there are others, for example, at Hospital St. Antoine in Paris, where they would be uncertain on the basis of one polyp in the rectum.

I as a treating clinician have an uncertainty as the polyposis number increases in the rectum. Each of us has a level of uncertainty.

If you decide that maximal large bowel therapy is indicated, you then ask the question of the physician, are you prepared to randomize as I'm going to indicate below? If the physician says yes, this is a physician who is at state of uncertainty about whether the morbidity and mortality and poorer function of a pouch is worth it in this patient or whether they should have an ileorectal anastomosis. They are uncertain. That is an individual physician's uncertainty.

At that stage, they are randomized to pouch or IRA with high dose celecoxib, which would be the maximal treatment. The pouch patients would be randomized to placebo or celecoxib, going on to an evaluation of pouch and duodenal adenomas in the longer term.

The Mayo might well be prepared to put in patients with one or two rectal polyps into that. Patients who might have 30 or 40 rectal polyps would be our level of uncertainty between doing a pouch or doing

an ileorectal anastomosis.

But the 1309 mutation -- none of us would be prepared to randomize like this. We would say, no, we're not prepared to randomize that mutation because the density of rectal polyposis is too high. And in those circumstances, they would have a pouch, but they would be randomized after the pouch into having placebo or celecoxib, again looking at pouch and duodenal adenoma rates.

This is an entirely inclusive study. It is the new generation in the United Kingdom of pragmatic studies with large numbers with simple endpoints. That would be the alternative way that you could look at this.

DR. RAGHAVAN: Well, I think that has been a very detailed and thorough discussion of a complex area.

The FDA presentation will begin in 30 minutes at 1:30.

(Whereupon, at 1:05 p.m., the committee was recessed, to reconvene at 1:30 p.m., this same day.)

AFTERNOON SESSION

(1:35 p.m.)

DR. RAGHAVAN: I think we'll reconvene the session, and I think Dr. Chiao is going to be presenting for the FDA. I'd ask you to take your seats please.

DR. CHIAO: Well, good afternoon, ladies and gentlemen, members of the ODAC committee. Thank you for giving me the opportunity to present the FDA review of the Celebrex supplemental NDA.

Celebrex was approved by FDA on December 31, 1988 for symptomatic treatment of adult patients with osteoarthritis and rheumatoid arthritis. The indication for this supplement is listed on this slide, that is, the reduction and regression of adenomatous colorectal

polyps in patients with familial adenomatous polyposis, also known as FAP.

NDA review is a team effort. This slide lists the members of the FDA review team. In particular, we wish to acknowledge the input of our GI consultants, Dr. Lewis from Georgetown University Medical Center, Dr. Mark Avigan and Dr. John Senior from the FDA Division of GI and Coagulation Products.

This slide shows the outline of my presentation today. Since this supplemental NDA is being considered for accelerated marketing approval, we will first go over the regulatory requirements for this type approval.

The federal regulation says accelerated marketing approval applies to certain new drug products that have been studied for the safety and effectiveness in treating serious or life-threatening illness and that will provide meaningful therapeutic benefit to patients over existing treatment. For example, the new drug product is able to treat patients who are unresponsive to or intolerant of available therapy or the new drug product represents an improvement over available

therapy. Approval requires adequate and well-controlled clinical trials establishing that the drug product has an effect on a surrogate endpoint that is reasonably likely to predict clinical benefits.

Accelerated approval is subject to the requirement that the applicant study the drug further to verify and demonstrate its clinical benefit. These so-called phase IV commitment studies should be carried out with due diligence.

The proposed indication is only being considered for patients who are suffering from FAP. The next couple slides summarize our understanding of this disease.

Familial adenomatous polyposis is an autosomal dominant genetic disease characterized by the presence of a germline Apc mutation. A variant of FAP is the attenuated form of FAP. Apc mutations in these patients are found at the 3 prime and 5 prime end of the Apc gene.

The attenuated FAP is clinically different from the classic FAP in that it is associated with the occurrence of less than 100 colonic adenomas and a later

onset of colorectal cancer, that is, later than the age of 40. Although the risk of colon cancer in these patients is still greatly increased, the management of these patients is different from the classic FAP, specifically prophylactic colectomy is not recommended for all gene carriers at the time of their diagnoses. If the patient has only a few adenomas, these adenomas may be removed endoscopically.

The hallmark of FAP colorectal polyposis is the presence of greater than 100 colorectal polyps. It is well known that 100 percent of these patients will develop colon cancer unless the colon is removed. I think that you also saw this from Dr. Levin's slide that 83 percent of patients with the intact colon will develop colon cancer by age 45 and 93 percent of these develop colon cancer by the age of 50.

In addition to the colorectal polyposis, there are other pleotrophic manifestations of the genetic defect, especially upper GI cancers and desmoid tumors. In a pooled data set on 1,255 FAP patients, there were 57 cases of upper GI cancers. That's 4.5 percent of the population. Strikingly, 36 of these 57

tumors, that is, 63 percent of these upper GI tumors, were periampullary cancers. The risk of periampullary cancers in FAP patients is approximately 100 times greater than that of the general population.

What are the current management strategies for FAP patients? It is important to note that prophylactic colectomy could prevent colon cancer in these patients and therefore is recommended for all FAP patients.

Two types of surgeries are performed.

Subtotal colectomy with ileorectal anastomosis removes the colon but leaves the rectal stump for functional purposes. Since the diseased rectal mucosa is not removed, these patients need vigilant follow-up. In the literature, 13 to 25 percent of these patients developed rectal cancer at about 20 years after initial subtotal colectomy. Some patients will develop rectal polyposis which is difficult to control or difficult to monitor. Repeated polypectomies may cause scarring. Overall, 25 to 30 percent of these patients will need to have the rectal stump removed eventually.

Another type of surgery is colectomy with ileal-pouch-anal anastomosis. The advantage of this type of surgery is that it removes all colorectal mucosa. However, this procedure is considered by some investigators or clinicians to be functionally less desirable. In addition, polyps have been reported to develop in the pouch. The malignant potential of these pouch polyps are not yet known due to the short follow-up of these patients.

It is important to remember that rectal cancer is not the only problem that FAP patients face. These patients often develop polyps in the upper GI tract. There has not been a consensus on the most appropriate surgical management of upper GI polyps. Furthermore, it has been reported in the literature that there is a high false negative rate in detecting upper GI cancers by biopsy in these patients, about 25 percent, which makes early diagnosis of cancers difficult. These patients' risk of dying of upper GI cancer is higher than the risk of dying from rectal cancer.

Study 001 is a double-blind, placebo-

controlled, three-arm study and is the only study submitted in the supplemental NDA. The primary efficacy endpoint is the mean percent change in colorectal polyp counts in the tattoo and marked areas. The secondary efficacy endpoint is the mean percent change in duodenal plaques in two focal areas. The entire treatment duration is 6 months; a total of 83 patients enrolled on the study using a randomization scheme of 1 to 2 and 2.

The next two slides address the differences in patient characteristics across the three treatment groups. You have already heard from this morning's presentation that the placebo patients tend to be older, about 10 years or so, and this gives you a breakdown on actually how many patients in each different category of age. As you can see, when you compare the patients in the placebo group to the patients in the 400 milligram group, 41 percent of the placebo group are over the age of 40, and 47 percent of the 400 milligram group are younger than age 30.

In addition to age differences, there are also differences in time from subtotal colectomy across three groups. Time from subtotal colectomy was

calculated using the date of the surgery and the randomization date. All placebo patients are 10 years or more out from the initial surgery. In contrast, only 71 percent of patients in the 100 milligram group and 44 percent of patients in the 400 milligram group are 10 years out from their surgery.

Now, why do we think this may be important? This is based on a case series from Cleveland Clinic. They looked at 88 patients who had subtotal colectomy and found that 26 percent had partial polyp reductions and some other partial polyp reductions after the initial subtotal colectomy. The reduction of polyposis after the surgery tended to occur more in patients who are young -- this means less than the age of 30 -- and also may be less further out from the subtotal colectomy.

We did covariate analyses using the ANOVA model, the age of the patients and the time from subtotal colectomy. We used this covariate in the model. They were used as a continuous variable. It didn't really impact on the primary efficacy point. We did not use the breakdown. That means we did not use

the age cutoff as a binary whether older or younger, they're 30 or 5 years or more out from the initial surgery.

This slide shows that the study population is heterogeneous. Most of the patients did have subtotal colectomy. However, you can see from the slides the significant patients. About 25 out of 83 patients had intact colon. All these 25 patients were from one site. That's the M.D. Anderson Cancer Center.

Also, the distribution of the intact colon tends to be a little bit more on the 400 milligram when compared to the 100 milligram and the placebo group.

Also, there are 13 patients with the attenuated phenotype of FAP. That's usually called attenuated FAP, which I alluded to early in my slides. They're clinically different than a classic FAP regarding the occurrence of less polyps and later onset of cancers.

We did put these covariates into the ANOVA model, and none of them has an impact on the efficacy results of the primary endpoint.

This slide lists the efficacy results of

study 001 based on the applicant's data set. As you can see, the mean percent change of colorectal polyp count is minus 28 percent in the 400 milligram group when compared to the placebo group with p equal 0.003.

The change in the duodenal plaques in the 400 milligram group is not statistically significantly different from the changes in the placebo group. In the 100 milligram group, there is an increase in the mean percent change in the duodenal plaques. This is primarily due to 2 patients who did not have any duodenal plaques at baseline, but developed some plaques at the end of 6 months.

We did polyp counts on still photographs with the help of our GI consultant, Dr. James Lewis. We reviewed and verified the methods of counting with Dr. Marina Wallace and Dr. Steinbach from the company. We counted polyps in 28 out of 40 patients at St. Mark's. We are not blinded to the timing of the photographs; that is, that we know which photographs are from baseline and final. However, we are blinded to the treatment assignment.

As you can see on the next slides, the FDA

efficacy results in these 28 patients are very similar to the applicant's.

The next two slides address safety issues of study 001. Celebrex appears to be well-tolerated for a treatment duration of 6 months. Most side effects are gastrointestinal and moderate in intensity. More patients on Celebrex either 100 or 400 milligrams appeared to have grade 2 diarrhea.

In terms of grade 3 toxicities, some of them are unlikely to be related to the drug. For example, the grade 3 in the placebo patients, a lymphoma, and the other two grade 3 in the 100 milligram. One is a suicide. The other is angiofibroma. However, there are two grade 3 toxicities in the 400 milligrams. One is allergic reaction, which you already saw from this morning's presentation. Also, there was incisional pain. The other grade 3 toxicity in the 100 milligrams is diarrhea and abdominal pain.

This slide lists the current Celebrex exposure in arthritis patients. Most of these patients were treated with a lower dose of Celebrex, that is, 100 milligrams or 200 milligrams once or twice daily. There

is no safety data in patients who have received 400 milligrams of Celebrex b.i.d. beyond 6 months at the present time.

We performed a number of exploratory analyses to answer two questions. These exploratory analyses are not prespecified and predetermined in the protocol. They're primarily performed just to answer two questions.

The first question is, what proportion of patients had at least a 25 percent decrease or 25 percent increase in colorectal polyp counts in focal areas?

As you can see on this busy slide, there are more patients on the 100 milligram group or 400 milligram group who have a 25 percent or more decrease in colorectal polyp count at focal areas. However, to our surprise, there are very few patients who had a 25 percent or more increase in the percent of polyp counts. This could mean that polyps at focal areas did not multiply rapidly in a significant number of patients during the treatment period of 6 months.

The second question we asked is whether

changes in polyp counts in one focal area in the rectum predict changes in the rest of the rectum. 74 patients had rectal videotapes at baseline and 6-month follow-up for global quality assessment of the rectum. The committee reviewed the videotapes without knowing the treatment group assignment and the timing of the examination. There are five members in the committee. Four out of these five are investigators of study 001.

The next few slides look at the rectal video assessment by the committee. As you can see on this slide, three members out of five members of the committee agreed in their ratings in 72 out of 74 patients. If you want four members' agreement, the number drops down to 52 patients.

I should point out that this analysis is different from what you saw this morning from the sponsor presentation because the sponsor, I believe, added all the numbers together and divided by 5 and came out with the mean scores. This is slightly different. That just looks at how many members agreed and what the agreed rating is.

This is a quite busy slide, but it lists the

number of patients with a specific type of rating by different degree of consensus. For example, if you look at the first row, which is here, that's telling us that if you go by four-reviewer consensus, there are 5 out of 30 patients in the 100 milligram group or 8 out of 29 in the 400 milligram group were rated to have an overall improvement of their rectal status.

Now, if you go by the consensus by four reviewers, the number dropped down in the 100 milligram to 2 people and from 8 to 6 at a 400 milligram group.

This is another way to look at it. The numbers here are percentages. These are not number of patients. So, in other words, what the percentage of patients actually was rated as better by a four-member consensus, and the numbers are 6.7 percent of the 100 milligram group and 20.7 percent of the 400 milligram group were rated as better. Again, you can see that most of the patients were rated as no change.

Well, this slide actually tried to look at how many of these patients by the four-member consensus were better actually had the predefined. This is again a post hoc analysis. It's not prespecified in the

protocol. The certain degree of decrease in polyp counts in a focal area.

So, this is a 25 percent or more decrease in polyp counts in one tattoo area in the rectum. You can see across from all three treatment groups there are 22 of these patients. And again minus 24 to positive 24 percent change in rectal polyp count in one area, we have more, and that's 38 patients. How many actually had 25 percent more increase in their rectal polyp count, there are only 7 patients. Now, out of those 22 patients, only 20.7 percent, 6 patients, were rated as having a global improvement of the rectal appearance by four members of the consensus. 22.7 percent were rated as same. 13.6 percent were rated as worse, and 36.4 percent the committee did not reach consensus.

So, this is basically sort of more analogous to a positive predict -- kind of value in a diagnostic test that among patients who had a greater than 25 percent decrease in rectal polyps in one area, 27 of those patients by this study had an overall improvement of the entire rectum by video assessment.

So, what are conclusions after review of the

study 001? The conclusions are the study enrolled a heterogeneous patient population, by which I mean they included patients with subtotal colectomy, patients with intact colon, patients with attenuated FAP.

The mean percent change in the colorectal polyp count is minus 28 percent in the 400 milligram group with p equals 0.003 when compared to the placebo group. This is supported by more patients in the 400 milligram had a greater or equal to 25 percent decrease in colorectal polyp count in focal areas when compared to the placebo group. It's also supported by more patients in the 400 milligram group had a better rating of rectal video by four committee members.

Celebrex at 400 milligram b.i.d. was well-tolerated for 6 months, but safety data for this dose beyond 6 months is not available at the present time.

Percent change in rectal polyps in one area does not appear to predict for changes in the entire rectum when the entire rectum is assessed by videotape by four viewers.

The durability of Celebrex effects on colorectal polyps cannot be assessed due to the short

treatment duration of 6 months.

Now, what is the real unresolved issue here? We think that this is an important question. Is reduction in polyps in FAP patients a surrogate likely to predict clinical benefit in these patients?

Some of this has already been discussed in the morning session, and I think we all would agree that reduction in rectal cancer, reduction in duodenal cancer, or reduction in any other FAP-related cancers would be a real benefit for these patients. Preservation of rectal stump without increasing the risk for rectal cancer, delay of prophylactic colectomy without increasing the risk for colorectal cancer are other benefits as well.

We have three comments on the study for your consideration. It's our thought that without a complete regression of all colorectal polyps, reduction in polyps by itself may not result in a decrease in colorectal cancer incidence in FAP patients. And the reason for saying that is that we know from the biology of FAP, the entire GI mucosa in these patients is at risk for developing cancer due to the germline Apc mutation.

Cancer may arise from the remaining polyps or even non-polypoid areas.

We find that the clinical significance of a partial reduction in colorectal polyps in FAP patients is difficult to assess from study 001. There's only one follow-up endoscopy that was done at 6 months. We do not know if these patients continue to receive Celebrex, whether the Celebrex effect on rectal polyps is going to be greater, less, or not. We don't really have the answer to that.

The final comment that we have is if the ODAC committee recommends accelerated approval, Celebrex treatment should be considered only as an adjunct to the usual care of FAP patients.

This concludes my presentation. I'm happy to answer any questions you may have.

DR. RAGHAVAN: Thanks, Dr. Chiao.

Questions from the committee? Dr. Jacoby.

DR. JACOBY: I wonder if you had considered an additional possible benefit. The polypectomies that are done for the larger polyps before the patient undergoes colectomy each have a risk. If you're

reducing the number of polyps, you're going to be reducing the risk of the polypectomy, which would include perforation, bleeding, infection. It seems obvious to me that if the polyp number goes down, the number of polypectomies required will go down. Wouldn't that be considered a clinical benefit?

DR. CHIAO: I think it would. I think it depends on what type of criteria that the surgeons will use or the gastroenterologists will use to remove the polyps. It's my understanding that some of the gastroenterologists will remove a certain size of polyps and not across the board. I think that probably should be best commented by the investigators.

DR. RAGHAVAN: Would the investigators like to comment?

DR. NEEDLEMAN: Would you restate that?

DR. CHIAO: Well, I think the question was asked whether a decrease in the frequency of polypectomy would be a benefit, and my comment is it would depend on what type of criteria that one would use to remove the polyps. If you say all the polyps ought to be removed, and then I think that's a different issue than you say,

well, only large polyps would need to be removed. If that is the case, we don't have any data at this point to say how many big polyps are there that would need to be removed.

DR. BERTAGNOLLI: The only thing that I have to say is the usual criteria were used, size of polyps and characteristic of polyp. In other words, is it feasible to remove it. If it is so flat that to remove it would cause a complication, we wouldn't obviously. So, if this agent were to lead to a reduction in the size of the polyps, which we seem to see from the data, then it would indeed reduce our usual indication for polypectomy.

DR. RAGHAVAN: Dr. Kelsen?

DR. KELSEN: We heard this morning -- or at least I think I heard this morning -- that the risk of getting colon cancer is now very small if you do a prophylactic colectomy. It's the trigger for doing the colectomy that I think is the issue here, and I think Dr. Phillips commented on this. It sounds to me like there is tremendous divergence of opinion as to when you do that. One polyps or 30 polyps, which is not such a

small difference.

How are you going to handle the analysis of whether or not this delays colectomy with such a phenomenal difference in when you do the colectomy? How are you going to analyze that? It's obviously at the discretion of the operating surgeon or the operating team. And we're going to be giving this drug, if it's approved, for X period of time presumably indefinitely until you reach that trigger point.

DR. RAGHAVAN: Mr. Phillips?

DR. PHILLIPS: I think that is an important question. The issue here is, are we dealing with cancer prevention by surgery or are we waiting for cancer to develop and then treating it?

When you have a patient with polyposis who has expressed the phenotype, then we know that surgery is the treatment, and we wouldn't consider that surgery should be changed.

But we know that in the younger patient, the risk of someone who's had a colonoscopy developing cancer over a period of time from historical data is really very, very low indeed. Therefore, we can

actually determine the timing of their surgery according to social circumstances, schooling, summer vacation, things that won't make them lose time out.

Those people who would consider that we could follow these patients endoscopically, one has to ask the question, what are they waiting for? Obviously the final answer and the final trigger is you're waiting for them to develop cancer and you're changing from cancer prevention to cancer treatment.

Now, we are not recommending any of that at all. We're recommending the use of this agent within the current guidelines for the management of a polyposis patient. Given that you do that, I think there could be significant benefit to these patients.

DR. KELSEN: Well, I appreciate that, but the difficulty I have is the statement that the decision as to when to do that seems to vary so much. We all understand we're not going to wait for cancer to develop. So, I'm struck by the comment -- I don't know who made it -- that at the Mayo Clinic you have one polyp --

DR. PHILLIPS: Yes. I'm sorry. That is the

type of surgery that's chosen rather than the decision to operate.

DR. KELSEN: I got that, but that's a surgical intervention.

So, is there broad agreement on the time to do the subtotal colectomy?

DR. PHILLIPS: Yes. The agreement that we have is when the phenotype is expressed, then you might as well get underway to prevent cancer. You fit it into their schedule.

Where the disagreement lies is in the magnitude of that operation. Because of the disadvantages of a pouch, there are groups, significant and very experienced groups, who feel that rectal preservation is useful.

I would comment on the rectal cancer rates that occur after ileorectal anastomosis. We're looking back in time. In those days, there was no other choice for the patient. These were patients being given ileorectal anastomosis who today would never be given ileorectal anastomosis because we have the opportunity to give them a pouch. So, those are the worst possible

rates.

And that is why a group of surgeons would be prepared with low density rectal polyps to have the evident benefits of ileorectal anastomosis, and if that could be supported by a drug, we would be even more comfortable and would feel that a number of our colleagues who are slightly more aggressive in their surgical management might be prepared to test this.

DR. CHIAO: Could I also ask? I think along with Dr. Kelsen's line, I think it's not totally clear to my mind either, is that if the patient started to express the phenotype and started to have polyps, is there an agreement on how many of the polyps are there that will prompt the surgeons that the colon needs to come out either by ileorectal anastomosis or the pouch procedure?

DR. BERTAGNOLLI: I can address that one. I think that what happens realistically, 50 percent of individuals at age 16 express the phenotype. Once they express the phenotype, we all agree entirely that surgery is indicated. One of the reasons we are proposing the younger group, even though that's

difficult, for the follow-up study is because we won't have confusion as to whether they need surgery or not.

Now, the timing of surgery is very crucial, and there is a variability in when that is done. It is clear to all of us who take care of these patients that ideally we would like to be able to delay, at times, the surgery, mainly to delay it to allow the kids to get through high school, to allow them to reach age of consent so that this major operation is something that they as adults can decide their own course for. If someone presents with a phenotype at age 16, it is often the case that we wait until they're 18 to 20 to do the surgery even though they're expressing the phenotype. So, again, that's why we, in particular, think this is the best population to do the follow-up study in.

DR. RAGHAVAN: Dr. Needleman?

DR. NEEDLEMAN: There were just two points in the review that we could comment on. We do have long-term safety experience, if you want to see it, with the 400 milligrams b.i.d. greater than 6 months.

Similarly, if we use the same criteria of the 25 percent change with the duodenum, while that wasn't

the primary objective, we certainly achieved significance and we could present that data if you would like.

DR. BERTAGNOLLI: I also forgot to say that there certainly is a degree of severity which triggers us to do surgery if the person is 12 years old, and we can all agree that anyone with severe dysplasia should have surgery even if they're at a younger age. Obviously cancer. Those are black and white. There also is a certain degree of polyp burden. We're surveying these kids sometimes every 6 months. If they're increasing.

These things could be in the context of a clinical trial, not that everyone does exactly the same thing, but agreed upon and standardized so that we could have some reasonable assessment when we're finished.

DR. RAGHAVAN: Dr. Margolin.

DR. MARGOLIN: I have several, but maybe we'll have to cycle around. Two are related questions, and I think Dr. Chiao could probably take a shot so we don't have to get any of the sponsors up here.

I'm still puzzled about what was brought up

by the patient letter as well as the way you analyzed the data looking at rectal disease. Do you feel comfortable that rectal cancer or rectal polyps or what happens to rectal polyps in the presence of an intervention is a surrogate for what's happening in the colon?

And the other part of that question is, are we focusing more on the rectum now? Because that is still the area that's the greatest threat in patients who have rectal preserving procedures and the colon is no longer an issue. You can get that out of the way when you have to.

I see the sponsors shaking their heads, so we know your answer.

DR. CHIAO: I can try to take a shot at your first question, and I think the first question is, are the changes in the rectum related to the changes in the colon? Right?

I think you can look at it in two ways. One is you can look at the focal changes in the rectum -- that was a defined method -- and compare it to see is that related to changes in the marked area in the colon

because in the colon, usually you have at least two marked areas using anatomical markers. One is the appendicial orifice. The other is the ileocecal valve, and depending on whether there are other polyps or not, you can have 1 to 2 tattooed areas as well.

I think it's difficult to assess because we only have 25 patients out of 83 had intact colon, and some of these patients do not have rectal polyps at baseline, to my understanding. If we just put everybody together and do a scatter plot and looking for correlation between the percent change of rectal polyps in the focal area and colon polyp change in the marked area, we do not find a correlation.

DR. RAGHAVAN: Dr. Blayney?

DR. BLAYNEY: Dr. Chiao, your presentation was the first one I had heard mentioned the attenuated FAP. I wonder, could the sponsor talk to us if they agree with your assessment that 10 to 15 percent of the patients in the study had attenuated FAP?

DR. CHIAO: Well, there is a database. They're looking at the attenuated FAP. What I did is I pulled out from the database. Out of 83 patients, 13 of

these patients were coded as having attenuated FAP and the distributions of these patients, 8 on the 400 milligram group and I think 2 on the 100 milligram and 3 on the placebo. That's what the numbers I gave in my slide.

DR. BLAYNEY: So, this was from the sponsor's submitted data.

DR. CHIAO: Yes, from the sponsor's data set. I did not have a separate data set on that. That was from the applicant's data set.

DR. BERTAGNOLLI: Attenuated FAP -- certainly there are genotype-phenotype correlations in FAP, and people with mutation cluster regions centering around codon 1309 have the most severe phenotype in general. Within individual families, all of whom have exactly the same genotype, though, there can be a very wide spectrum in the phenotype of the disease. We all have individuals in a family who clearly have the gene who look like an attenuated patient even though their brother may have expressed the classic phenotype.

We believe that there are certainly modifier genes for FAP like there are for every other disease we

know about, and everyone gets genes from their affected parent as well as their non-affected parent. So, in general, these correlations are good, but they don't always hold up. And we don't know about the effect of modifier genes on the effects of these drugs, for instance.

DR. CHIAO: Well, I just wanted to -- I'm sorry. I'll just make my comment. I just wanted to let you know that we did put that into a covariate analysis and that did not impact on the primary efficacy results. I think the reason that we did that is that -- the point I think I wanted to make is that this is a heterogeneous patient population because we had patients with a clearly clinically different variant of FAP or have intact colon or have subtotal colectomy. But we did put these variables into an ANOVA model and it did not affect the primary efficacy result.

DR. BLAYNEY: I'm struck. I've taken care of two patients within the last year as a medical oncologist who are dead or in the process of dying of this illness that slipped through, as the term was used, the surveillance. Those people concern me, and I'm

concerned that we may lower the vigilance of our screening and our surveillance by the availability of such an easily available, low tech mechanism for dealing with these people.

DR. RAGHAVAN: I just have to make a comment from the Chair. One of the things about the way the FDA works and this process is one of fairness to all sponsors. The way the rules of this presentation apply is the sponsors have actually been given the opportunity of commenting. In fact, I gave them extra license in question time by allowing an hour and a half of question time.

In other presentations, the sponsor can be invited to respond, but I'd like to make it clear that it is not an automatic right for the sponsors to respond to questions from the committee. So, I would appreciate it if you don't jump to your feet. I'm sure many of the people helping the sponsors haven't sat through this process. I'm not trying to be doctrinaire, but we operate under a set of rules. And I apologize for that, but the rules are the rules.

Dr. Avigan, you had a question.

DR. AVIGAN: I wanted to just address myself to the plan for the phase IV study, which is that the proposed study will have two arms. Both will be treated with the drug. Basically what will be scored will be surgeon behavior. We just heard before that surgeon behavior, to some extent, is a discretionary phenomenon.

I want to hear again from the sponsor perhaps and perhaps from the speaker. Given that there is no placebo in the proposed study and that we would be using historical controls, given the weakness, is it possible -- I want to hear from the sponsor what their considerations are on that subject.

DR. JOHNSON: I wonder if I might interject a clarification on that, Derek, because actually this is an important element of this deliberation. I don't take care of these patients except when they have metastatic disease. I too am struck by this question. But I would prefer to ask not the sponsors. I want to know what our consultants here have to say about this, number one.

First of all, does it matter if you delay surgery 8 months? I'm not sure it makes any difference, and I don't really care what the sponsors think about

what they think is a relevant endpoint because that's like asking a barber if you need a haircut.

(Laughter.)

DR. JOHNSON: I want to know what our consultants think is a relevant time frame. How long should one delay the surgery? That to me is important, if that's going to be their endpoint.

Secondly, I'd ask the consultants. We heard the sponsor's experts tell us that, oh, they would use Celebrex in a heartbeat. Well, it's approved. I hope they're using it now. But I want to know what were you using before Celebrex was available.

Sulindac has been shown to reduce polyps in this disease, and have you guys been using that, or is that just a drug that no one uses? I just am interested.

DR. RAGHAVAN: Would one of our consultants like to comment?

DR. JACOBY: I'd be happy to make some points on that.

Actually we have been using Sulindac and other drugs in that category for these patients. In my

clinic I would say the majority of patients actually want treatment of that type. They perceive that there are severe problems that are not adequately dealt with by the surgery.

In fact, in some circumstances the surgery can be harmful. The reason that we may want to delay until the age of consent and until the child is more fully developed is because some surgeons believe that the results are better when they operate at an age closer to, let's say, 18.

The other issue that hasn't been brought up but has been a big problem in my clinic is the problem of desmoids. We have no good treatment for desmoids. Some people think that the desmoids may actually be triggered by the surgery. So, a delay in the surgery may delay death due to desmoids.

I see Henry Lynch nodding his head back there. Maybe he'd like to give another --

DR. JOHNSON: Actually that's a very important point that Dr. Santana tried to touch on this morning with the sponsor, and they skirted the issue. Basically what happens to the extra-intestinal

manifestations, if anything? Now, 6 months is a short period of time, but that seems to me to be a very relevant issue to the indication here. Does it, in fact, delay desmoid development and what impact, if any, does it have on desmoids? Those are issues that are important.

But let me ask you again the question that I want to know. If you delayed surgery from age 18 to 20 -- 18, by the way, -- actually 12 in Tennessee is the age of consent. Right?

(Laughter.)

DR. JOHNSON: For marriage I mean.

(Laughter.)

DR. JOHNSON: But seriously. Does delaying it a year make a difference relevant? I could see if one delayed surgery from age 16 to age 65, that might make a difference, or age 16 to age 35, that too might make a difference. Getting one through one's reproductive years, all the issues of genetic counseling aside, that makes sense to me. I'm just asking for the experts to tell me. Dr. Lewis maybe or Dr. Jacoby.

DR. RAGHAVAN: I think Dr. Jacoby's

suggestion of a comment from Henry Lynch, notwithstanding that he's sitting over to my left, he's a renowned expert. It would be a shame to have him come all the way and not maybe respond to Dr. Jacoby's comment. Dr. Lynch, would you like to respond to his question?

DR. LYNCH: Well, I do take the position that delaying surgery could have a potential benefit in the case of desmoids. Particularly this would be apropos in those families where we see aggregates of desmoid tumors and where there are some hot spots in the Apc gene. This goes back to my own experience with patients that have developed desmoids.

In one case -- and I'll make this very brief, it was a 14-year-old boy from a classical FAP family that I had recommended a prophylactic colectomy on at age 14. A couple of years later -- that's pretty close to the average when the cancers occur after surgery. Actually it's about 5 years, but this happened about 3 years -- he began developing a desmoid which absolutely became massive, and he went under all types of therapy. There was no response. I finally put him on adriamycin

and DTIC. He had what we thought was a complete response. Some of you may have seen this because of a paper I published several years ago in the American Journal of Gastroenterology.

Anyway, to make a very long story short, we didn't know when we should stop because the CT-scan did not let us know whether we were dealing with necrotic tissue in that desmoid. There was still a mass there. So, we did a laproscopic evaluation, and within a year he died of desmoids in each of the trocar sites. This was very convincing to us that what is well known is that the surgical effect correlates very strongly.

If I can make just one more comment on the attenuated FAP. Again, there will be many patients that you really will not have to do prophylactic colectomies on. When you look at some of these extended pedigrees where they do have the germline mutation and only have three or four, five, six maybe adenomas. The gastroenterologists in my group at Creighton are able to do polypectomies, and these patients are going on into very long lives. So, I think there could be some benefit to the drug in those situations.

DR. RAGHAVAN: Thank you.

Dr. Margolin, did you have one more?

DR. MARGOLIN: One of them was asked. Well, I'll just ask it again just to confirm. The multivariate analysis that you did, you feel comfortable that all of these major heterogeneities and presumed risk factors that may have affected the outcomes of the primary endpoints have been taken care of by the multivariate analysis?

DR. CHIAO: Well, the covariate analysis looked at the age of the patients, years since subtotal colectomy, intact colon or not, attenuated FAP or not. However, the age and the years from subtotal colectomy were used as a continuous variable in the covariate analysis. We did not do a breakdown. We did not say, we'll cut off 5 years less or more and 30 years or less, because if you look at the Cleveland Clinic series, they used a cutoff of 30. They see more young patients had spontaneous regression of rectal polyposis after the subtotal colectomy. It's seen more in the young.

The reason that we were looking at that is just because there's some literature suggesting that may

have an effect. The surgery itself may cause spontaneous regression, as a matter of fact, in about 50 to 60 percent of the patients. But we didn't do a breakdown with a binary variable. We certainly could go back and relook at it. But using them as a continuous variable of age and years, we do not find any impact on the statistical significance of the primary endpoint.

DR. RAGHAVAN: Dr. Temple?

DR. TEMPLE: I have two questions. The first is to find out whether you described a difference of opinion about a result from the sponsor. They showed a scatter plot relating effect on the overall gut and the effect on the sentinel lesions -- they were all in that upper right quadrant -- and described a significant p value relating those two. You said specifically you didn't think there was a correlation between outcome on the overall gut and the effect on the sentinel lesion. Were you reflecting a redo of that analysis or your own analysis on the four-person consensus?

DR. CHIAO: It's our own analysis. Our analysis is different with the sponsor's because I personally feel it's very difficult for me to interpret

a value mean score of 0.2 or minus 0.2. We're talking about a five-member consensus. So, instead of using a mean score, why look at how many of them agreed and use the agreed rating as the endpoint?

I put up two slides, I think the first showing you the agreed upon rating by three members and four members, and the second was showing by four members in a percentage of patients. My number is derived from the consensus four members.

DR. TEMPLE: No, I understand. So, you're really talking about two different analyses.

DR. CHIAO: A totally different thing. Our analysis is not a correlation. We're just looking at what proportion of patients had a certain magnitude decrease in the rectal polyps only or do they have other colon polyps as well, and how many of them actually had a better overall rating by four members of the committee.

DR. TEMPLE: Right.

The second question is, did you or did -- I guess this might go to the sponsor -- do an analysis of the two centers separately of their results? That's one

way to sometimes gain support for a single study, to show a consistent effect across studies. Did anybody look at that?

DR. CHIAO: Well, we included the treatment centers in the covariate analysis, and I believe that the sponsor did that as well. I do not think it has an impact on the statistical significance of the primary efficacy endpoint.

DR. TEMPLE: No, that's not what I'm asking. I just wondered if you did an analysis of each center separately.

DR. CHIAO: No, I did not try to separate the two centers. We would have had less than 10 patients if I separated them all out.

DR. RAGHAVAN: Would the sponsor like to comment?

DR. GORDON: Can I have the slide that shows the focal responses by center?

This is the graph that you've seen several times today in which we've looked at the outcome by site, looking at the percent change from baseline, showing St. Mark's in white and M.D. Anderson in green.

Again, you can see, for the most part, there's a pretty good distribution of the responses between the two centers.

DR. TEMPLE: I had a simpleminded question and probably Dr. Simon will be irritated I'm even asking it. Did you do a statistical analysis of each center separately?

DR. GORDON: We looked at each center separately, yes. Dr. Jordan will address that.

DR. JORDAN: Yes. The p value for the U.S. center is 006 and for St. Mark's it's 095.

DR. TEMPLE: Thank you.

DR. RAGHAVAN: Dr. Nerenstone? No?

Other questions? Dr. Lewis.

DR. LEWIS: We're talking about post-marketing studies if the drug is approved. I'm wondering why we're hung up on the one study that we've heard you've proposed which is in a totally group that was studied. These would be adolescents.

What about the patients in this study? I would ask the clinicians what are they currently doing with these patients. I understand they were taken off

of this study at the end of 6 months because they're probably in other studies, but if they weren't in other studies -- how are we going to get the information on whether there's additional polyp reduction after 6 months in these individuals and for the patients with the intact rectums, et cetera? How are we going to know how to prescribe this drug and for how long a time in the study population?

DR. RAGHAVAN: Mr. Phillips, do you want to take that?

DR. PHILLIPS: I think you've caught the answer in your own question. The problem is that these are rare patients and we can't keep them in any one study on the off chance that we've got efficacy. We have to move on to new studies until we have an established treatment which is accepted. We would now accept this is an established treatment, and so that's what we want to have is the established treatment.

Up until this time in our own practice, the problem with Sulindac with the rectal suppositories and the rectal taken orally is we've had problems with side effects. So, we've ended up not doing that as a

standard except in patients with severe rectal polyposis, whereas we would like to take this back into people with more mild disease.

I don't know if I've fully answered your question.

DR. LEWIS: Well, I thought it was fairly simple. I mean, this is a 6-month study. We have a 28 percent reduction in the number of polyps, and we're using that as a surrogate endpoint to suggest that this is going to change in some significant way the way we'll have to manage these patients in the future. We may have fewer proctectomies and fewer rectal cancers perhaps in the future. I don't know how we'll know that in this particular group without an additional long-term study.

I guess part of my question is to the FDA and that's as part of an accelerated approval, do we need to stick to the study population to keep studying that population as opposed to try a totally different study population?

DR. TEMPLE: We've certainly accepted studies in other stages of the disease. For example, we've

approved drugs for refractory disease and accepted studies as primary therapy as the, quote, confirmatory study. So, that could be possible.

DR. GORDON: If you would like to see it, we do have a couple of other study designs that, in fact, address the population that was used in this study.

DR. LEWIS: I think it would be helpful to see those.

DR. GORDON: Can we have slide 129?

This is a trial that's focused on individuals who have a need for secondary surgery. So, again, for a discussion point, I'll say this would be a group of individuals who had an ileorectal anastomosis and might go on to lose that because of progression of disease. I'll show you two versions of this.

This is a double-blind, controlled study in these individuals with previous FAP surgery, looking again at two doses of celecoxib. We're estimating that this trial will require 1,958 individuals randomized 1 to 1. The goal would be until the need for subsequent surgery or until 368 events occur.

If we go to slide 130, this shows the

assumptions that underlie that. We're estimating an event rate of approximately 5 percent per year that would have subsequent FAP surgery, again a 10 percent effect of the drug, reducing that down to 4.5 percent at 100 milligrams twice a day, and a 30 percent effect at 400 milligrams, bringing that down to 3.5 percent, a dropout of 10 percent, a power of 80 percent, and a p value of .05.

If we go to the next slide, this would be the same trial now with a placebo arm as opposed to an active celecoxib control arm and, again, celecoxib at 400 milligrams twice a day. This would require 1,194 patients, and we would continue the trial basically until 240 events occurred.

On the following slide are, again, the assumptions that underlie this, the reasoning behind 1,194 individuals.

So, such a trial obviously has the complication of being an extraordinarily large trial for FAP.

DR. RAGHAVAN: Dr. Albain?

DR. ALBAIN: I would just like to ask the

FDA's opinion on the use of a 100 milligram b.i.d. control arm in any of these subsequent trials based on your review of the data.

DR. CHIAO: It doesn't occur to me that this drug is effective by the prespecified primary endpoint or secondary endpoint in the 100 milligram group. I think, as a matter of fact, we saw 123 percent increase of the duodenal plaques in the 100 milligram group, but I pointed out that this is primarily due to patients who did not have duodenal plaques at the baseline, but who developed. But logistically speaking, these 2 patients should be included because they had progression of disease. They should not have developed the plaque if 100 milligrams is effective.

DR. JOHNSON: But even if those 2 patients were taken out, didn't the sponsor's data show --

DR. CHIAO: Yes. It's only minus 47 percent and that's not statistically significant.

DR. RAGHAVAN: Dr. Margolin.

DR. MARGOLIN: My questions are sort of follow-ons to both of Dr. Lewis' questions.

Thank you for showing us your planned trials.

I'm not sure they answer all the questions.

I think part of the problem is how this indication would work in terms of how to select patients and then what to do after they meet certain endpoints. If you believe that this drug works against this very important intracellular pathway and that that's involved in all or most of at least the gastrointestinal pre-malignant manifestations of this phenotype, would you have patients continue to take the drug even after they've had each of the serial surgeries that maybe you have succeeded in delaying, but not 100 percent avoiding?

Okay. That's the answer. Thank you. I'll keep it short. That was a yes for those who didn't see the sponsor shaking their heads.

But the other question I have has to do with sort of the FDA's stance on how post-marketing studies to validate accelerated approvals are designed. I think the pediatric or adolescent trial is interesting and really shifts gears maybe to a more intrinsic and very important question, but it doesn't really necessarily validate the findings from this tantalizing, but

extremely limited study that we just heard about. I think what we need to know is, do they need it lifelong? What are the potential down sides? Some more details about what we can achieve from the approach that was taken today.

DR. RAGHAVAN: Who are you asking?

DR. MARGOLIN: I'd like to hear the FDA on that.

DR. CHIAO: Well, I agree with you that the population is very different because one is the young adolescents who have not had any type of surgery. I think the purpose of using that is to suppress the phenotype, if we could, and to delay the surgeries. That's not, to my mind, equivalent to what we were studying here in the adult population who have already had subtotal colectomies.

That's again what I was pointing out, this heterogeneity of the patient population because if you subtract 25 from 83, we're only left with 53 patients with subtotal colectomy which is, by the way, the majority of the patients on the study. For these patients, the clinical benefit would not be to avoid

colectomy because they already had it. I mean, it would be to prevent subsequent development of FAP-related cancers or to reduce the need for rectal stump removal or decreased polypectomy and that sort of benefit.

So, I think the clinical benefits in the two populations are different, and I'm not sure they can be addressed by one surrogate endpoint.

DR. RAGHAVAN: Any additional comment on that point from the sponsor? Maybe Mr. Phillips would want to address that or oncologists from Brigham? No. Okay?

Other questions?

DR. GORDON: We could do a composite endpoint, if that would help.

DR. RAGHAVAN: Go ahead. I think we're trying to get clarification of a complex area, so I'll give you some license.

DR. GORDON: Could I have slide 135, please?

What we had originally envisioned as a follow-up trial was, as I think had been mentioned earlier, was trying to look at a composite endpoint in which one would try to, in this very complicated patient population, capture a number of events. We sort of

break these down into two major areas, if you will. One would be FAP-related death, any FAP colorectal or duodenal cancer developing, a secondary FAP-related surgery, such as loss of a retained rectal segment.

Other important endpoints that we think that would need to be followed and you've heard somewhat in the discussion would be the development of high grade dysplasia either in the colorectum or the duodenum, and as mentioned by Dr. Bertagnolli and discussed up here a little bit would be the accumulation of a number of large polyps as a potential endpoint to use in a study.

Having this sort of composite endpoint, while I can't give you an exact number, would obviously have some impact on the study design and the study size.

DR. RAGHAVAN: Dr. Temple.

DR. TEMPLE: Well, that sort of answers Dr. Margolin's question because that is the population that was studied. But what's the design of that trial? Are you going to randomize between 100 and 400? And would you care to say why you think, given the available results, people will enter that trial?

DR. GORDON: I think you're asking questions that we tried to address somewhat when we initially laid out the questions and wanted to have this discussion about trial design. We think it is challenging to envision why people would enter a trial with a placebo arm. We have a drug that appears to be very well-tolerated. It's a drug that's available. It's a drug that appears to be safe, and we're convinced that it's a drug that has activity. So, why anybody would enter a placebo-controlled trial is challenging, and I think that's why we're here for discussion is to say this is a very difficult, very challenging area that has some real unmet need to it and how do we best address it.

DR. TEMPLE: Before you leave, why do you think they'd enter one where they were randomized to 100 milligrams? I'll tell you what my concern is. You're making that sound possible, and I wonder whether it really is.

DR. GORDON: I'm not even going to show the dot graph again because I think we've shown that several times and it's shown that there have been individuals who have responded at the lower dose. I think you're

exactly right, though. That's why we need to have this discussion, is to get the guidance on how would you do this trial to get the follow-up information that's needed.

DR. NEEDLEMAN: If we weren't satisfied with the historical control, of course then you would have done the one arm open --

DR. TEMPLE: No. It's a terrible problem, and if you could do the 100 versus 400 and show a difference in those outcomes, everybody would say, hurray, hurray. But that doesn't solve the combined ethical and practical dilemma that people will criticize it for using a less effective therapy. And the more difficult question here is even if you get over that, will patients, properly informed, enter a trial like that, and that's a very hard question.

DR. GORDON: And that really does put us back to using the historical data as a basis of comparison.

DR. TEMPLE: Which raises the question of whether one is forced to study an entirely different population, like sporadic polyps or something like that.

DR. NEEDLEMAN: We, indeed, originally

approached the FDA that the logical trial was to go back into the SAP population as the next way with a properly balanced trial to go forward with its appropriate controls, and it was a desire from the FDA to have some extension of this outcome with this dilemma.

DR. RAGHAVAN: Dr. Pelusi?

DR. PELUSI: I think I have some very basic questions to the experts that we have here. A couple of things come to mind as we do this discussion.

One, it seems like we're having multiple discussions, which are needed, but getting back to the actual study itself, I would like some input from the experts, if you will, in terms of when we're looking at this follow-up study, if you will, and we're looking at this drug in the use of children between 12 and 19, what benefit or from the information that we got from this first study gives us the indication or gives us the ability to make the decision that, number one, it would indeed be helpful based on the fact that we don't know what the reduction in the number of polyps ultimately means and we don't know what the histology is in terms of, as they go through changes or are reduced, if that

makes a change in their overall survival or outcome. If we don't have the data on the use of this drug in younger children, especially with contraception and everything else, does that set us up, if you will, to perhaps do more harm than good?

Then the last part. I agree with Dr. Lewis. I am confused with these 83 people that all of a sudden are now out and we have no way to follow them because you would think these people really do have some valuable data long term.

DR. RAGHAVAN: Well, I've looked at Dr. Surawicz. She's going to comment when she has the formal discussion role.

Dr. Brand, do you have any answer to some of those questions?

DR. BRAND: I think the first question is a very tough question. Clearly you could cure all these patients of their colon cancer by just taking out the colon. I think there's no one here that doubts that.

I think the importance comes in finding these drugs that will address the duodenal cancers. The data does suggest that there's some improvement. It didn't

come to statistical significance when you look at it, and I think it would be nice to go out longer, a year. Maybe they weren't on it long enough to see whether or not there would have been statistical significance.

I think clearly future studies need to address the upper GI tract lesions just as much, maybe even more than these other issues of prevention.

From the standpoint of do these polyps then reflect changes in the duodenum, that I just do not have a good feel for, and I don't know if anyone does in this room.

DR. RAGHAVAN: Do any of the other experts want to comment?

(No response.)

DR. RAGHAVAN: A couple more questions, then I'd like to move to the next component. Ms. Forman, then Dr. Nerenstone.

MS. FORMAN: Are there statistics on the percentage of FAP patients who actually develop cancer?

DR. RAGHAVAN: Do you want to take that? Proportion of FAP patients who develop cancer at certain time points.

DR. CHIAO: Well, it depends on what type of cancer you wanted to know. I think in patients who have had subtotal colectomy -- that means the diseased colon has been taken out -- you're looking at the 13 to 26 percent risk of developing rectal cancer during a long period of time. That's about 20 years.

DR. AVIGAN: The numbers that I remember are that patients who had ileorectal anastomosis, 25 percent of them go on to their second surgery for lesions that are worrisome or tumorigenic. About 4 to 6 percent of them get duodenal cancer. If they don't have their colon out by age 50, they all get colon cancer.

DR. RAGHAVAN: St. Mark's experience?

DR. PHILLIPS: Of 222 patients who had colectomy and ileorectal anastomosis, 9 died of upper gastrointestinal cancer, 11 died of cancer with an unknown primary, probably upper gastrointestinal, and 5 died of rectal cancer, 5 died of desmoid disease, 5 died of perioperative complications, not necessarily now, but if they get readmitted in the longer term because of an intestinal obstruction related to that primary surgery. It might be more apparent after pouch surgery. They

get that.

If you just take the issue of rectal cancer, of 222 ileorectal anastomosis, 22 developed rectal cancer, but only 5 died of it. We only have 1 rectal cancer under the age of 30. The risk of rectal cancer by the age of 50 is 8 percent, and between 50 and 60, it rises to 30 percent.

DR. RAGHAVAN: I'd just like to interrupt the discussion now. We've had questions. Thank you, Dr. Chiao. We have a little time for further open public hearing. We have listed Carolyn Aldige from the Cancer Research Foundation of America. Did she come? Is she here? If you can come up to one of the microphones, Ms. Aldige and please let us know if you've received any support from the sponsor for your appearance.

MS. ALDIGE: Good afternoon. As you heard, I'm Carolyn Aldige, President and founder of the Cancer Research Foundation of America.

Since 1985, the Cancer Research Foundation of America has relied on its NIH-approved, competitive, peer review process to provide cancer prevention research funding to more than 200 scientists at more

than 100 leading academic institutions across the country. We also sponsor a large number of cancer prevention public education programs, including our recently established National Colorectal Cancer Awareness Month, which by unanimous consent on November 19th, the Senate so designated.

I speak to you as an advocate for cancer prevention, research, treatment, and education. Searle has not provided any financial support to me or CRFA in connection with today's presentation.

15 years ago, when I founded the Cancer Research Foundation of America, cancer prevention, as you know, was not commonly discussed in the biomedical or clinical research communities. Today I'm delighted to say prevention research, including clinical trials of promising agents, is booming and enormous gains are being made.

The approval of tamoxifen for prevention of breast cancer in certain high risk women, something you all know very well, was certainly a watershed event, and now celecoxib for use in familial adenomatous polyposis is another important step forward. I believe this

application also deserves your recommendation for approval.

Though by any definition rare, FAP is a compelling public health challenge for three important reasons.

First, those persons born with the disease will contract and likely die of colon cancer.

Second, their care and treatment, combined with that of family members, causes both significant mental and financial hardship.

Third, slowing the development of colon cancer in people living with FAP represents a highly credible target for pharmacologic intervention. Immediate clinical benefit can be had now. At the same time the science of cancer prevention is pushed forward another step.

People living with FAP have few therapeutic options. Indeed, they too often face a highly uncertain future marked by colectomy and premature death. Available treatments provide little clinical benefit.

Celecoxib represents a new class of interventions, one which shows bright promise for

patient benefit. We at CRFA applaud the National Cancer Institute for sponsoring rigorous trials in this patient population. We recognize the limitations of trials in orphan diseases and we remain unsure if what is to be shown here today can truly be characterized as cancer prevention.

We believe, however, that a new and compelling therapeutic option has been shown. Celecoxib in FAP may one day show a survival benefit. In the meanwhile, patients living with this rare disorder will have access to an intervention that should slow the development of polyps, providing tangible clinical outcome. This is another step in the right direction, one which I hope will receive your endorsement.

Thank you very much for allowing me to speak to you today.

DR. RAGHAVAN: Thank you.

Are there any other people who would like to submit information or views to this meeting?

(No response.)

DR. RAGHAVAN: So, we'll move on now to the committee discussion and vote. Our habit has been to

identify two primary discussants for this phase of the process who try to set into context what we've heard today and what some of the remaining issues for discussion should be. Those two members are Dr. Christina Surawicz and Dr. David Kelsen.

Dr. Surawicz.

DR. SURAWICZ: I would like to congratulate not only the patients and family members and Ms. Aldige who just spoke, but also the sponsor and the FDA on the quality of their presentations, the answers to the questions, and the willingness to tackle difficult and diverse issues.

I think as we look at how we're going to prepare to answer to vote, I think we need to ask ourselves whether we consider this study to be adequate and well-controlled. If you'd like my opinion, I will say that I think that it is.

Given the results of this study, does celecoxib appear to be effective in the treatment of this proposed indication for FAP? Certainly I believe there is convincing evidence that it is effective.

Given the observed toxicity, does the risk-

benefit appear to be acceptable? I don't think we saw evidence of very much risk at all whatsoever.

Then I think the most difficult is, should this new drug application be approved and should there be any labels or tags or any modifiers that we need to do, should we vote to approve?

DR. KELSEN: I think the most revealing comments I heard this afternoon came from Dr. Phillips on the incidence of death from cancer. If I have the numbers correct, there are about 220-some patients in the data that he presented, approximately 4 or 5 percent of whom died of duodenal cancer and only a small percentage who died of rectal cancer. So, it's pretty clear that surgical intervention by colectomy and careful observation will prevent death from cancer in this disease, and I don't think we're faced anymore, except in the undiagnosed and unsurveilled patient, with a high, high risk of dying of colorectal cancer.

So, the data that was presented today really addresses the issue, as I think the sponsor and we have talked about now this afternoon, as to whether you can delay doing a major surgical procedure in a young

population. We don't know that from the data they presented today.

We're being asked -- and I think with some justification -- to make an assumption that a decrease in the number of polyps and a tiny decrease in the size of the polyps -- I think it's about 5 percent in the 400 arm -- significantly will change a doctor's opinion as to when he's going to perform a surgical procedure in a young boy or girl.

I'm not quite sure I've got that firm as to when they're not going to do this. I think what I've heard is they're going to do the colectomy pretty much on schedule in the vast majority of patients because they're not really comfortable that a 30 percent reduction decreases the risk of cancer, and they're going to be deciding to not do the rectal resection. And that's what that is really based on.

The follow-up study then would be crucial to me because if the follow-up study showed you prevented the polyps from developing and then you would prevent any surgical procedure, that would really be a major impact. I was hoping to be able to say that we could

follow the upper GI tract as a really tough surgical indicator to change outcome, but with 4 or 5 percent of patients dying of duodenal cancer over an observation period in a very rare disease, I don't think you'll ever be able to show that.

So, I think they showed evidence that you can decrease polyps. We're faced with the issue, do we assume then that they will change outcome on the basis of that decrease in polyp number, not so much size.

DR. RAGHAVAN: Would people from the committee like to free associate or make any comments? Dr. Nerenstone.

DR. NERENSTONE: This is obviously extraordinarily complex, but I wanted to ask the FDA what they thought about a registry of patients with this disease who were treated with Celebrex if we decide to go ahead and approve it.

I have to disagree with Dr. Kelsen in that he said, well, only 10 percent die of colorectal cancer. I agree I'm not sure what the observation period is. I think that's a big number. 10 percent die of it, but in fact there are even more cases that develop it and are

cured surgically but another surgical procedure.

If you kept a registry relatively simple of polyps or the need for polypectomy, of the need for revision of the rectal stump, of the development of cancer, and of the death rate, in 5 years would that give us some understanding of the evolution of this disease of rare patients who were treated in a novel way over time so that 5 years from now we can look back and say, this is the historic control?

We're not going to be able to say, oh, yes, there's a 5 reduction in the need for surgery. That's ridiculous, but we may be able to say instead of 10 percent of patients dying from cancer, we now have 2 percent of patients dying from cancer, or instead of 50 percent of these people needing revisions of their rectal pouch, we now have 2 percent.

It's not a study. It's really observation of clinical treatment over time. Is that something that the FDA might be interested in?

DR. TEMPLE: Well, in some sense you're describing a single-arm cohort study and the question will be how good the data are that are collected. It

would be better if you had everybody enrolled in an actual study, maybe less good if you tried to extract it from registries.

As you could probably tell from my questions, I'm not sure what the alternatives for the present population are actually going to be because I despair of the randomization proposed. It doesn't mean you couldn't study other things, of course.

Sure, that is one kind of thing we could think of. It would basically be an historically controlled study. There apparently are lots of data on outcomes, and if the difference was large enough, you might detect it credibly with all the limitations of those kind of data, however.

DR. RAGHAVAN: The difficulty with historically controlled studies, as you know as well as I do, is we have changing dietary patterns and carcinogens in the community and smoking and so on.

DR. TEMPLE: And surgical procedures and everything.

DR. RAGHAVAN: Dr. Blayney, you want to comment.

DR. BLAYNEY: Yes. I'm again quite concerned and struck by my recent experience with this illness professionally. But I think as an oncologist, when I read the package insert for cytotoxics, it says these must be used by a physician skilled in the management of cytotoxic agents. I think here we have equally as lethal and natural condition, not cytotoxics, but I would strongly advocate that if this is approved, that the label be very clear on who can use this agent because this big bugaboo, managed care, may say that, well, you don't need these expensive surgical procedures, we have a pill for this condition. So, I would strongly encourage you.

Secondly, this registry business does make some sense. There is this thalidomide registry that I have to participate when I use thalidomide for, admittedly, an off-label indication. There is a registry so this is not without precedent.

Thirdly, if you are going to the trouble of a clinical trial, I think it would be again useful to me as a clinician to not have -- and I think it would be a difficult IRB issue to get an IRB to approve a 100

milligram versus 400 milligram dose, given what we've seen. I would encourage you to think of a way to use another drug as a control. We've heard that Sulindac is used clinically and perhaps that would be a more legitimate control that would yield some clinically useful information.

DR. RAGHAVAN: Dr. Temple.

DR. TEMPLE: Well, the trouble is, for all the reasons I explained this morning, it's not easy to interpret that unless you have a very well established belief system about Sulindac.

DR. RAGHAVAN: Dr. Margolin.

DR. MARGOLIN: Well, as far as the design, I think none of us is going to be smart enough to figure out how to design the best post-marketing study and how to actually make it happen because some of the most brilliant ideas are most difficult in practice.

But I'll throw one out, which is just like with tamoxifen chemoprevention, or whatever word you want to use for it, what we think we can achieve is improvement in a significant morbidity endpoint. We don't know yet whether we're going to impact on the

mortality endpoint. As a person who has tried and has been through many of the frustrations of trying to do randomized trials and knows how difficult it is, I wonder if some kind of a now-versus-later design, where there truly is still a placebo control or a non-treatment control, would be possible where you can still use this mechanism to validate the surgical and polyp type endpoints, but also recognize the fact that we have not proven an alteration of the survival endpoint. You could probably say, well, if you cross over, then that's going to muddy that too, but it's at least just one suggestion.

DR. RAGHAVAN: I just would like I think to make the comment Dr. Johnson's grand-daddy once commented that a camel was a horse designed by a committee.

(Laughter.)

DR. RAGHAVAN: And I'm not sure that our role is necessarily to try to help the sponsor identify what studies it could or should do. We have some fairly focused things that we need to address, and I really would like to restrict the discussion to the questions

that the committee has seen and not to now move off into the province of what trials could we design for the company. I think they can come and discuss that with members of the committee afterwards. So, things that relate specifically to the questions.

Dr. Sledge.

DR. SLEDGE: I'm not sure I agree with you there.

DR. RAGHAVAN: Okay.

DR. SLEDGE: Because my understanding was that we were being asked to give significant input in terms of trial design from what I was reading.

My real question is whether or not any trial is possible. What I heard one of the clinicians say was that they were going to put 100 percent of their patients on this drug when it became available. I heard a member of our committee saying that basically everyone who is currently available and who isn't having a toxicity issue is on Sulindac or some drug like it.

It strikes me that the standard from all of our experts is that they're going to use these drugs and that they consider it ethical, appropriate, and nontoxic

to do so. That being the case, is any such trial ever going to be possible? It's a realistic concern.

DR. RAGHAVAN: Dr. Temple?

DR. TEMPLE: This has happened before in other oncologic settings. It may be possible to study arenas where people are not quite so sure, for example, sporadic polyps which was a suggestion that we have made.

You also wonder whether if you took people who were pre-phenotypic and had an endpoint that was phenotypic display of the disease, that might be an endpoint you could study now. I don't know how convincing it would be, but that's another possibility.

Or the other thought was that you could take people who had had their colon and rectum fully removed, therefore have no obvious benefit because the duodenal effect is really not established. Maybe that's a study people could enter into.

But you're right. It needs a lot of thought.

I guess I would like to say that you don't have to reach any of these things if you don't agree that they've shown a benefit. There was a sequence to

our questions that did suggest --

DR. RAGHAVAN: Which was what I wanted to address.

DR. TEMPLE: First see if you absolutely agree that there is a benefit shown of some kind and then move on to the really impossible questions.

DR. RAGHAVAN: Dr. Simon.

DR. SIMON: I agree with Dr. Sledge. I think we're asked to vote on accelerated approval, and that means that we believe that there's something that will hold up in some subsequent confirmatory trial. So, although I don't know that we can design a subsequent trial in any detail, I think for me to go forward with voting for accelerated approval, we have to have some confidence that there can be some follow-on trial that shows some clinical benefit to some subset of patients.

My concern is for any drug, there's a window of opportunity when you can do clinical trials, and I'm concerned whether that window has passed here and that what should have been done was a clinical trial that showed some clinical benefit during the period of time when that window that was open.

DR. RAGHAVAN: I guess my response, however, is that if you can't design the appropriate trial, that may not be a reason to hold up the drug. So, you end up with a circuitous argument. If one takes to its logical extension what you've said, you could hold the company ransom. If can't design a trial that we like, we're not giving approval. And that's not necessarily in the patients' best interests.

Dr. Temple, do you want to comment?

DR. TEMPLE: Well, but even before that, we laid out the questions not accidentally, with the first one being, are you persuaded that there's a finding here. Actually you don't have to worry about those other things which are extremely interesting and difficult if you don't think they've got a persuasive finding. There's only one study. Judy suggested various things to think about and how convincing it is. That still seems like, in some ways, job 1.

The second thing for accelerated approval is that you believe that what has been shown is, quote, reasonably likely to predict clinical benefit.

After all that, then you get to the question

of whether they can ever validated it, and that's a very thorny question. But those first two things seem to us to come first.

DR. RAGHAVAN: And I agree with that. I think we need to address the early questions first, and then we can open up our free association.

So, we have written, as usual, tons of information and graphs, pictures, and so on. So, in the interest of time, I'm going to go straight to the preamble to the questions to the committee.

We have concluded that there was, on the Celebrex 400 milligram b.i.d. arm, an approximately 25 percent reduction, compared to placebo, in the identified focal areas in a single controlled study. Blinded committee assessment of videotaped endoscopies of the rectum revealed that 21 percent of evaluable patients rated "better" by four reviewers at 6 months compared to baseline on the Celebrex 400 milligram b.i.d. arm. There's no apparent effect on the duodenum, an important source of malignancy. There are no data that address the issue of persistence of effect beyond 6 months. Reliance on these data as a basis for approval

of Celebrex for "reduction and regression of adenomatous colorectal polyps in FAP patients" poses a number of difficult questions about the persuasiveness of the finding, the clinical meaningfulness of the finding, the specific population the drug would be indicated for, the precise use to which the drug would be put, and how dangerous behavior could be avoided/prevented, and the ability to assess the ultimate clinical value of the treatment.

I have to apologize to Mr. Phillips that I did not write the syntax or grammar in that question.

(Laughter.)

DR. RAGHAVAN: Persuasiveness of the finding.

DR. TEMPLE: What's wrong with it?

(Laughter.)

DR. RAGHAVAN: That's the essential difference between the United Kingdom and the United States.

(Laughter.)

DR. JOHNSON: Actually that's not exactly true, Dr. Raghavan. Let me tell you the equine statement of my grandfather just for your edification.

(Laughter.)

DR. JOHNSON: You can't teach a jackass English is what he used to say. So, just in case you want to know.

(Laughter.)

DR. RAGHAVAN: So, question number 1. Persuasiveness of the finding. The study endpoint reflects changes in colorectal polyps in focal areas. Overall assessment of videotaped colorectal endoscopies showed improvement in some patients. Treatment was not associated with a statistically significant reduction in duodenal plaque-like polyps compared to placebo.

The question: Do you believe the observed focal effect on colorectal polyps is a reasonable indicator of the effect in the whole colon and rectum or in the whole GI tract?

Dr. Surawicz? Dr. Kelsen?

DR. KELSEN: Well, I'd break them into two since they broke them into two. I would first say do you believe that the observed focal effect on colorectal polyps is an indicator of the effect on the whole colon and rectum. I think that it is in the order of

magnitude that they describe based on their graphs showing videotaped correlations between the focal area and the rest of the colon.

DR. RAGHAVAN: Do people have comments that they wish to add as opposed to reiterate?

(No response.)

DR. RAGHAVAN: So, let's divide the question, with permission from its author, whoever that may be --

(Laughter.)

DR. RAGHAVAN: -- to restrict itself to the colorectum. Those who believe that there's the observed focal effect on colorectal polyps is a reasonable indicator of the effect in the whole colorectum. Those who do believe it? Hands held high.

(A show of hands.)

DR. RAGHAVAN: Thank you.

Those who don't believe it?

(No response.)

DR. RAGHAVAN: Those who abstain.

(No response.)

DR. RAGHAVAN: So, there seems to be a consensus.

Then for the second part of the question, as modified by Dr. Kelsen, da-da-da-da-da, for the whole gastrointestinal tract, implying the duodenum as well. Those who believe it? Hands high.

DR. JOHNSON: Again, we're talking about the 400 milligram b.i.d. dose.

DR. RAGHAVAN: Correct.

(No response.)

DR. RAGHAVAN: Those who abstain?

(No response.)

DR. RAGHAVAN: Those who don't believe it?

(A show of hands.)

DR. RAGHAVAN: So, I think again consensus, if I read -- yes, I think Dr. Surawicz voted --

DR. SURAWICZ: I should have abstained. I just don't think I have enough information.

DR. RAGHAVAN: Okay. So, 1 abstention.

All right. So, I think therefore I would call that a yes and therefore we proceed through the cavalcade of questions.

Question 2, the study lasted 6 months. Do you believe it provides adequate evidence of a

persistent effect on colorectal polyps?

Any discussion?

DR. SURAWICZ: Can you clarify what the question means?

DR. RAGHAVAN: I'm sure Dr. Temple could clarify.

DR. SURAWICZ: Can you clarify? The question means that it had an effect over that 6-month period? Because we don't know what happened after that. That wasn't part of the study.

DR. TEMPLE: Well, that's the question. But you're going to have to decide, somehow or other, whether you think this represents an effect that should be presumed to be chronic. Obviously, you don't have further data, but if you thought, for example, that the effect abruptly ended, hit a wall at 6 months, this would not be very attractive.

DR. RAGHAVAN: Dr. Margolin?

DR. MARGOLIN: Well, just a clarification of a clarification. I assume that you mean that if they were to stop therapy, that the polyps that regressed or that didn't would come back. Is that correct?

DR. TEMPLE: No. I mean if you presume that the drug will be continued for longer than 6 months, is it a reasonable assumption that it will continue to be effective. That's what I really mean.

DR. RAGHAVAN: So, just to clarify the question, the study lasted 6 months. So, we have data that extend out to 6 months. And the question that Dr. Temple is trying to address is, if one accepts the data from the 6 months of study, do we then assume that if you continue this medication ad infinitum, control of the polyps would be retained for a significant period of time, maybe not permanently, but for a long, protracted period of time, or would it be an effect that would run out shortly thereafter?

Dr. Nerenstone.

DR. NERENSTONE: I think that's what we don't know, and that's why I think so many of us feel uncomfortable answering that question, even if we think this drug should be approved, without attaching some sort of follow-up on those patients.

DR. RAGHAVAN: Why don't try to answer the question as it's phrased? And then we can discuss the

implications --

DR. SURAWICZ: But we do know from Sulindac patients that when you stop the Sulindac, the polyps come back. That we do know.

DR. RAGHAVAN: From Sulindac patients, do we have an upper end of control of the polyps that you can quote?

DR. SURAWICZ: No. I think most of those studies are pretty uncontrolled, but when you stop the drug, the polyps come back. So, that piece of information biases my answer to this to say yes.

DR. RAGHAVAN: Dr. Kelsen?

DR. KELSEN: Well, I think the way it's written, it's very straightforward. Do you believe that the study provides adequate evidence of a persistent effect on colorectal polyps? The study does not provide adequate evidence to me of a persistent effect because they didn't look.

DR. TEMPLE: I should probably modify it. Really, we're going to have to conclude, one way or another if we were to say yes to this, that there's reason to believe the effect will persist. That's

really what I'm trying to elicit. I'm sorry for the imperfection. I acknowledge the imperfection of this one part.

(Laughter.)

DR. RAGHAVAN: So, therefore, I think what I was hoping to hear from you, Dr. Temple, is what the question was getting at. So, what the question is getting at is, is it a reasonable presumption of this committee that if this drug passes muster and is out there for a period beyond 6 months, is it a reasonable assumption that the effect will continue for a clinically relevant time? It could be a year. It could be 5. I think we've all agreed we don't know the answer to that. But is it a reasonable presumption of this committee that if they drug is out there and patients take it, that they will get more than a short-term benefit?

Dr. Johnson.

DR. JOHNSON: I have a question to ask, again, the experts, and that is, in those individuals who get Sulindac, continue on Sulindac, don't have any of the adverse effects, what percentage of that patient

population redevelops polyps or new polyps while receiving the Sulindac? Is the answer none, some, a lot, all?

DR. JACOBY: I think the best study on Sulindac was published by Giardiello in the New England Journal. In that study, when patients were stopped, the polyps grew back, indicating that at that period of time there was still an effect of the drug.

In the MIN mouse, I've done studies that aren't published where we start and stop treatment and we still get an effect even at later time points.

I think if you look at the epidemiologic data where people were treated with this same category of NSAIDs for 10 or 20 years, there's still a benefit found after that interval.

DR. JOHNSON: The only other question that I would have -- and maybe it's too late to ask the sponsor this question, but one of the questions we might have asked is why did those polyps that failed to regress fail to regress. They did, in fact, biopsy some of these. Were COX-2 expression levels different and what happens in the MIN mouse when that happens?

DR. JACOBY: Well, not all polyps express COX-2.

DR. JOHNSON: Correct. So, maybe that was the situation here?

DR. JACOBY: It would be an interesting thing to study, whether the ones that failed to respond failed to express COX-2.

DR. RAGHAVAN: Does the sponsor have an answer to that? No, okay.

So, can I put the question? The study lasted 6 months. Do you believe it provides adequate evidence of an anticipatedly likely persistence of effect if the drugs are continued indefinitely?

Those who believe that it provides adequate evidence of a persistent effect, hands up high.

(A show of hands.)

DR. RAGHAVAN: I can't tell whether, Dr. Sledge, are you scratching your nose?

DR. SLEDGE: Scratching my nose.

(Laughter.)

DR. RAGHAVAN: So, I'm sorry. Being old, I don't see well. So, high high.

(A show of hands.)

DR. RAGHAVAN: 6.

Those who believe it does not show that?

(A show of hands.)

DR. RAGHAVAN: 6.

Abstentions?

(A show of hands.)

DR. RAGHAVAN: 3, and that's 15.

There is only a single study supporting effectiveness. Is the single result so persuasive that you believe it should be accepted as evidence of a sustained reduction in focal polyps?

DR. TEMPLE: In this case, obviously, sustained for the 6 months of the study. Do you believe the finding?

DR. RAGHAVAN: Discussion? Dr. Surawicz? Yes or no?

DR. SURAWICZ: I'm looking at my watch as someone with an eye on the non-stops to the west coast, so I'm going to keep my comments and just vote.

DR. RAGHAVAN: You have no comment, okay.

Dr. Kelsen.

DR. KELSEN: I think the way it's ascribed, did it last for 6 months, if I've just interpreted your comments correctly, is it that the patients responded and they stayed in response for the 6 months?

DR. TEMPLE: No. This is a question about the weight of the data. Ordinarily we expect findings to be replicated. There's provision in a recent modification of the law to accept an unreplicated finding based on various reasons. Is this one of those cases where you find that study still strong or supported by other things you know or something that it should be believable without replication? That's the question.

DR. KELSEN: Well, since I voted yes on 2, I would vote yes on 3. And I assume the 6 people who voted no on 2 will vote no on 3.

DR. RAGHAVAN: Well, let's see.

DR. JOHNSON: No, because it's a clarification that this was limited to the 6 months.

DR. KELSEN: Yes. I think there was an effect for 6 months in the study, and I think a single study in this disease is adequate.

DR. RAGHAVAN: So, we'll put the question. Those who do think it should be accepted as evidence of a sustained reduction in focal polyps through the period of 6 months.

(A show of hands.)

DR. RAGHAVAN: Those who do not?

(A show of hands.)

DR. RAGHAVAN: And Dr. Margolin, do you or don't you or are you abstaining or not voting?

DR. MARGOLIN: The sustained effect for 6 months?

DR. RAGHAVAN: Yes.

DR. MARGOLIN: Why don't I abstain?

DR. RAGHAVAN: All right. Dr. Margolin abstained.

Let us continue on. Rather than reading a lengthy preamble on significance or meaningfulness, as it's described, I'd like to go straight to the next question which is do you believe that a reduction in colorectal polyp count in FAP patients in focal areas of some magnitude is "reasonably likely" to predict benefit, assuming that all other aspects of patient care

are unaltered? Explain what clinical benefits might be predicted.

All right. Well, let's do the first part of that. Do you believe that a reduction in colorectal polyp count in FAP patients in focal areas of some magnitude is "reasonably likely" to predict benefit, assuming that all other aspects of patient care are unaltered?

Does anybody want to comment, or can I put it straight to the vote? Straight to the vote, okay.

So, those who believe that a reduction in colorectal polyp count is reasonably likely to predict benefit?

(A show of hands.)

DR. RAGHAVAN: 11? 12.

Those who do not?

(No response.)

DR. RAGHAVAN: And those who abstain? There should be 3 abstentions.

(A show of hands.)

DR. RAGHAVAN: We'll have to do that one again, and please, hands up. I know you're tired but we

want to get this done.

Those who do?

(A show of hands.)

DR. RAGHAVAN: 12. 12 who do.

Those who do not.

(No response.)

DR. RAGHAVAN: Those who abstain?

(A show of hands.)

DR. RAGHAVAN: So, no negatives and 1 abstention? 2 abstentions.

DR. RAGHAVAN: Do you believe that the observed reduction, about 25 percent at 6 months, is likely to predict benefit in FAP patients, assuming treatment is otherwise unaltered?

Dr. Temple, I guess I missed the subtlety of the difference.

DR. TEMPLE: Well, sorry. The first question is about whether you think that endpoint is potentially credible. The second is whether you think this magnitude of reduction, 25 percent, is credible.

DR. RAGHAVAN: So, is 25 percent enough? Those who believe it is, a 25 percent reduction is

enough?

(A show of hands.)

DR. RAGHAVAN: 12.

Those who do not believe it's enough?

(No response.)

DR. RAGHAVAN: Abstentions?

(A show of hands.)

DR. RAGHAVAN: 3 abstentions.

Dr. Temple, I'm going to come back to explain what clinical benefits might be predicted because I suspect that's going to head us to free association, and we'll try to get the answers done so Dr. Surawicz can get away.

If the answers so far are yes -- and by definition, they must be for us to be here -- do you believe, without further data, that we can be reasonably sure or can draft labeling or other mechanisms to allow assurance that treatment will not be altered because of a belief that it is now "safe" to delay surgery?

DR. TEMPLE: Can I add one modification? When I wrote this, I figured you'd do surgery exactly when planned. It's fairly obvious from the discussion

that people would be unaltered in the sense that they'd use the same criteria for deciding on surgery, and I think that's a more realistic question, even though that raises some risks too.

DR. RAGHAVAN: So, just say the question as you'd like it modified.

DR. TEMPLE: You can answer the question with that modification I think.

DR. RAGHAVAN: Okay. So, those who would answer in the affirmative with Dr. Temple's modification?

(No response.)

DR. RAGHAVAN: I don't think anybody understood your modification.

(Laughter.)

DR. TEMPLE: People have sounded these alarms all through the day. They're afraid that people will not follow patients as rigorously, will make assumptions about how protected they are, won't look at the duodenum, all kinds of stuff like that. Do you think we can convince people who will not 100 percent probably be specialists in this to behave properly so they don't do

harm? And some suggestions would be good.

DR. JOHNSON: It's the very thing that Dr. Blayney touched on earlier. It's the very issue about if you approve this, will every oncologist in America suddenly or family practitioner or internist or gynecologist start giving Celebrex to the patients in just hopes that they don't --

DR. RAGHAVAN: So, do we believe that people believe and obey package inserts is the question.

DR. TEMPLE: Or other mechanisms, patient inserts, for example.

DR. RAGHAVAN: Okay. So, those who do believe that we can do that? Those who do believe it?

DR. SURAWICZ: We have to believe in education.

(A show of hands.)

DR. RAGHAVAN: Slowly rising hands. 11 who believe it.

Those who do not believe it?

(No response.)

DR. RAGHAVAN: Those who abstain?

(A show of hands.)

DR. RAGHAVAN: 4, okay.

DR. SLEDGE: Could I make a comment on that? I do think that half of all doctors graduate in the bottom 50 percent of their class.

(Laughter.)

DR. SLEDGE: I guess my question is whether or not this is a case for a black box in the indication.

DR. PAZDUR: But remember also other drugs, nonsteroidals, are being used in this area right now.

DR. TEMPLE: That's not a reason not to black box. It also is a reasonable candidate for patient labeling because the patient should be part of inducing monitoring, and we can certainly do that.

DR. RAGHAVAN: Do you recommend approval of Celebrex under the accelerated approval rule for some treatment of FAP?

Those who do recommend approval?

(A show of hands.)

DR. RAGHAVAN: Those who are opposed to it?

(No response.)

DR. RAGHAVAN: Those who abstain?

(A show of hands.)

DR. RAGHAVAN: Dr. Simon. 1 abstention.

If yes, please consider the indication that should be approved, for example, for use as an adjunct to usual care, not as a substitute for any aspect of monitoring or surgery that would ordinarily be used, in the treatment of FAP. We would add details of what has and what has not been shown. We would add details, the FDA. Also consider needed warnings and precautions that should be included in product labeling.

All right. So, please consider the indication that should be approved.

DR. KELSEN: Well, it sounds like the information we received is for patients who have already developed the phenotype and not for patients who simply have the genotype. So, I would think that the data we've seen is on patients who have polyposis, not patients who are in families who may already have had a harvest and been shown to carry the gene. So, I would suggest that the indication be for patients who have established phenotypic FAP, not for patients who are in the follow-up study who have genotypic FAP but are not yet phenotypically presenting.

DR. SANTANA: As a follow-up to issues of warnings and precautions, I think we mentioned earlier the lack of substantial pediatric data on safety issues of pregnancy and also drug interactions because if these patients are going to stay on this drug for an extended period of time, they're going to be taking a lot of other stuff.

DR. RAGHAVAN: So, I'm not sure how the committee can do this, Bob. Give me some advice.

DR. TEMPLE: Well, we're asking you to sort of free associate, which you're doing on things --

DR. RAGHAVAN: So, that's helping you?

DR. TEMPLE: -- and sort of general comments about the indication would be helpful. You can go as far as you want. We've had one specific one to say it's only for people with phenotypic disease.

DR. RAGHAVAN: We have experts here who spend their lives delving into the bowel. You guys should be giving some advice here. So, come on. Dr. Jacoby.

DR. JACOBY: I think one thing that is useful, there was a paper also by Giardiello out of Hopkins looking at the use of genetic testing in FAP,

and he commented on the fact that more frequently there was misinterpretation when it was done by general practitioners than when it was done by specialists in the field. I think this is another situation where it would be advantageous to have the patient referred to a center that is familiar with treating these patients.

DR. RAGHAVAN: Other indicators? Dr. Lewis, anything to add?

DR. LEWIS: Not really. I think the appropriate indication is what has been mentioned in the statement. It's to be used as an adjunct, and I agree that it should probably be restricted to the patients that have been studied to date.

They're planning to do, I think, the adolescent study. It will have to go into the labeling, its use in people under the age of 18 has not been studied, and that should be studied. But I would leave it at that.

DR. RAGHAVAN: Any other advice for the FDA? Dr. Brand?

DR. BRAND: My only concern, which I share with Dr. Jacoby, is it's going to bring a lot of people

doing genetic testing, and that's going to open up a whole other can of worms here, unless you stick to phenotypic expression. I think dealing with some pancreatic cancer families, it's a big issue too. We haven't even touched upon some of the ethical issues that the use of this drug is going to bring into play.

One factor I don't know is whether it impacts my surgical colleagues with the use of this drug in terms of a preference of going for ileorectals over ileal-anals with this, which can certainly benefit quality of life. I don't know if anything can be commented about that with the course of this study design.

DR. RAGHAVAN: Probably this is not necessarily the place to do that, but I'm sure they have friends they can call.

Do you have enough information on the answer to this?

DR. TEMPLE: Yes.

DR. RAGHAVAN: Needed warnings and precautions?

DR. TEMPLE: Oh, no. Anything else you have

to say about that.

On the question of post-approval study, there has been a lot of comment on that. So, if you feel you can go further and say more, that's fine, but I think we've heard a fair amount. It's a difficult problem obviously.

DR. RAGHAVAN: Any particular warnings or precautions people would like to advise Dr. Temple and his merry men and women?

(Laughter.)

DR. SLEDGE: Well, it is kind of unusual, but I really do think, as I suggested before, that you really ought to include a black box that specifically mentions the need for being unaltered in terms of your diagnostic and surgical approach.

DR. RAGHAVAN: If responses regarding the persuasiveness or meaningfulness of the finding are no -- they were yes.

DR. TEMPLE: Skip that.

DR. RAGHAVAN: So, we can pass on.

If accelerated approval is recommended, the applicant is required to study the drug further to

verify and describe its clinical benefit. Please comment on the acceptability of the applicant's proposed post-approval study, including the study population, adolescents with FAP; choice of control; and primary efficacy endpoint, the proportion of patients who require colorectal surgery by age 21.

We've had some discussion and raised significant concerns. I think Dr. Sledge raised or vocalized the concern that once this has got to this stage, it may be difficult to design the appropriately controlled study. That window has opened and closed I suspect.

What do you want us to add to that concern? Because it's probably the limiting one.

DR. TEMPLE: Well, I think you have talked about this a lot. Either now or by letter later, if you have any bright ideas about what studies can be feasible or how to do them, that would be helpful.

What I've heard so far is at least one possibility that you're not going to do much better in the population studied than an open study. That may be true. It sounds like you could study people who are

pre-phenotypic. That's seems interesting and it's a pediatric study which is interesting.

I wondered whether you could still carry out a study in people who had had both colon and rectum removed, for whom there is no terribly obvious benefit. You can even imagine randomization to one of several different surgeries, one more onerous than the other, and people willing to be randomized in the presence of the drug to see if you can get away with the less onerous surgery.

So, we'll think about a lot of this stuff. I'm sure the company will. Any thoughts anybody has are welcome.

DR. RAGHAVAN: Do any of the regular committee or experts have advice to the company, advice to the FDA, advice to patients, something to add beyond the discussion?

(No response.)

DR. RAGHAVAN: Do you agree that the proposed study is adequately designed to demonstrate clinical benefit of Celebrex therapy in FAP patients?

What do you think? Dr. Kelsen?

DR. KELSEN: You're sort of asking what we've already discussed.

DR. TEMPLE: Yes. I really think you've --

DR. RAGHAVAN: Are you happy with that? Okay.

DR. TEMPLE: You've done it almost all.

DR. RAGHAVAN: There's probably nothing to do in 4.

The one thing that we glossed over earlier, which again was in the free association category, explain what clinical benefits might be predicted. I mean, we've sort of covered that as well, really.

So, Dr. Temple, we've all worked hard. Dr. Blayney and I have missed our taxi to the airport. But my question remains. Is there anything else that relates to the topic at hand that the FDA needs advice on from this panel?

DR. TEMPLE: No. Just Thank you.

DR. RAGHAVAN: I'd like to thank the sponsor for an elegant presentation of information, the committee members for working hard. Thank you.

(Whereupon, at 3:35 p.m., the committee was

adjourned.)