

Committee Questions on Rofecoxib to Dr. Braunstein & Dr. Villalba

JOINT MEETING OF THE ARTHRITIS ADVISORY COMMITTEE AND THE DRUG SAFETY AND RISK MANAGEMENT ADVISORY COMMITTEE

February 16-18, 2005, Hilton Gaithersburg, 620 Perry Parkway, Gaithersburg, Maryland.

Highlights.....	1
Discussion Text.....	6

Highlights

NEED TO PRESENT ALL SIGNIFICANT RISKS:

- **SHOW PULMONARY EDEMA DATA:** Dr. Wood said that “since the primary object of this Committee is to evaluate all the risks and benefits that these drugs can produce” he was “very surprised” not to see the Kaplan-Meier curve for pulmonary edema from the APPROVe study. Dr. Nissen added that “heart failure and pulmonary edema would be helpful”. Dr. Braunstein showed the slide and said that this was a recognized NSAID side effect and was in the labeling.
- **SHOW HAZARD RATIOS OF 4.6 & P<.05 IN PRESENTATIONS:** Dr. Wood pointed out that this had a hazard ratio of 4.6 and a p value of <.004. Dr. Wood said “It’s important for the Committee, and this goes for all the speakers I think, that if there are other hazards with a hazard ratio of 4.6, that we see these as they are presented, so that we can make a cumulative estimate of what the hazards are from these drugs”.

CONCOMITANT ASPIRIN:

- **RELATIVE RISK WITH ASPIRIN ON BOARD:** Dr. Schafer followed up an earlier question from Dr. Nissen and asked about the relative risk with rofecoxib when aspirin was on board. Dr. Schafer pointed out that in the APPROVe trial with aspirin on board the relative risk is 3.25 “with a confidence interval which is wide, as Dr. Fitzgerald had suggested it might be, because of small numbers” but it goes from 0.98 to 13.81. Merck interpreted the VIGOR results as based on an “aspirin-like effect” of naproxen but Dr. Schafer suggested

that these APPROVe results “essentially disprove the Naprosyn hypothesis”.

- **EARLY AND LATE CV EFFECTS MAY HAVE DIFFERENT MECHANISMS:** Dr. Braunstein suggested that the mechanism for the cardiovascular effect may be different in VIGOR (where the effect occurred earlier) and APPROVe and that if one looks at the APTC endpoint “the difference actually seems to go away”.
- **ROFECOXIB WORSE THAN NAPROXEN EVEN IN ASPIRIN SUBSET: COULD BE BP EFFECT:** Dr. Temple commented that the aspirin group was a “baseline subgroup” of about 1,000 patients and they were randomized to either rofecoxib or naproxen. With combined aspirin and rofecoxib “there is plenty of COX-1 inhibition” so that the combination represents non-selective NSAID therapy. Thus,

the difference in cardiovascular hazard ratio between rofecoxib and naproxen does not seem to be explained by the COX-2 hypothesis, and he wonders if the increased risk could be related to something like a blood pressure effect which would not be prevented by aspirin. Dr. Wood pointed out that patients were not randomized to aspirin. Dr. Braunstein said their aspirin data “were not robust enough” and there were not many events (Dr. Temple interjected that there were 16).

- **IN CV EFFECT, NAPROXEN-ROFECOXIB DIFFERENCE MUCH GREATER THAN WITH ASPIRIN-PLACEBO:** Dr. Nissen pointed out that naproxen would have to be a lot better than aspirin in preventing cardiovascular events to explain the increased events on rofecoxib on the basis of a beneficial effect of naproxen – and that this seemed unlikely.

BLOOD PRESSURE:

- **BP EFFECT INSUFFICIENT TO EXPLAIN CV EFFECT?** Dr. Braunstein said that the blood pressure effects “would not appear to explain the magnitude of the cardiovascular findings”. Dr. Temple responded that “one of the reasons to worry is that people with diabetes or heart disease are probably more susceptible to the blood pressure effect”. Dr. Nissen also expressed concern about the greater blood pressure effect of rofecoxib compared with other coxibs.
- **MORE BP DATA PRESENTED:** Merck provided more blood pressure data on rofecoxib since they had only

shown data from APPROVe (which showed the “expected” NSAID difference from placebo) (see end of this section for slides).

- **BP EFFECTS OF ROFECOXIB & CELECOXIB SIMILAR IN ONE TRIAL:** Merck described a 14-day trial in the elderly that compared the BP effects of rofecoxib 25 mg/day and celecoxib 400 mg/day (that “at that dose have similar inhibition of COX-2”) as well as naproxen 500 mg bid and placebo and showed similar mean changes in systolic and diastolic blood pressure with each NSAID

that was greater than that seen with placebo.

HEART FAILURE:

- **OA DATABASE SHOWS HEART FAILURE SIMILAR WITH ROFECOXIB AND NSAIDS:** Merck showed an additional slide from their osteoarthritis database that showed that “the incidence of heart failure was low and was generally similar to ibuprofen” and also diclofenac. However, there was one epidemiology study that suggested a higher rate with rofecoxib.
- **HEART FAILURE IS AN EARLY EVENT:** Dr. Abramson asked if heart failure was an early or a late event in the study. Dr. Braunstein said that it was an early event.
- **SHOULD HEART FAILURE BE INCLUDED IN COMPOSITE CV ENDPOINT?** Following the presentation by Dr. Villalba (FDA), Dr. Furberg suggested that heart failure should be included together with myocardial infarction and stroke in evaluating “the cardiovascular signal”. Dr. Villalba said that apart from an increase in heart failure with rofecoxib versus naproxen in the Vigor study, we do not have heart failure data on other NSAIDs.

MORTALITY:

- **INCREASE IN TOTAL MORTALITY:** Dr. Hennekens expressed concern about the increase in total mortality with rofecoxib in addition to the increased cardiovascular risk. Following the presentation by Dr. Villalba (FDA), Dr. Fleming pointed out that more than half of all-cause mortality occurred in the VIGOR and Alzheimer’s studies which is where “we are seeing the signal”. However the Advantage study “still had one more death” and the cardiovascular events were 4 to 0 in the wrong direction.

OTHER:

- **PROBLEM INTERPRETING ROFECOXIB CARDIAC SAFETY LABEL:** Dr. Wood and later Dr. Shapiro expressed concern about the wording of the rofecoxib cardiac safety label change, and how a physician was supposed to interpret it.
- **BASIS FOR ROFECOXIB WITHDRAWAL:** Dr. Crawford asked about the level of signal that made Merck decide to withdraw

rofecoxib. Dr. Cush asked whether, if Merck had known that the cardiovascular risk was a “class effect” at the time the decision was made to withdraw rofecoxib, “would you have made the same decision” Dr. Braunstein said “I couldn’t go back and speculate”.

- **INCLUDE/EXCLUDE NAPROXEN FROM POOLED NSAID ANALYSIS:** Dr. Schafer asked about the justification of doing a separate analysis of rofecoxib and naproxen rather than including all comparative NSAID groups together.
- **RISK-BENEFIT:** Dr. Gibovsky asked if Merck had calculated the risk-benefit profile in colorectal polyps. Dr. Braunstein said they had not done this. Dr. Fleming asked if Merck had tried to compare the risk-benefit profile for gastrointestinal versus cardiovascular toxicity. Dr. Braunstein said no.
- **LACK OF FOLLOW-UP MIGHT UNDERESTIMATE RISK:** Dr. Fleming and Dr. D'Agostino expressed concern that the absence of follow-up beyond 2 weeks post-discontinuation might be underestimating the level of risk. Dr. Villalba (FDA) said “that is a good question for the sponsor” but that they did try to follow-up the patients.
- **PRECIPITANTS OF MI:** Dr. Cannon asked if myocardial infarction appeared to be precipitated by some intervention such as CAPG

surgery. Dr. Braunstein said this did not appear to be the case.

- **ROFECOXIB INDICATED FOR JRA:** Dr. Ilowite pointed out that rofecoxib was one of the few drugs indicated for Juvenile Rheumatoid Arthritis and that without it patients had fewer options.
- **ABSOLUTE RISKS IMPORTANT AS WELL AS RELATIVE RISKS:** Dr. Platt emphasized that it was important to consider absolute risks in different populations, rather than just relative risks.
- **IS GI SAFETY CLAIM WITH ROFECOXIB RELATED TO COMPARATOR?** Dr. Cryer asked that, since rofecoxib’s unique labeling for increased GI safety was based on the VIGOR trial (which compared rofecoxib with naproxen) the GI safety difference might have been a result of the comparator selected rather than the degree of COX-2 inhibition – particularly since the rofecoxib comparisons with diclofenac did not show a clear GI safety advantage. Dr. Braunstein said that he thought rofecoxib would have been superior to diclofenac in GI safety “in an adequately powered study”.

Incidence of Congestive Heart Failure in OA 6-Week/6-Month Population

	Rofecoxib 12.5 mg N=1215	Rofecoxib 25 mg N=1614	Rofecoxib 50 mg N=476	Ibuprofen 2400 mg N=847	Diclofenac 150 mg N=498
Incidence (%) of CHF ¹	0.4	0.1	0.0	0.4	0.8

¹ Congestive heart failure and left cardiac failure

1722

Change from Baseline for Blood Pressure (mm Hg) at Day 14 of Treatment

Treatment	LS Mean Change from Baseline (mm Hg) ¹ at Day 14	
	Systolic	Diastolic
Rofecoxib 25 mg (N=17)	3.4	0.3
Celecoxib 200 mg BID (N=16)	4.3*	0.8
Naproxen 500 mg BID (N=15)	3.1	-0.4
Placebo (N=14)	1.3	1.2

¹ Change from baseline (day 1) is the average of 8 am, 12 noon, 4 pm and 8 pm blood pressure measurements on day 14.

* Significant within treatment change from baseline (p<0.05).

Pharmac 908

232

Discussion Text

DR. WOOD: Great! Thanks very much. As I am sure you would agree, the primary job of this committee is to assess all the risks and benefits that these drugs can produce, and we have certainly been encouraged to do that by everybody who has spoken so far. That being the case, I was very surprised not to see the Kaplan-Meier curve for pulmonary edema. can you show us that from the APPROVe study?

DR. BRAUNSTEIN: Certainly. That would be slide 213.

DR. NISSEN: Yes, heart failure and pulmonary edema would be helpful.

DR. BRAUNSTEIN: We certainly examined that. You know, the question that has been on the table--we believe the hypothesis we were exploring was the incidence of thrombotic cardiovascular events. Pulmonary edema is a mechanism-based effect of selective COX-2 inhibition that has been well appreciated and, in fact, is already described in product labeling. So, we did see an effect. This is in our publication. As shown here, we saw an effect. This is a combined endpoint of congestive heart failure, pulmonary edema of cardiac failure, so all congestive heart failure type of events that we observed in the study.

DR. WOOD: And this had a hazard ratio of 4.6 and a p value of less than 0.004. Right?

DR. BRAUNSTEIN: Yes.

DR. WOOD: So, I mean, it is important for the committee--and this goes to all the speakers I think, that if there are other hazards with a hazard ratio of 4.6, that we see these as they are presented so that we can make some cumulative estimates of what the hazards are for these drugs. Just because they are in the label does not mean we shouldn't hear about them here, it seems to me. The second question I have, which has always worried me, is when you go back to the original label change that you made, you know, when you changed the label to say caution should be exercised when Vioxx is used in patients with a medical history of ischemic heart disease, as a physician what am I supposed to do with that? Am I supposed to say to patients take the drug slowly, or swallow it with milk, or only take it with the lights on? Tell me what I am supposed to do with that information. I am not being facetious here because as we go through this process we are going to have to decide how we make whatever labeling changes we make, if that is the decision we make, and that doesn't seem to me to have been helpful. But maybe you knew something that I didn't. So, what did you intend me to do with that information?

DR. BRAUNSTEIN: At the time when we conducted negotiations with discussions with FDA on that labeling, there were no specific data that showed a statistically significant increased risk in one patient group or another. However, given the uncertainty in the data, it was felt to be prudent to recommend that caution be exercised in that patient group

if you are considering using the drug. What we meant by that was that you need to carefully weigh the risks and benefits of the different treatment options. We think that when evaluating the options on patients it needs to be done on an individual patient case-by-case basis. Patients differ with respect to their cardiovascular risks, with respect to their GI risks, with respect to their history of allergies and with respect to how they responded to these different medications in the past, and all of that information needs to be taken into consideration when assessing and determining what type of therapy should be used versus another. And, we felt that one of the things that should be considered was cardiovascular history, and that is what we meant by that.

DR. WOOD: Okay. Other questions? Stephanie?

DR. CRAWFORD: Thank you. I appreciate the presentation. I heard both the speakers say that the sponsor, Merck in this case, made the decision to voluntarily withdraw rofecoxib in the interest of public health although the drug could have been continued on the market. When we look at adverse events we desire to predict uncontrollable events and control controllable events. The bottom line question which is really important to me as we consider these issues when we look in this case at the issue of hazard of cardiovascular events is how much is too much? In other words, how did the sponsor come to the conclusion that the evidence was so compelling as to take the step of voluntarily removing the drug product from the market?

DR. BRAUNSTEIN: Well, at the time when we saw the increased risk compared to placebo there were not data to allow us to conclude that this could be a class effect, and we felt that there were other options available to patients, including therapies that adverse event not known to have this increased cardiovascular risk. So, given those options and alternatives, we felt that the responsible action at the time was to withdraw Vioxx.

DR. CRAWFORD: Excuse me, but I am asking specifically what was that signal that was at the level where, in the interest of caution or whatever the mechanism was, you said this level is unacceptable at this time based on the given evidence?

DR. BRAUNSTEIN: Well, we saw overall a two-fold increased risk and that was seen versus placebo so it was something that we knew was statistically significant. The magnitude of the risk was on the order I think of one or two percentage points, but still at the time the other agents--it was a determination that amongst the choices that patients had available to them there were other agents that were not known to have this risk and, given the ability for patients to have alternatives that they could discuss with their physicians, we felt that we should withdraw Vioxx at that time.

DR. WOOD: Dr. Shafer?

DR. SHAFER: Two questions. I will make them fast. Do any studies show improved analgesia on Vioxx?

DR. BRAUNSTEIN: No. I mean, all of our efficacy studies show very similar results at comparable doses to NSAIDs.

DR. SHAFER: Okay. The other thing is can you go to slide number 36?

DR. BRAUNSTEIN: Yes?

DR. SHAFER: I just can't help but notice, but the upper bounds of the confidence intervals for the first two groups encompass the mean of the naproxen comparison. Does that give you pause in justifying excluding naproxen as a separate comparison group? If you take a look at the upper bounds, they include the mean of naproxen which might suggest that statistically those groups really shouldn't be segregated as you have done.

DR. BRAUNSTEIN: Well, when you look at this, if you were to combine all the data one would not see a statistically significant difference. It would tend to obscure the naproxen finding, and we felt, given what we observed in VIGOR and what we had observed all along the program, that that wasn't the right way to go, especially given the difference pharmacologically. I mean, in terms of looking at the data we also were taking into context what we understood about the pharmacology of these agents and the ability for naproxen to provide that kind of inhibition of COX-1 that Dr. FitzGerald talked about. So, we thought that not only were there differences in the clinical data but there were differences in the pharmacology data that supported keeping naproxen separate.

DR. WOOD: Dr. Gibofsky?

DR. GIBOFSKY: One of the stratifications we are asked to do during the next three days is, of course, the

risk/benefit relationship. I am wondering if you have calculated the risk/benefit of cardiovascular thrombosis outcome versus the benefit of cancer prevention in the population. I can understand where the relative risk of 1.92 is. I understand what it means when the relative risk goes up above a certain number above 1.0 but, you know, you can't go much below 1.0. So, have you calculated to what extent your risk of cardiovascular events is related to your protection against cancer?

DR. BRAUNSTEIN: Well, we didn't actually study cancer as an outcome. We were looking at polyps which are precancerous lesions.

DR. GIBOFSKY: The same question basically.

DR. BRAUNSTEIN: Well, even there, you know, polyps are easily--there is a different mechanism. There is an alternative therapy available for the treatment of polyps. So, in order to evaluate risks and benefits one has to compare the risks and benefits of one treatment option versus the risks and benefits of another treatment option. In doing so, I think that this wouldn't--

DR. GIBOFSKY: Well, let me ask it another way then, if you did not see a cardiovascular signal in APPROVe would you have concluded that the reduction in risk in polyp formation was efficacious?

DR. BRAUNSTEIN: We concluded that the reduction in risk in polyp formation was efficacious regardless of the cardiovascular finding. Are you asking whether the overall risk/benefit would have been favorable???

DR. GIBOFSKY: Yes.

DR. BRAUNSTEIN: That would be speculative for me. We haven't looked at the data with that specific question in mind. I think we would need to take a look at all the patients that we looked at in all the different subgroups to see if that remained the case. You know, you saw some congestive heart failure. We say NSAID type typical effects that one would see in one of these studies, not just cardiovascular risk but there was a small increase in ulcers, not as much as one would anticipate to see with a non-selective NSAID but still present. There was a small increase in other NSAID type effects like edema and hypersensitivity. So, we haven't made a formal risk/benefit assessment.

DR. GIBOFSKY: Just one last point, you stressed the concept of their being other modes of therapy available and so that factored into your decision to take this agent off the market. But there are other ways of treating polyps as well, which leads me to question in that context the rationale for the APPROVE study.

DR. BRAUNSTEIN: We thought this was an interesting and important scientific question that had been raised in the literature.

DR. WOOD: That sounds like a retrospective question so I will let you off the hook. Let's move on. Ralph?

DR. D'AGOSTINO: Two quick questions. In slide 48 you, I think quite sensibly and again post hoc, split out the cardiovascular risk and redid the analysis. Now, if this were replanned

and I got a result like that I would say that this is great; this shows me that placebo is better no matter what I do. I mean, the cardiovascular does increase a bit but the placebo is still maintaining itself even in individuals without cardiovascular risk.

DR. BRAUNSTEIN: This slide shows the relative risks in each of these groups. It is not placebo and then rofecoxib.

DR. D'AGOSTINO: Well, it is all against placebo.

DR. BRAUNSTEIN: It is all compared to placebo, yes.

DR. D'AGOSTINO: Right, and placebo wins everywhere. So, no matter if you have cardiovascular risks or not, still placebo was better. Am I misinterpreting this?

DR. BRAUNSTEIN: You know, in this we only see trends for some subgroups and in others we don't identify particular subgroup factors where there is an important difference.

DR. D'AGOSTINO: Well, that is a subgroup and it sort of indicates consistency to me. In slide 42 there was consistency regardless of CV risk. In slide 42, if I look at those numbers on the bottom, I presume those are individuals available. You are dropping about 100 individuals after 12 months or so. Do we know anything about the loss to follow-up on these individuals?

DR. BRAUNSTEIN: We did not see differences, for example, in changes in cardiovascular risk associated with patients who discontinued--

DR. WOOD: Wait a minute, these are not all patients who dropped out, are they?

DR. BRAUNSTEIN: These are all the patients who remained in the study.

DR. WOOD: So, some of these patients may not have advanced to the end of the study.

DR. D'AGOSTINO: Well, if you start at the beginning--that is my question, I mean it is randomized, right? So, there must have been about a 50-50 break so you would think at each point you would have approximately the same numbers in the two groups.

DR. BRAUNSTEIN: Well, there is a differential dropout due to adverse experiences for example that one would normally see in an NSAID trial against placebo.

DR. D'AGOSTINO: Well, why couldn't they be followed for CV events? Why wasn't it like an intent-to-treat analysis or something?

DR. BRAUNSTEIN: Yes, the way we had prespecified the analysis was that all events were determined up to 14 days after discontinuing therapy. The only intention-to-treat analysis was one done for mortality overall.

DR. WOOD: Dr. Nissen?

DR. NISSEN: Yes, I think Ralph's point is very, very important. We need to see an intent-to-treat analysis. You are telling me that 14 days after they dropped out of the study these folks were not followed beyond that?

DR. BRAUNSTEIN: We are following patients who are off-drug, who terminated treatment in the study, and we don't have data yet on that.

DR. NISSEN: Because there are a lot more people dropping out of the rofecoxib arm and the question is why are they dropping out and what happened to them. The signal here could be a lot stronger than we see using this somewhat selective analysis. I am used to an intent-to-treat analysis, Ralph, for a trial like this and I am confused as to why it was done in this way. You know, a cardiovascular hazard, if this is a pro-atherogenic therapy, is going to persist quite a while after you stop the drug. So, I think we really do need to see--I mean, to clear the air here we have to see that intent-to-treat analysis. I would track those people down and find out what happened to them. As a cardiologist, I obviously use a lot of low dose aspirin so I am very familiar with the low dose aspirin literature, and we see in low dose aspirin perhaps up to a 20 percent reduction in cardiovascular risk in individuals who are at risk. So, what I am really confused about is that you attributed what you found in VIGOR to the beneficial effects of naproxen, but you are talking about a 4- or 5-fold difference in myocardial infarction rates and I just want to know how you came to the conclusion that that amount of difference could be explained by naproxen. Naproxen would have to be a lot more effective than aspirin. We know aspirin inhibits platelets as well as anything else out there. So, how did you guys arrive at that conclusion that it was naproxen related?

DR. BRAUNSTEIN: Well, other than in addition to the data that support that

naproxen can have this effect, and specifically with regard to the magnitude that you are pointing out in myocardial infarction, there were only 24 events in VIGOR. The cardiovascular outcome studies that you are referring to oftentimes have hundreds, if not thousands, of events that they are assessing and that allows one to very carefully and with precision identify what the relative risk reduction is. In VIGOR we had fairly wide confidence intervals and, in addition, VIGOR studied exclusively patients with rheumatoid arthritis. These are patients with chronic inflammatory disease, elevated C-reactive protein and in those patients we know that the effect of aspirin is also magnified. So, given those factors, we felt that it was certainly compatible with an aspirin-like effect.

DR. NISSEN: Again, I am not sure I buy that. You know, post-MI patients have a very elevated risk and the most we ever expect from aspirin is perhaps a 20 percent reduction in recurrent events. Even with dual platelet antagonism with aspirin and clopidogrel we don't get a whole lot more than that. So, this story about naproxen, as I think Garret FitzGerald amply discussed--it doesn't stand the test of any kind of scientific rigor. I guess the other question I wanted to challenge you on is this comment that you made that the blood pressure effects in APPROVe were consistent with what is seen in other NSAIDs. I hope many of you have had a chance to look at the Archives manuscript that compares a meta-analysis of blood pressure effects. It sure looks like rofecoxib is an outlier here, showing a weighted mean difference of about 5.5 mm Hg or almost 6 mm Hg compared to NSAIDs which are substantially smaller. Is it your

position that rofecoxib does not produce greater degrees of hypertension than comparable NSAIDs?

DR. BRAUNSTEIN: Most of the studies that are referenced in that analysis, unfortunately, are confounded by dose. We think it is very important when one looks at a pharmacologically mediated effect, especially one that is known to have a dose-dependent association, that the drugs be assayed at doses that provide pharmacologically equivalent degrees of inhibition of COX-2. For example, for rofecoxib and celecoxib that would be 25 mg of rofecoxib and 200 mg twice a day of celecoxib.

DR. NISSEN: Okay. I want to clear the air on one more thing and, obviously, this drug has been the subject of a great deal of public attention and I think it would be a great opportunity for you to explain, from your perspective, why did it take 14 months, from February of 2001 to April of 2002, for the label to change? Were you fighting the FDA? Was there a big battle over what the wording ought to be of the label? I mean, it seems like 14 months is an awfully long time after an advisory committee meeting that recommended a warning to take for agreement to be reached about what that warning ought to say.

DR. BRAUNSTEIN: The advisory committee--

DR. WOOD: I think that is something probably we should let him pass on--unless you want to... Go ahead.

DR. BRAUNSTEIN: No, no, no.

DR. WOOD: Go ahead.

DR. BRAUNSTEIN: After the advisory committee there were a lot of discussions with FDA. There were data requests from them which we provided to them. We submitted at that same time the NDA supplement for rheumatoid arthritis because we felt it was important. As you know, VIGOR had been conducted in rheumatoid arthritis patients at 50 mg and it was important to communicate to physicians that the appropriate dose in those patients was 25 mg. So, there was a lot of information for the FDA to review. They also asked for updated analyses of all our safety data. So, they had a lot of work cut out ahead of them, and we worked diligently with them to provide the information, conduct the analyses that they requested, and collaborated in that way to make sure they had that information, and then we worked assiduously to conclude a label. So, I don't think, considering the wealth of information, that the time frame is unusual.

DR. WOOD: And after 14 months, it was "take the tablets slowly."

DR. BRAUNSTEIN: Well, after 14 months the advice was that cardiovascular risk factors, cardiovascular history should be taken into account--

DR. WOOD: Well, that is not what it said. It is most important to remember it didn't say you shouldn't give it to people with cardiovascular risk factors. It didn't say it shouldn't be given to people who had had an MI or any other expletive statement like that. It said caution should be exercised in patients with history of heart disease. That is quite different.

DR. BRAUNSTEIN: What I tried to say or at least what I was trying to communicate was that the risk/benefit assessment we felt needs to be done on a patient by patient basis and, in addition to taking GI risk into account, one should also take cardiovascular risk into account given the uncertainty of the data that was available at that time and, as the label said, the clinical significance of these cardiovascular findings were unknown and that, therefore, the cardiovascular information should be taken into account when considering the use of rofecoxib.

DR. WOOD: Dr. Hennekens?

DR. HENNEKENS: I would be interested in knowing the total number of deaths in the randomized trials of rofecoxib against all other comparators and then against placebo, non-naproxen NSAIDs and naproxen.

DR. BRAUNSTEIN: You have that on your slide. The numbers of deaths are underneath the rows. I don't have the numbers at the top of my head. We would have to do a quick tally. Also, the only problem with looking at the numbers is that the numbers themselves don't take into account imbalances in exposure, which is why we showed them as rates per 100 patient-years because it certainly takes into account the differences in exposures. Compared to the NSAIDs we did not see differences in the rates, and compared to placebo we did not see differences except, as I pointed out, in the Alzheimer's disease study where there was a statistically significant higher rate with rofecoxib.

DR. WOOD: Dr. Cannon?

DR. CANNON: You mentioned in the VIGOR and APPROVe clinical trials that the major driver for the increased cardiovascular events on rofecoxib was acute myocardial infarction. My question is were these myocardial infarctions apparently random events or was there any setting in which they seemed to occur more frequently? For example, in relationship to a procedure, including a coronary interventional procedure, or surgery, or were the myocardial infarctions random events? I am thinking in terms of Dr. FitzGerald's presentation and the recent valdecoxib experience with bypass surgery.

DR. BRAUNSTEIN: We haven't identified any kind of associations such as you are asking. But I am not sure that we have specifically looked at the question the way you are asking. So, I am not 100 percent sure.

DR. WOOD: Dr. Abramson?

DR. ABRAMSON: Yes, I guess one of the surprises or unexpected findings in APPROVe was that it took 18 months for these curves to separate with rofecoxib. I was unaware of the heart failure and pulmonary edema data until this morning. Often fluid retention occurs early in the course of putting people on NSAIDs. So, I am wondering could you tell us more about when those heart failures occurred over the course of time. Were they early events, or was this also something that took some time to appear in the population?

DR. BRAUNSTEIN: As one would expect from an NSAID, fluid retention, and heart failure were early events. If you look at discontinuations for example due to edema-related adverse

experiences, including heart failure, patients tended to discontinue--if they were going to discontinue, they discontinued early and then the two groups continued in parallel. But, yes, it was an early finding as you would expect.

DR. WOOD: Tom?

DR. FLEMING: Could you show us the curves back from the VIGOR trial that looks at complicated confirmed upper GI?

DR. BRAUNSTEIN: Complicated confirmed upper GI?

DR. FLEMING: Correct.

DR. BRAUNSTEIN: We don't have those.

DR. FLEMING: You just quickly referred in your presentation to the results being positive.

DR. BRAUNSTEIN: The results were that the two curves showed the same V-like difference and they continued to separate over time. I am just looking here and apparently we don't have that slide.

DR. FLEMING: You showed us the confirmed upper GI and those cumulated to rates of 4.5 against 2.1. The data we have been provided separately for the complicated confirmed upper GI are 1.4 against 0.6. So, it is the same relative risk but a much less frequent event.

DR. BRAUNSTEIN: Sure, yes, and those were mostly GI bleeds.

DR. FLEMING: I was just curious to see a pattern as to whether that is, in fact, cumulatively increasing or more apparent early in time. Let me go on to the next point. That reflects approximately numerically almost exactly the same number of prevented cases of complicated confirmed upper GI as there were excess numbers of thrombotic cardiovascular SAEs. In essence, what you have said is that the analgesia was comparable. So, essentially what we are really looking at is relative safety profiles and the goal here is to reduce the upper GI. And, we are essentially preventing an equal number of upper GI complicated events for equivalent numbers of excess events in the thrombolytic cardiovascular arena. Yet, essentially I think you were saying the latter didn't seem as established yet numerically it was the same. There were also in the trial excess numbers of deaths of 22/15 and when you presented the Alzheimer's data you gave us I think slide 35 that indicated that when you looked at the Kaplan-Meier curves for confirmed thrombolytic cardiovascular events that didn't seem to reinforce the excess rates that you were seeing with VIGOR and, yet, it did reinforce the excess mortality as you have now circled back and reported at the end. In 2003 the excess mortality is quite significant but it was also significant in 2001. The latter date is in Tab G, page 2 but the former data is in Tab F, page 39 where excess mortality was significant at 33/20 and the cardiovascular were 8 to 4. So, you were seeing from these two sources excess mortality and you were seeing excess numbers of thrombolytic events that were equivalent in number to the number of prevented complicated confirmed upper GI events. Am I correct on this summary?

DR. BRAUNSTEIN: Well, no. There are a couple of points I would disagree with. First, in VIGOR the difference in mortality was not statistically significant and also in terms of looking at the causes of death, cardiovascular mortality which is the difference we would see was not different between the two groups. There were 7 on rofecoxib and 6 versus naproxen. So, I am not sure--

DR. FLEMING: Well, I don't think we disagreed. I am not talking about statistical significance here. I am talking about what the data are actually suggesting in what is available--

DR. BRAUNSTEIN: Well, I must say there is a lot of data that you pointed out to me and--

DR. FLEMING: Well, just to summarize the essence, while you have emphasized appropriately the upper GI events being decreased, when you look at the actual number prevented in complicated confirmed upper GI it is numerically almost identical to the number of excess thrombolytic cardiovascular SAEs that were seen in VIGOR. You also saw a numerical increase of a relative risk of 1.5 on mortality, which was also seen in the Alzheimer's study which you were saying at the time was contradicting the sense of concern related to the overall thrombolytic excesses. And, what you were seeing at the time, even back in 2001, was a statistically significant excess in death rates with a doubling in cardiovascular-related deaths.

DR. BRAUNSTEIN: Let me ask Dr. Reicin because she perhaps has a better handle on it and I am sort of getting lost in the mass of data that is coming up.

DR. REICIN: I think there are two issues that I think you brought up.

DR. WOOD: Sorry, just for the record, can you identify yourself?

DR. REICIN: I am Dr. Alise Reicin, Vice President of Merck Research Labs. In terms of looking at VIGOR, I think you are correct. There was excess in cardiovascular events on Vioxx and there was a decrease in the complicated GI events on naproxen.

DR. FLEMING: Which numerically were almost identical.

DR. REICIN: And I think that that is also fair to say. If you compare our data versus diclofenac and ibuprofen at the time, there was no difference in cardiovascular events. In fact, numerically it was in favor of Vioxx and, yet, there was a significant reduction in GI events. So, that takes care of that. So, versus naproxen, I think you are right, there was excess in CV, lower in GI versus ibuprofen and diclofenac, however, no evidence of an increase in CV and a reduction in GI. In terms of the mortality data that we had at the time, we had a significant reduction in mortality on Vioxx versus non-naproxen and the NSAIDs that we had in our Phase III OA studies, and at the time we actually did not make a lot of those. We thought it was potentially by chance. That was actually driven by CV mortality in the non-naproxen group. In VIGOR there was a numeric imbalance, 22 to 15 in deaths, but cardiovascular mortality was similar. In terms of Alzheimer's I don't think there was statistical significance back at the time of VIGOR. There was a numeric

imbalance. In terms of cardiovascular I think the numbers were 8 versus 4. They were put in the label. So, pretty small numbers. The rest of the difference that we saw was due to things like poisoning, electrocution and other things that we thought were not drug related.

DR. FLEMING: You are correct, it was 8 versus 4 in cardiovascular related deaths, but it was statistically significant in total mortality at that time as well. It was 33 against 20, with p values reported, depending on the method, of 0.007 to 0.26. Now, the final data are significant but even the early data were significant and reflected the level of excess mortality that VIGOR was establishing but not in a significant fashion.

DR. REICIN: Again, we didn't see it though in any of our other data sets. In fact, in the early data sets statistically it went the other way, non-naproxens had higher one. I think you can see that in RA also there was no evidence of an excess. In ADVANTAGE there was no evidence of an excess. You see now in ViP and--

DR. FLEMING: But there was in ADVANTAGE. There was an excess.

DR. REICIN: Not in overall mortality.

DR. FLEMING: Yes, in overall mortality--oh, I am sorry, in Alzheimer's.

DR. WOOD: Tom, have you finished?

DR. FLEMING: Yes.

DR. WOOD: Dr. Shapiro?

DR. SHAPIRO: I guess I want to follow up on a comment that you, Dr. Chair, made. I am still concerned about the label change and how helpful or not helpful it was, not only because it may not have been as helpful as it might have been to clinicians but also to patients in the informed consent conversation. What else was made available or should have been made available or could have been made available to clinicians to make some sense out of this, caution should be exercised when Vioxx is used in patients with a medical history of ischemic heart disease?

DR. BRAUNSTEIN: Were you addressing me or the Chairman? Me? What we made available were the data. I mean, I think that is the answer to the question in terms of the labeling and in terms of what we had published.

DR. SHAFER: So, the VIGOR and Alzheimer's results were made available. You just weren't going to analyze them to make any more definitive statements at that time about what clinicians should take away?

DR. BRAUNSTEIN: Well, by 2002 we were also starting to implement our outcome study. We thought the important message to clinicians was that there is a GI benefit and there is also a cardiovascular finding that we don't understand given the differences between the two data sets. It says the clinical significance was unknown and that this information needs to be taken into consideration when assessing the risks and benefits of these drugs in individual patients. Individual patients differ in terms of their risk profiles and that decision on which drug to be used is best made on a patient by patient basis.

DR. WOOD: Dr. Ilowite?

DR. ILOWITE: Rofecoxib was pulled from the market approximately three weeks after its approval in children with juvenile rheumatoid arthritis. I have two quick questions. Were there any cardiovascular events in any of the trials in children?

DR. BRAUNSTEIN: No.

DR. ILOWITE: Second, did you give any consideration to the fact that there were no other COX-2 inhibitors, other than one NSAID that was available as a liquid, before you made the decision to pull it from the market?

DR. BRAUNSTEIN: The focus I think was on the list we had seen versus placebo in the adult patients. This kind of disease, cardiovascular disease, is not very common in children and we hadn't seen anything like that in our population.

DR. WOOD: Dr. Boulware?

DR. BOULWARE: I want to go back to the previous question. What I heard was a discussion about an offset between complicated GI events and it sounds like non-fatal MIs. If I understood the discussion back and forth here, they are roughly comparable. Now, in patients requiring an NSAID, and I am not talking about the APPROVe data here but in patients requiring NSAID treatment if there is roughly comparability of complicated gastrointestinal events with non-fatal MIs, it sounds like Merck's thinking was that the risk of a non-fatal MI far outweighs, in a patient requiring NSAID treatment, the risk of complicated GI

events and that that was what drove the decision. The reason I am interested in this is that obviously this meeting is entirely about how you make a risk/benefit calculation. So, your thoughts in September about this issue are I think helpful to us in thinking about these risk/benefit issues.

DR. BRAUNSTEIN: I wouldn't put it exactly the way you stated it, and that is because the individual patients at risk for these problems differ and there were alternative approaches for patients with GI risk that were available at the time. Now, we recognize that rofecoxib had met the highest standard. Well, yes, it had met the highest standard but there were alternatives available and we did not have data on what one could do for more studies. The data was unclear as to the mechanism so we felt that given those options, the withdrawal made the most sense.

DR. BOULWARE: Can I just make a little follow-up comment? It sounds like you are trying to have your cake and eat it too. On the one hand, you would have liked to have said pre-September that rofecoxib was the only COX-2 selective drug that had demonstrated effect in reducing GI toxicity. Now you are saying, after you pulled it from the market, there are lots of other alternatives that are almost just as good. I don't really understand.

DR. BRAUNSTEIN: I couldn't say "almost just." There haven't been head-to-head studies to answer that latter part of your question. There were alternatives. We did not know that there is a class effect for cardiovascular.

DR. WOOD: Dr. Manzi?

DR. MANZI: This question actually may better be answered by Dr. FitzGerald, I am not sure--is he here?

DR. WOOD: Here he comes, just in time.

DR. MANZI: He eloquently pointed out that there is clearly variability in individual dose response with regard to COX-2 inhibition. Since we are grappling with this issue of class effect versus a specific drug effect, is it feasible or helpful to look at the degree of COX-2 inhibition in association with these events?

DR. WOOD: You are up, Garret. Just take that microphone.

DR. FITZGERALD: I would say yes amongst all those things.

DR. WOOD: Amongst all those things? I don't understand.

DR. FITZGERALD: I mean one of the issues that you would hypothesize is relevant to outcome is the degree of selectivity attained in an individual.

DR. WOOD: You mean amongst other things related to the drug?

DR. FITZGERALD: Amongst other things related to the drug and underlying--

DR. WOOD: Sure. Dr. Platt?

DR. PLATT: Compared to other NSAIDs, do I understand properly that 98 out of 100 patients who take the drug would have about the same outcome? That is, the significant difference

between the regimens is approximately-- 2-fold means about a 2 percent absolute difference.

DR. BRAUNSTEIN: Which outcome are you referring to?

DR. PLATT: To the GI outcomes.

DR. BRAUNSTEIN: There is a range. There is a small range because it does seem that we have a larger difference-- you know, if you line them up it is a little larger with naproxen and a little smaller with diclofenac but I would say on average it is about two-fold.

DR. PLATT: Right, but that 2-fold translates into about two patients out of 100 having a different outcome than they would have if they had taken the comparator. I am trying to get at the question of whether we can identify those two patients with greater certainty than just treating everyone. And, I would ask the same question about the cardiovascular complications. That is, in this complicated business of risks and benefits, can we do better than we have at guiding both clinicians and their patients in having at least semi-quantitative estimates of what the risks will be and what the benefits will be so they can make an informed judgment?

DR. BRAUNSTEIN: We know that from the VIGOR results because we looked at patients with different baseline risk for GI disease, and this is something that is well understood, what the different risk factors are for GI disease, including things like prior history of a GI event, and we saw the same 50 percent reduction across all the different risk factors. In terms of cardiovascular, we are still introduction he process of

trying to see if we can identify particular risk factors that would correlate. So, that is still an open question based on our data.

DR. PLATT: But saying 50 percent really obscures the fact. Some people may have a baseline risk of a serious GI event of 20 percent or 30 percent, in which case 2-fold is a very big improvement for them--

DR. BRAUNSTEIN: Yes, of course.

DR. PLATT: If we knew enough we would know that most people have effectively a zero risk. So, there is very little benefit for them. Have you put the data together in a way that helps us identify the people who stand most to benefit and the people who stand most at risk, and is it possible that those are different groups?

DR. BRAUNSTEIN: Dr. Reicin can I think provide more information on the VIGOR results because she was involved extensively in the VIGOR study.

DR. REICIN: Dr. Laine may come up to help me if I don't remember something. We actually published a paper on looking at specific subgroups in the VIGOR study. What we found is very similar to what Byron talked about during his discussion. Patients with typical risk factors, age more than 65-- do you want to add something?

DR. LAINE: I agree absolutely. The reason I actually took these data and published this paper with the VIGOR results is that I have had the same idea. Relative risk isn't important in practice; it is the absolute change, the number

needed to treat. So, we looked at that with absolute incidence of number needed to treat and for clinical events, for instance, if you had a prior event you only have to treat ten people for one additional event. But if you don't have a prior event you have to treat, let's say, 60 or 70. The same with age, if you are over 75 you only need to treat ten people for one additional event. But if you are under 65 you need to treat 50 or 60. So, I agree absolutely that at least with the VIGOR data, we stratified by these different clinical risk factors that Byron showed earlier.

DR. WOOD: We have three more questions. Dr. Shafer, Dr. Cush and then Dr. Temple.

DR. SHAFER: Two questions. Can you go to slide 48?

DR. BRAUNSTEIN: That is the subgroups, yes?

DR. SHAFER: Yes, is the one on various subgroup analyses. Can we show the slides? Just to highlight what the question is, in slide 48, this is following on the comment by

Dr. Nissen regarding the aspirin use, what you show in the APPROVe trial is that the risk factor for those with aspirin on board, in fact, is 3.25 with a confidence interval which is wide, as Dr. FitzGerald has suggested it might be because of small numbers, but it goes from 0.98 to 13.81. Now, the hypothesis behind VIGOR and interpreting VIGOR as an aspirin-like effect, was that aspirin was going to confer safety. Doesn't the data on slide 48 essentially disprove the naproxen hypothesis in VIGOR?

DR. BRAUNSTEIN: No, there is no naproxen in the study--

DR. SHAFER: Right, but the hypothesis was that naproxen was acting like aspirin.

DR. BRAUNSTEIN: Yes.

DR. SHAFER: Yet, here in the presence of aspirin to provide the safety, you are not seeing benefit.

DR. BRAUNSTEIN: I would argue that the mechanism for what we saw in VIGOR, which was a very early difference between the two treatment groups, is qualitatively very different than what we see in APPROVe. So the mechanism for the cardiovascular difference in the two studies is not necessarily the same and, therefore, whatever difference we are seeing here or not seeing with aspirin doesn't really relate to what we saw in VIGOR. I would also point out, as you have already pointed out, there are wide subgroups. I think Dr. Villalba has pointed out that when we looked at the APTC endpoint, which was just myocardial infarction, stroke and vascular death, the difference actually seems to go away but, again, there are very small numbers and we don't want to over-interpret at this point what the data say.

DR. WOOD: But the major point here, just to help you here, is that these people were not randomized to aspirin. So, people who were on aspirin were a different subset than the people who were not on aspirin in terms of cardiovascular risk and so on. So, it is not like naproxen.

DR. BRAUNSTEIN: Yes. Yes, of course. Sure.

DR. WOOD: The one thing I would say while you have that slide on there is that I think is going to be important for us is that our job is not to identify groups that are at particular risk, Richard. Our job I think is to see if we can identify patients who are at low risk--

DR. PLATT: Yes.

DR. WOOD: I am not arguing with you. I am just making a generic point and it is not clear to me that there is such a group identified there.

DR. PLATT: Well, it seems to me that there will always be risk--

DR. WOOD: Right.

DR. PLATT: --the question is can we help inform decisions that patients have to make?

DR. WOOD: Dr. Cush?

DR. CUSH: Dr. Braunstein, a few times you mentioned that you made this decision based on the signal that you found in the alternatives existing, and not knowing if it is a class effect. If you knew that this was a class effect would you have made the same decision? And, knowing what your COX-2 potency is, does that factor into that?

DR. BRAUNSTEIN: I couldn't go back and speculate what decision we would have made based on a different set of data.

DR. WOOD: I think that is a fair answer. Let's move on to Dr. Cryer.

DR. CRYER: I would like to come back to a consideration of the potential gastrointestinal benefits of COX inhibitors and specifically Vioxx, and I am going to use your slide 33 to help me with my questions and comments. You repeatedly made the point that Vioxx, rofecoxib, was unique in its labeling with respect to its gastrointestinal benefit and that was a label revision that was largely derived from a discussion of the data in the VIGOR trial in which naproxen was the comparator. I want to underscore that the conclusions reached may be as much of a reflection of the comparator as they could be a reflection of properties intrinsic to the COX-2 specific inhibitor. As I look at the pooled analyses from the rofecoxib experience and specifically look at diclofenac, it does not appear that the difference in reduction compared to diclofenac is statistically significantly different. So, the question that I have for you is do you think that the revisions in the label would have been the same with respect to the GI observations in VIGOR had diclofenac been the comparator rather than naproxen?

DR. BRAUNSTEIN: In an adequately powered study. I think the failure here in these confirmed events, in order to have the confidence interval narrow enough we would need enough power to do that. In fact, when we looked at investigator reports of these events, in all, including the unconfirmed, we did have statistical significance. So, I think that, yes, in an adequately powered study we would show a difference from diclofenac.

DR. WOOD: Bob?

DR. TEMPLE: Actually, I wanted to pursue something Dr. Shafer raised. The aspirin subgroup is a baseline subset. People are probably reasonably well randomized to whether they get-

DR. WOOD: They didn't get aspirin.

DR. TEMPLE: No, I know. They were different populations from people who were on aspirin but they are randomized to the two treatments, and there are about a thousand of them. From everything that I would have understood from Dr. FitzGerald's talk, when you are on both aspirin and rofecoxib you are not on a selective drug anymore, or probably not because you have plenty of COX-1 inhibition. But the hazard ratio there is higher than the other people. I wonder whether that is easily explained, or it could be explained by blood pressure effects which, of course, aspirin will not reverse. Because I think it needs some kind of explanation.

DR. WOOD: So, is that addressed to Garret?

DR. TEMPLE: Either.

DR. BRAUNSTEIN: With regard to aspirin data, they are not robust enough. We are talking about a total of 11 events, as I recall, in that analysis for the APTC. There are not a lot of events in that analysis.

DR. TEMPLE: There were 16.

DR. BRAUNSTEIN: Right, 16 events. There are very wide confidence intervals, as you know. So, I think it is difficult to draw specific conclusions about aspirin. With regard to blood pressure, as I indicated, when we looked

at that the blood pressure changes that we observed would not appear to explain the magnitude of the cardiovascular findings that we observed in APPROVe.

DR. TEMPLE: One of the reasons to worry is that people with underlying heart disease or diabetes are probably more sensitive to blood pressure effects. There is some evidence of that. Anyway, just a thought.

DR. WOOD: Garret?

DR. FITZGERALD: I would just say one can over-parse extraordinarily small amounts of data in retrospect, and that there is enough flexibility in what one would expect to see to account for that. For example, we don't actually know if inhibition of COX-1 has no impact on the blood pressure response to a COX-2 inhibitor. In fact, from what I showed you in mice, one would anticipate if one actually designed a study to address that question that the answer would be yes. So, I think that, coupled with the fact that aspirin, even if one had loads of data, would be expected to modulate rather than abolish the hazard through this mechanism really means that it is not an answered question rather than an answered one.

DR. WOOD: Great! Well, let's stop at this point and break for lunch. We will restart at exactly one o'clock.

(Lunch recess.) A F T E R N O O N P R O C E E D I N G S

DR. WOOD: Merck has a couple of slides they wanted to show to address the blood pressure issue that came up in the previous discussions. So, let's go ahead and do that first quickly.

DR. REICIN: The first was in relation to the issue about congestive heart failure, which is a known side effect of all NSAIDs and COX-2 inhibitors and is reflected in their labeling. Since the only data we showed was from APPROVe, we had an expected difference from placebo but if you look at this slide you see that in our OA database the incidence of congestive heart failure was low, and it was generally similar to ibuprofen. You can see that it ranged from 0.1 to 0.4 percent on rofecoxib; 0.4 percent on ibuprofen; and 0.8 percent on diclofenac--so, generally similar to the NSAID comparators. I will acknowledge that there was one epidemiologic study that suggested that the rate was higher on rofecoxib.

DR. WOOD: But the data in the APPROVe study are up to 1.5 in the rofecoxib group, and this is for serious heart failure--congestive heart failure, pulmonary edema. Right?

DR. REICIN: It was versus placebo. The rate was higher in that study than we have seen in other studies.

DR. WOOD: Right. But it is not a question of whether it is against placebo or not. The underlying rate is much higher.

DR. REICIN: The rate was higher in that study. We didn't see it as high in our other studies. The other slide, 232--it was a question about whether rofecoxib had effects on blood pressure that were very different than the other NSAIDs. This was a study done in elderly patients. It was not an ambulatory blood pressure study but blood pressure was measured in these patients 4 times a day. If you look, rofecoxib 25 mg was

compared to celecoxib 200 mg BID. That is the highest recommended chronic dose for both of these medications. The medications at that dose had similar inhibition of COX-2. We compared it also with naproxen 500 BID and placebo, and I think you can see that the changes in systolic blood pressure and diastolic blood pressure are similar among the active treatments and greater than placebo.

DR. WOOD: Okay. Thanks very much. These are helpful comments. Are there any questions specifically and only on these two things? Steve?

DR. NISSEN: Could you show us the use of antihypertensive agents in the two arms of APPROVe? I would be interested in seeing if there was a difference in use of antihypertensive drugs. I am also interested--you know, these mean changes are useful but it is also useful to know the fraction of patients that had sustained increases of, say, 15 mm or more because that is the kind of level of increase that would constitute a substantial risk. So, I am interested in use of antihypertensive drugs and I am interested in the number of people who had greater than a 15 mm sustained increase in each arm.

DR. NORGAN: Kevin Norgan, Merck. The use of antihypertensive drugs in the APPROVe study, at baseline it was approximately 30 percent. It was 30 percent in one treatment group and 29 percent in the other treatment group. Then, during the course of the study the numbers increased to approximately 40 percent in the rofecoxib group and approximately 35 percent in the placebo group. The actual numbers are in the publication that is on the Internet.

DR. NISSEN: And was that difference statistically significant?

DR. NORGAN: I don't recall. I think it was but we would have to check.

DR. WOOD: Then 25 patients dropped out because of hypertension versus 7 in the placebo group. Right? So, that should be added to the number that actually ended up on antihypertensives in the APPROVe study.

DR. NISSEN: What about the issue of the 15 mm or greater? Do you have any data on that? Bob Temple, isn't that something you guys like to look at in the FDA, the sort of 15 mm outlier group?

DR. TEMPLE: I don't know. I think we look at mean just as often.

DR. NISSEN: All right, but I would like to know because I didn't see that.

DR. REICIN: We will get back to you later with that data, Dr. Nissen.

<Break for Dr. Villalba's FDA presentation on Vioxx>

DR. WOOD: Thank you very much. Could you go back three slides, and then I am sure Dr. Fleming will want to ask you a question?

DR. VILLALBA: Which one?

DR. WOOD: The third last slide in the handout. That one.

DR. VILLALBA: This one?

DR. WOOD: Yes. Am I right?

DR. FLEMING: You read my mind. I wanted to follow-up on this because it is also a follow-up to a question I asked this morning. Just to get a sense of what the totality of the data is telling us about whether there is an all-cause mortality risk increase, and the two studies on the left definitely strongly suggest that there is. In the discussion this morning it was pointed out that there are other sources of data that might complicate the interpretation, the ADVANTAGE trial being one of those. But if you look on sponsor slide 54, which we won't go back to now, the other studies are all very small relative to the numbers of events. More than a half of the total deaths in the meta-analysis of all the studies are from the VIGOR study and the Alzheimer's studies and that is where we are seeing the signal. The ADVANTAGE study that we were told about that didn't show significance still had one more death, and you said in your presentation it was 4 versus 0 in the wrong direction. There are 2 in the cardiovascular. I guess my concern here is that when I look at this it is on-drug, and I think it is getting back to a question Ralph was asking earlier today. All of these analyses, are we correct, are only giving us that deaths that occurred within--what?--30 days of being on drug?

DR. VILLALBA: Two weeks actually.

DR. D'AGOSTINO: Yes, I raised that morning. I mean, why weren't these individuals followed till the end of the study to find out about mortality?

DR. VILLALBA: Well, actually that is a good question to the sponsor because we know that they were followed as much as they could do it, but it was not

mandatory. They tried to collect all the data they could but it was actually--I would prefer them to answer.

DR. WOOD: Well, let's not involve motivation right now. Let's just keep going with the facts. So, Tom, keep going.

DR. FLEMING: Well, that is the essence that I wanted to get at. It was just to understand that this is just on-drug and there is nothing else you can provide us in terms of a true ITT? Is that correct?

DR. VILLALBA: There were more deaths also after but there was not a balanced exposure.

DR. WOOD: No, what he is asking is do you have an intention-to-treat analysis?

DR. FLEMING: Correct.

DR. VILLALBA: No, I don't have it with me. That is why I said this is still under review. There is pending information.

DR. WOOD: But before we leave this slide though, it is important to remember why we are here. I mean, this is a drug whose indication is a safety indication, and the reason to give the drug was to reduce an adverse event which is always thrown up as causing this terrible outcome, although the outcome has improved substantially over the last 10, 15 years. It is certainly worrisome when a drug that is supposed to produce a safety benefit, in fact, is producing an increase in mortality, it seems to me, and that is worthy of some discussion. Certainly, an ITT analysis would have been important.

DR. VILLALBA: Again, I completely agree. We are concerned, but we don't know how other NSAIDs would look here.

DR. WOOD: I understand.

DR. VILLALBA: We need to put it into context.

DR. WOOD: That is what my teenaged kids say as well. Curt?

DR. FURBERG: I was wondering whether you, within the agency, considered the risk of heart failure. I mean, when I look at the tables and in your presentation you are using the term heart arrest signal in a narrow sense. There is nothing in your tables on heart failure. It is an issue. As the Chairman found out a little bit earlier, in the APPROVe study, a 4-fold increase in a long-term trial. Do you have information from the Alzheimer trials on heart failure? If you look at the adverse effects of the drug, we shouldn't just narrow it to heart attacks and stroke. Let's broaden it to heart failure and make that part of our evaluation.

DR. VILLALBA: Yes, I don't have slides with me regarding congestive heart failure but, again, we don't have the data for other NSAIDs. That is the only thing that I can keep saying. But there was more heart failure, for example, in VIGOR clearly as compared to naproxen.

DR. WOOD: I sense that there is a response coming from the sponsor. Do you want a couple of minutes to think about that before you get up? You can take a couple of minutes and we will take another question, if you want. Take

your time; we won't forget you. Dr. Bathon?

DR. BATHON: I am a little confused about the aspirin issue. On your slide 35 you showed a decreased hazard ratio or relative risk for the aspirin users compared to non-aspirin users. But in Dr. Braunstein's presentation it was the opposite. I realize that the outcomes were measured a little bit differently.

DR. VILLALBA: That is a very good point. These are APTC endpoints and the way that Dr. Braunstein showed it was all cardiovascular/thrombotic events that included also peripheral events, unstable angina and TIA. So, the point of this slide is precisely that when we design a study that is going to address these issues in the best possible way we need to choose the right endpoint. And I don't know what that endpoint is because if you look at all cardiovascular events you may see more than if you look only at APTC.

DR. WOOD: Dr. Shafer?

DR. SHAFER: I know it is always easier in retrospect to try to make sense of things than prospectively when you are looking at many possible adverse outcomes and trying to figure out where to focus one's attention. But if you could go back to slide 23, what we see here, in slide 23, is a lot of suggestions of danger signals. Dr. Braunstein made an interesting point earlier when he said that it would take about 30,000 patients to demonstrate an increased risk, and yet we see danger signals in very small studies of short duration. So, that has obviously to be a cause for concern. Then along comes the VIGOR trial. As I understand, basically VIGOR had a 2-

5X increase in serious adverse cardiovascular events depending on the endpoint you chose to look at. Now, there are two possible interpretations of that. One interpretation was that rofecoxib increased risk. At the time you had this background worrisome signal rate which was consistent with the mechanisms that

Dr. FitzGerak spoke about, and if that were the true state of things, then potentially millions of patients were being placed at risk. The converse choice is that Naprosyn decreased risk. There were very weak data to support that. As we heard from Dr. Nissen, the effect was too large to be really explained by any known effect of aspirin. And, the safety data that were used to support the safety of rofecoxib was far less than the 30,000 patients that would be required to significantly show the difference. By Dr. Braunstein's own statements, you know, it would take far more patients to really statistically significantly show that up. What I first thought was the company and the FDA chose to give pretty good credence to the naproxen hypothesis. It sounds from the comments today that that is still the position of Merck. What would it have taken, what kind of data would it have taken, given the results of the VIGOR trial and the two alternative hypotheses, for the FDA at that point in time to either put a black box warning or perhaps even remove Vioxx from the market? What kind of data would you have had to have in addition to what you have?

DR. VILLALBA: I cannot answer that question. What I can tell you is that this was as compared to naproxen. We never bought the naproxen theory, but we also

did not have evidence that Vioxx was worse than placebo or other NSAIDs.

DR. SHAFER: You have great evidence in VIGOR though.

DR. VILLALBA: I completely agree but it was naproxen, and I think the presentation tomorrow with the epidemiologic data on naproxen will be very informative about how confused we are until today. Regarding the signals, yes, those were observed but that was after VIGOR, not before. Again, we have that long-term, placebo-controlled data in Alzheimer's patient elderly population that had shown no difference in myocardial infarctions or strokes. There was that signal of cardiovascular death that, by the way, was put in the label. But there were 8 versus 3 events and we didn't know what to make of that.

DR. KONSTAM: Hi, there. I am Marv Konstam. I am from Tufts University and I am here with Merck as a consultant. In 2001 I was first author on the overall pooled analysis for the entire rofecoxib database so I just think I want to speak to it, and the interpretation of VIGOR and where the company I think was, and the world was, at that point. I think it is really difficult to look at individual studies with very, very small numbers and find signals, and one can draw all kinds of conclusions from them; and there may be signals in the other direction in some of the other small studies. So, what was done at that time was, you know, there was a signal from the VIGOR study. This finding was unexpected. It showed an adverse effect on cardiovascular endpoints. Now, one thing I want to stress about that is that of all of the information that could be

brought to bear, I think the point estimate for the hazard ratio from that is probably the least important to me. You know, you are looking at very small numbers of events, unexpected finding, wide confidence intervals. So, I just want to point that out. What was done at that point was that the entire rofecoxib database to that point was reviewed in a systematic way, and all of the data were pooled. They were divided, as you heard, between Naprosyn comparator, other NSAID comparator but, most importantly, the placebo comparator. Because VIGOR was an active controlled study and none of us to this day know exactly to what extent the result was contributed to by an adverse effect of rofecoxib, a favorable effect of Naprosyn or a combination. So, the most valuable data are the placebo-controlled data. And, reviewing all of the placebo-controlled data to that point, pooling all of those data, there was 3000 patient-years of follow-up, there was not a hint of an adverse signal--not a hint of an adverse signal. Now, granted, there were confidence intervals around that signal so that is real. We still didn't know, and I think we know a lot more today thanks to the APPROVe study, but at that point in time if you look at all of the placebo-controlled data that existed there was not a hint of a problem, which I think led me at that time and I think led others at that time to say this may be contributed to by a significant beneficial effect of Naprosyn.

DR. WOOD: Just let me make sure I understand. Are you saying that that is still your position?

DR. KONSTAM: No, no. That was the position at that time. One might then ask, okay, what is different between the

APPROVe data, and I might say that I was on the data safety monitoring board for APPROVe, and why is APPROVe different than the pooled placebo-controlled at that time? I think that is a really cogent question to ask and I have asked myself that question. I believe the difference now, in retrospect, is exposure time. From APPROVe we see no evidence of a hazard in the thrombotic events through 18 months and then there is a separation. The median follow-up in that pooled analysis that I just referred to is relatively short. I don't know what it was exactly but it was months. It certainly wasn't the 9 months that was there in the VIGOR study or the 2.4 years in the APPROVe study. So, that is a substantial difference. There are other differences, but to me that may be the explanation for why the pooled analysis, back in 2001 and as it went forward, showed no problem but APPROVe then came and did show a problem. I think it probably was the exposure time.

DR. WOOD: But just to be absolutely clear, you are not saying that you still believe the VIGOR study was due to a totally protective effect of naproxen, are you?

DR. KONSTAM: No, no, I am not.

DR. WOOD: Good. I just wanted to be clear on that.

DR. SHAFER: While you are there, Dr. Konstam, in terms of the relative risks of the two possible choices--either rofecoxib increases risk or Naprosyn decreases risk--was that part of your thinking as well? What are the possible outcomes of the two competing

hypotheses? The truth is probably somewhere in between.

DR. KONSTAM: Well, first of all, let me just add one other point that I should have mentioned. The other point about the VIGOR study was the dose. So, there was a very high dose used in VIGOR and there were lower doses in the pooled analysis. APPROVe was 25 mg; an intermediate dose. What was your question again? I am sorry.

DR. SHAFER: There are somewhat different potential concerns with the conclusion that rofecoxib increases risk as opposed to the conclusion that Naprosyn decreases risk. Was that part of your decision analysis at the time?

DR. KONSTAM: Yes, thinking back at that time, there was no adverse signal from the placebo-controlled data. I don't think, you know, most people were completely satisfied with that. If you look back at what the company did at that point, first of all, there was a warning put on the label and we can argue whether that was good enough or not. But then we embarked on a large placebo-controlled program with a prespecified adjudication process for cardiovascular events, and that is the process that led to the definitive finding of APPROVe, even with a much smaller N than they were planning to do so they had a much larger program planned and we decided to stop APPROVe because we saw it in APPROVe.

DR. WOOD: Let's go on. Dr. Nissen?

DR. DOMANSKI: We have heard a lot of discussion about who know what, when, and we have seen a tremendous amount of data presented, and in the end

this committee is going to have to make some recommendation about what to do going forward. I am very interested, if we could, in hearing from each of the pharmaceutical manufacturers, as well as everybody else of course, before they sort of go away into the distance. I am very interested, given the totality of data that are currently available--not what you knew when or who should have known what, how or when or who should have done something else--I am very interested in what you think ought to be done now going forward, knowing what we know. What recommendation would you make? What would you like to see come out of this? Or, maybe what do you think we should see come out of this?

DR. WOOD: Dr. Nissen?

DR. NISSEN: Yes, a quick comment and a question. The comment is--and I think for people in the audience who may not fully understand why we are drilling down on this intent-to-treat aspect of the analysis--that it may be that the individuals who are dropping out of these trials because of adverse events, that received the COX-2 inhibitor, they may be pharmacogenomically more susceptible to the adverse effects of COX-2 inhibitors. So, you are taking out of the trial the people that are at greatest risk. If you don't follow those people you may not find that out. This idea of censoring events after two weeks--you know, I think we have to all learn something from what happened here, and this is the first time I really realized that that was the way these studies were conducted. That was a mistake. Once a patient is exposed to drug you ought to follow him as long as you can because there may be a persistence of risk and we

learn something from that. So, a lesson is learned. I think it is a useful lesson to learn. I guess the second question--and, you know, you may or may not want to answer this but if you had to do it all over again would you do it differently?

DR. WOOD: Let's keep the tense in the future tense. Let's not keep regurgitating that. Bob, do you want to say something in the future tense?

DR. TEMPLE: Yes. I just want to remind people that intent-to-treat analyses are generally loved by people because they are conservative analyses. They tend to make effects go away. That is why we like them. If you are worried about informative censoring and other stuff like that--

DR. WOOD: But it tends to make efficacy effects go away.

DR. TEMPLE: That is correct. They also make time effects go away.

DR. WOOD: Not if you are dead. (Laughter)

DR. TEMPLE: No, no, you have to count the deaths. It is not that you shouldn't follow people up but the analysis that includes all people long after they are off the drug has a very high likelihood, I believe, or not showing the effect of the drug. You have to remember it is a conservative analysis for looking for effects. Before we get too enthusiastic about it, if I make the effect look less when it really doesn't deserve to look less--

DR. FLEMING: Could I just quickly add to that? Historically we look at safety and often we do truncate follow-

up after two weeks or a month. That is based on the premise that safety risks are acute. If they are, in fact, acute, then you are going to get a clear sense of what is going on with the type of approach you are talking about. Mortality effects, I would think, are much more difficult to justify as being purely acute. There is a basis to what you are saying. If you follow everybody for a long time after they are off therapy there could be some diluting. Nevertheless, if you want an unbiased assessment of the truth you need to do what Steve is talking about, an ITT analysis, and then make your judgment as you look at the hazard ratio over time.

DR. D'AGOSTINO: We are sitting here and we don't know the answer. It may have washed it away and it may not have.

DR. TEMPLE: I am not saying don't get the analysis but, for example, our ordinary position in an outcome study is that we want to see the intent-to-treat analysis.

DR. WOOD: Let's hold this for the discussion. Let's just keep focused on the questions right now. Any further questions for the speaker? I am not forgetting about you. Hang on just a moment. Dr. Holmboe?

DR. HOLMBOE: I just had a question to the speaker. Again, we are trying to give you some advice and some guidance as to this. Given, as Alastair said earlier, that this drug was really evolved for a safety indication, therefore, being compared to another class of drugs, in retrospect learning that those comparisons were based on drugs approved prior to new knowledge that

has been accumulated, such as presented by Dr. FitzGerald, it would be helpful for me to hear what has the FDA learned about the process or form? What can you tell us that might help in the future when you are faced with these sorts of things? For example, the diclofenac is a perfect example, well, it turns out that maybe it is not, you know, your run-of-the-mill NSAID. A lot of what you presented in the original data was, like, well, it was between ibuprofen and diclofenac, therefore, we determined it was probably okay. So, I would be anxious to hear what you have learned since you have been with this project now for seven years.

DR. VILLALBA: Well, actually we wanted to have the recommendation from you to know how to proceed now because we have close to 20 approved NSAIDs so what do we do with them?

DR. WOOD: Ever the optimist, right! Dr. Domanski?

DR. DOMANSKI: I guess before Merck gets away I would still like to hear their view of where we should go from here. I am really quite curious about that. I understand about intention-to-treat. We do clinical trials. But I would just like to hear what their thoughts are.

DR. WOOD: Their thoughts on what?

DR. DOMANSKI: What their thoughts are on where we should go from here.

DR. WOOD: I thought we were talking about where we have been. I am happy to hear them on where they should go. Do you have thoughts on that, Bob?

DR. DOMANSKI: No, I am asking that of Merck.

DR. WOOD: Oh, I am sorry.

DR. BRAUNSTEIN: I think we showed that on our last slide. Can I see our last slide, 57? I mean, for the short term what we are trying to do is trying to better understand our data; trying to better understand which patients were at increased risk for the events that we observed in APPROVe based on both the clinical data and also the specimens that we have from these patients. We also are working with various people to try and explore different hypotheses for the data, and we are collaborating with others who are looking at the data across all the drugs in order to get a better feel to see if we can understand when we pool all the data because I don't think any one data set that we have is powerful enough to address these questions. So, hopefully, by pooling the data we will be able to get a better feel for this. The last is that we think we need to do comparative outcome studies to better understand the relative risks of the selective COX-2 agents with the traditional NSAIDs. There are not long-term data on the traditional NSAIDs to really establish what their cardiovascular risk profile is, and we think that the study that we are doing, for example the MEDAL study is one such study in the right direction.

DR. WOOD: Merck wanted to present some other data. Right?

DR. REICIN: I think there was a question about congestive heart failure in the Alzheimer studies.

DR. WOOD: Right.

DR. REICIN: So, just put up slide for us 12-22 and then we will go to 12-28. I showed you this slide just at the beginning and you noted that in our 6-month population--so this is a shorter population than either APPROVe or what I am going to show you in Alzheimer's--the rates were quite low and they were similar to the NSAIDs. If you go now to 12-28, in the Alzheimer's studies, in protocol 078 which was a 4-year study, interestingly, the rate of congestive heart failure was similar between the two groups, 2.2 percent on rofecoxib 25 mg, 2.6 percent on placebo. In 091, however, which was a one-year study the rate was a little bit higher on rofecoxib, 3.2 percent versus 1.4 percent. I think these rates are more what you would expect in an elderly population. The mean age of this patient population was 75 years old.

DR. WOOD: Thanks.

DR. REICIN: One other thing, there was a question about ITT mortality. In APPROVe we are following patients in an ITT way for mortality. That is still ongoing. To date, there were 3 thrombotic events in each treatment group following that 14-day period.

DR. WOOD: Dr. Paganini?

DR. PAGANINI: I have a question on the comparative data with other NSAIDs. Is there not a post-approval period of time for drug review, and from that post-approval Phase IV type studies can you not draw anything from that to compare to?

DR. VILLALBA: Phase IV commitments are made at the time of

approval. If there were not specific agreements between the FDA and the company to conduct those studies we have no legal power to mandate any kind of studies. So, some studies are done basically pursuing different--I mean with promotional, advertisement or whatever there are many studies. But those are really not usually large outcome studies. They are short studies with small numbers of patients. I don't know if I answered your question.

DR. PAGANINI: You did in a way. One of the issues that I think we are going to have to face is how do you compare these things, both things that have already been approved and new, to the same standards when they were approved back then to current standards? Perhaps one of the ways around that might be an approval comparison with longer Phase IV commitments by companies to follow-up on what is happening to that drug over time. That way, you would have the ability to compare a new to a similar in a similar population of patients.

DR. VILLALBA: Absolutely. That is something that we learned, yes.

DR. WOOD: Ralph?

DR. D'AGOSTINO: I am all for torturing data and during Lent I always read Dante's "Inferno." (Laughter) But shouldn't we be impressed with the APPROVe study? You leave us with a table that compares a lot of studies and you throw out some obviously important questions, but shouldn't we sort of look very seriously at the APPROVe study? It was well designed--

DR. VILLALBA: Of course.

DR. D'AGOSTINO: --and shouldn't we sort of diminish in our view some of the previous studies?

DR. VILLALBA: The Alzheimer's studies, do you mean? Now, yes. What I was saying is that these are different populations and I do not have a good explanation for why we didn't see the same in an elderly population.

DR. D'AGOSTINO: Well, could it be that the APPROVe study was going after a particular set of outcomes and the others weren't, and it was more retrospective?

DR. VILLALBA: No, because in the Alzheimer's studies they also used the same standard operating procedures to adjudicate the event.

DR. D'AGOSTINO: But do they have the same ascertainment? You know, in designing a placebo-controlled study where you go after something retrospectively, looking at that and trying to say the ascertainment might have been the same.

DR. VILLALBA: You are completely right. That is possible but that is a question to the sponsor, if the ascertainment could have been different in the Alzheimer's studies.

DR. WOOD: Dr. Hennekens? Actually, I would like to ask Marvin a question. Marvin, the APPROVe study was scheduled to terminate at about 6 weeks after the early termination on the basis of the board's recommendation that you were on. As I recall, the numbers of events were 45 and 25 at that time. So, was the board unanimous in its decision

to terminate, and was the basis clearly related to that particular endpoint?

DR. KONSTAM: Yes. Yes, that is exactly right. I would say that the reason, if I might say why we recommended termination--the reason we recommended termination is that we felt at that point in time that we had a definitive piece of information that wasn't going to change. The reason we recommended termination was that we felt the patients in the APPROVe study were not aware of this and had not been consented to this adverse effect. So, in our judgment, you know, from an ethical viewpoint if you were going to continue you would have to go back and re-consent them and that certainly wasn't practical at that point in time. So, that is the specific reason we recommended termination.

DR. WOOD: But you told them that caution should be exercised in patients with heart failure. Right?

DR. VILLALBA: May I say something?

DR. WOOD: Sure.

DR. VILLALBA: I don't want to leave you with the impression that we think or I think that APPROVe is not important. I just want to show you how puzzled we were with all the data. So, until APPROVe we didn't have a firm reason to really take a regulatory action that was different from what we had done up to that time.

DR. WOOD: Ralph again?

DR. D'AGOSTINO: In terms of the Alzheimer's study, do you have information on the all-cause mortality? I forget what you said. Do you have anything about CVD, cardiovascular mortality when off drugs?

DR. WOOD: Let's take that under advisement unless you have it right there. Do you? No? All right, we will get back to that. Any other questions? Yes?

DR. TEMPLE: Actually, I wanted to respond to Ralph. The thing about APPROVe is that it was longer than the rest of the studies and most of the effects were seen sort of late. So, it provided the kind of information that really didn't exist before.

DR. D'AGOSTINO: When we come to the discussion of designing the trial, there is so much emphasis on how many events we should have and I am always bothered by that because I would like to make sure people have taken the drug for a long enough time. I think this is a case where you are seeing where length is where something is happening.

DR. WOOD: Unless there are any other questions, let's stop our discussion of Vioxx at this point, rofecoxib, and take a ten-minute break. We will reconvene and start on celecoxib when we get back.