

# Committee Preliminary Discussion on Questions 1-3: End of Day 2

## 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.....	8
Question 1 Discussion.....	8
Question 2 Discussion.....	24
Question 3 Discussion.....	37

### Highlights

#### QUESTION 1: DO COX-2 DRUGS INCREASE CARDIOVASCULAR RISK?

- **DESPITE A CLASS EFFECT THERE ARE DIFFERENCES BY DRUG AND BY DOSE:** Dr. Gross said that, although there may be a “class effect”, the drugs to be discussed are sufficiently different so that “one or more of the drugs .... should be marketable .... and one or more of the drugs ... should not.” Dr Nissen said that there was a “class effect, but I think that there is clearly evidence of a gradient in risk, and that gradient is not only by drug, but also by dose.”
- **SEPARATE QUESTION INTO ARTHRITIS AND NON-ARTHRITIS:** Dr. Gibovsky interjected to ask “that we discuss it in the context of patients with arthritis versus patients with other conditions”
- **POINT ESTIMATES OF RISK WITH EACH COX-2:** Dr. Fleming gave an overview of the statistical evidence and concluded: “There certainly is an effect that is going on

here...” Although the point estimates are not precise, he estimated hazard ratios of 1.4-1.5 (rofecoxib), 2.5 (valdecoxib), 1.3 (celecoxib), 1.6 (etoricoxib) and 1.2 (lumiracoxib). Dr Hennekens said that Dr. Fleming’s estimates of hazard ratio are “lower than one would have guessed based on the early data-dependent stopping of some of these trials”. Dr. Wood suggested that the favorable results in the arthritis studies might be “diluting the effects” when Dr. Fleming made his hazard ratio estimates. Dr. Fleming responded that “the best analysis ..... is one that will drill down in all of these dimensions as best we can.” Dr. Wood’s response was: “It is not likely to be less than the numbers you gave. It is likely to be more, right?” Dr. Fleming responded that “There will be settings where it is less. There will be settings where it is more.”

- **LOW DOSE CELECOXIB:** Dr. Fleming said “Can you, for example, give celecoxib at a low dose with a short enough duration that in wide settings, it would be safe. That is still entirely possible within the context of what we have said”. Dr. D’Agostino expressed concern that “if we give this global statement” we “can't get back to the question you just raised: Can we look at Celebrex at a low dose?” Later, Dr. D’Agostino said “it may be a low dose of Celebrex, that may be viable, and not unsafe”.
- **DEFINITION OF COX-2 DRUG:** Dr. Shafer asked “What is a COX-2?” and suggested that this might include meloxicam, Sulindac and diclofenac. Dr. Wood said “Let's stick to the drugs for which we have got evidence.” Dr. Abramson echoed Dr. Shafer’s concern and pointed out that the newer drugs are the only ones with long-term studies and that the doses used in those studies were higher than the usual doses of comparators. Dr. D’Agostino said his “comment is very much the same”.
- **IF IT IS JUST COX-2, THEN ALL NSAIDS IMPLICATED:** Dr. Abramson pointed out that all the NSAIDs, including the older non-selective ones, inhibit COX-2, so that if the increased cardiovascular risk is related to COX-2 inhibition per se, it should occur with all NSAIDs. Dr. Gross expressed concern that “if we conclude that there is a class effect for the selective COX-2 inhibitors ..... then people are not going to want to use COX-2 inhibitors at all and they will be using the non-selective NSAIDs, which from the data presented, doesn't look as though many of them are better from a risk point of view”.
- **DIFFERENTIAL RISKS WITH DIFFERENT DRUGS:** Dr. Fleming’s take on the traditional NSAIDs was that “naproxen is a winner” and that “the theory that was put forward that diclofenac is COX-2-like is at least supported by the trials where it was studied”. Dr. Hennekens said that “the picture that ... begins to emerge to me is that rofecoxib, ibuprofen, and possibly valdecoxib are in one bin, diclofenac and celecoxib in another bin, naproxen in a third bin, and then aspirin in the fourth bin, going from concerns about hazard to neutrality to benefit.” Dr. Manzi commented that she had “a problem making inferences about diclofenac and naproxen in studies where I think we have a difficult time feeling comfortable with the results in relationship to the COX-2s..... So, to feel that we don't have enough information to really feel comfortable with COX-2s, and then to try and extrapolate to the comparators in those, I think is dangerous”.
- **IMPORTANCE OF BLOOD PRESSURE EFFECTS:** Dr. Nissen emphasized the importance of blood pressure effects and pointed out that lumiracoxib and rofecoxib, both very COX-2 selective drugs, have very different effects on blood pressure and that “the relationship between relatively small differences in blood and cardiovascular morbidity and mortality is rock solid across a huge number of drugs and interventions, and you almost can predict what will happen”. Dr. Temple said that comparative studies of blood

pressure effects with these drugs could be done with 20-40 patients per group.

- **ASPIRIN DATA MAKES MECHANISM UNCLEAR:** Dr. Shafer again expressed concern about the fact that risk did not seem to be reduced by concomitant aspirin which seemed to raise doubts about “the Fitzgerald hypothesis”. Dr. Temple agreed that aspirin is “a fly in the ointment”.
- **HIGH COX-2 DOSES IN TRIALS MAY NOT HAVE BEEN WISE:** Dr. Temple commented that “We encouraged everybody to study high doses to rule out GI distress. Whether that was wise in retrospect, I am not sure, and I think I probably had something to do with it a long time ago. I am not sure that was the best thing.”
- **CORRELATES OF INCREASED RISK:** Dr. Farrar said that “the ones with some of the higher risks are not necessarily the most COX-2 selective”. Dr. Cryer said that increased GI risk and increased CV risk go together in the same patient. Dr. Cryer said that lumiracoxib which is not yet on the US market showed GI benefit. Dr. Wood pointed out that such benefit was

seen at doses that had an increased cardiovascular risk.

- **PROBLEM WITH LACK OF FOLLOW-UP:** Dr. Holmboe and Dr. Fleming expressed concern that we don’t have cardiovascular safety data collected for the patients who dropped out early.
- **RISK NOT LARGE ENOUGH TO CONTINUE HOLD ON ONGOING TRIALS:** Dr. Farrar said that with “all of this discussion about risk, I don’t want to imply that I think that this risk is big enough to actually warrant the continued hold on all the trials that we have going, and I think, in fact, what it suggests to me is that we need to continue with trials to understand better what the data is telling us.”
- **CHART TO COMPARE DRUG DROPOUT RATES:** Dr. Day suggested that “a giant chart” be prepared “before tomorrow” that lists the percentage dropout rates for each drug together with the reasons for drop-out. Dr. Temple said “The question is how reliable it is..... it isn’t by any means always properly done”.

## QUESTION 2: OBSERVATIONAL STUDIES

- **OBSERVATIONAL STUDIES ARE ONLY HYPOTHESIS GENERATING:** Dr. Nissen said that we should only “use observational studies as hypothesis-generating studies. If you see a signal in an observational study, it is an indicator that you need to do a randomized controlled trial..”. Dr.

Fleming agreed that observational studies should be considered hypothesis generation; passive surveillance can detect rare events but active surveillance is needed to detect a change in frequency of common events; however, in all observational studies “recorded covariates are just the tip of the

iceberg, so you are left with a great deal of uncertainty about bias” and these studies can give “us a guide because we can't do randomized trials in every setting... but the ultimate answers in most cases really come from the randomized trials.” Dr. Wood also agreed and emphasized the point made earlier by Dr. Nissen that observational studies had fooled everyone into thinking that estrogen therapy prevented heart disease in post-menopausal women. Dr. Hennekens also agreed and in this respect quoted Bradford Hill who he had the privilege of knowing. Dr. Farrar said that “Every single one of them is confounded by indication. The best example is the indomethacin one where it is only used in people who are sicker than people who aren't.”

- **DESPITE LIMITATIONS, OBSERVATIONAL STUDIES HAVE ADVANTAGES:** Dr. Stemhagen made some general comments on observational studies: Despite their limitations, they do provide additional information not available from the randomized trials: dose-response data, data in a setting of adequate follow-up, real-world patients and real-world doses. Dr. Holmboe said that “a poor randomized controlled trial actually may be worse than a good observational study” and randomized studies can have poor follow-up and other problems. Similarly, meta-analyses of randomized trials need “a test of heterogeneity to see if they really truly could be combined”.
- **FDA EMPOWERMENT TO MANDATE STUDIES:** Dr. Cush suggested that Congress empower FDA to mandate post-marketing

trials, including observational studies.

- **LOWER RISK BECAUSE OF INTERMITTENT DOSING IN OBSERVATIONAL STUDIES:** Dr. Bathon suggested that the risk of drugs in observational studies might be lower than in randomized trials since patients tend to take their drugs only on an intermittent basis.
- **“WILD ARM” OBSERVATIONAL ANALOG IN CLINICAL TRIALS:** Dr. Paganini suggested that randomized trials include a “wild arm” that consists of patients receiving “what is usually and customary done” when prescribing the drug. This would provide an observational study analog in a controlled setting. Dr. Domanski was intrigued by the “wild arm” concept. Dr. Paganini gave the example of an ongoing NIH-VA dialysis study which has three arms: high dose, low dose, and “what is usually and customarily done”. However, in further discussion with Dr. Domanski, he said that the “wild arm” was really just a prospective registry and that patients were not randomized to this third arm. Dr. Friedman thought that the “wild arm” concept was not a good idea “when we were dealing with the ARDSNet issue”.
- **VARYING RISK LEVEL ACROSS TIME NOT EVALUATED IN OBSERVATIONAL STUDIES:** Dr. Elashoff pointed out that where the risk level varies across time “the standard way that observational studies lump it all into patient years is bound to be misleading”.
- **SHOULD RECOMMEND GOOD OBSERVATIONAL STUDY**

**DESIGN TO FDA:** Dr. Friedman suggested that it would be useful to advise FDA on the best use of observational studies rather than just focusing on “all of the difficulties in using observational studies”.

- **DETECTING COMMON VS. RARE EVENTS IN OBSERVATIONAL STUDIES:**

There was discussion between Dr. Platt, Dr. Wood and Dr. Temple about the ability to detect a change in the frequency of common events if one were to mandate post-marketing observational studies. Dr. Platt felt that if one used an automated recording system in “the model of the large linked databases” to identify certain pre-specified common events, the problem of under-reporting could be handled. Dr. Temple felt that this issue with chronic use drugs was always cardiovascular and agreed that Dr. Platt’s system could probably be put in place. Dr. Wood expressed concern about mandating such a study for every chronic use drug but Dr. Temple felt that “we have access to databases, whether it’s California Medicaid or whatever, and one can do that”. “Maybe that is something that we could think about.”

- **OUTSIDE PRESSURE TO BAN DRUGS ON BASIS OF INVALID**

**OBSERVATIONAL STUDIES:**

Dr. Jenkins (FDA) expressed concern that, going forward, “this going to be a mining exercise for everyone who does observational studies in the world probably. They are going to be looking to do another COX-2 or another NSAID observational study” and he was interested in what the committee thought FDA should do with “the next observational study that is touted as wow, this really shows something”. Dr. Nissen suggested that “We have seen some strange things go on, like the warning around naproxen, that was clearly based upon pretty weak evidence. So, I think having a good standard is where you have to kind of hold your ground.” Dr. Wood said that “...it certainly seems to be in the public interest that you [FDA] should have the power to ensure that that kind of a study gets done, and that is something certainly people should hear and hear loudly...”. Dr. O’Neill said that one of the problems was that observational studies are often poorly designed and “it is not even clear what the prespecified hypotheses were”.

**QUESTION 3: COX-2 BENEFITS VS. NON-SELECTIVE NSAIDS**

- **QUESTION 3:** Discuss the available data regarding the potential benefits of COX-2 selective non-steroidals versus non-selective non-steroidals, and how any such benefits should be weighed in assessing the potential benefits versus the potential risks of COX-2 selective agents.

- **FDA PERSPECTIVE:** Dr. Jenkins (FDA) asked for the committee views on the GI advantages of COX-2 drugs, efficacy in pain relief, and about “the value of choice” by making multiple drugs available.
- **EFFICACY, RISK WEIGHTING & COMPASSION:** Dr. Nissen said

“I haven't seen any compelling evidence that in terms of pain relief, that the drugs are actually more effective...”, that “the GI events here are serious events” but that “I have to give them less credence than the kind of hard, permanently disabling effects of MI and stroke”, and that “compassion has to come into our decisions” for “patients who just don't tolerate the conventional NSAIDs”. Dr. Fleming said that in “a crude estimate” the COX-2 drugs might be preventing “7 per 1,000” GI events but that “how is that up against 4 events that are strokes, MIs, or cardiovascular deaths?” However, the benefit-risk may be different in certain patients and it would be important to “more scientifically, rigorously establish certain subpopulations where there really is a differential relief”. Dr. Fleming said that if there is no efficacy advantage to the COX-2 drugs, “there should be an incredibly low threshold for what you would accept in additional cardiovascular events, because the only thing you are getting relative to nonspecific NSAIDs then would be a very small GI.”

- **COMPARE TO NSAIDS + PPI:** Dr. Hoffman said that it is important to obtain 2-3 year data comparing the GI safety of non-selective NSAIDs in combination with proton pump inhibitors to COX-2 drugs alone.
- **IS LOWEST SAFE DOSE STILL EFFECTIVE?:** Dr. Hoffman also said that If COX-2 drug dosage is going to be reduced to “the lowest safe dose” it needs to be established that such doses have adequate efficacy and retain the GI protective effect of higher doses. In addition,

although “it is our obligation to provide patients choice” we “shouldn't give people a choice if we think that choice is uninformed and potentially does harm”.

- **MANY OTHER CHOICES AVAILABLE:** Dr. Cryer said that even without the choice of COX-2 drugs, patients would still have the choice of “20 other NSAIDs available in the U.S”.
- **COX-2 GI BENEFITS SMALL:** Dr. Cryer also said that he thinks that “the GI benefits are less than previously speculated”. The GI advantage of COX-2 drugs versus non-selective NSAIDs may be greatest in patients at low GI risk and less in those patients who are at greatest GI risk. In addition, in “the face of low-dose aspirin, there is no apparent GI benefit”. Dr. Wood said that for the two drugs presently on the US market (celecoxib and valdecoxib) “we have no clear randomized data that show GI benefit”.
- **MERCK'S VIEW:** Dr. Kim (Merck) said that rofecoxib has “unique benefits” and that the increased cardiovascular risk appears to be a “class effect”. Dr. Wood asked if he was saying that “if we think the cardiovascular effect is a class effect, you would consider putting Vioxx back on the market”. Dr. Kim did not answer directly. Later in this session, Dr. Crawford asked Dr. Kim about the “potential reintroduction” of Vioxx by Merck. Dr. Kim again responded ambiguously, seeming to link any reintroduction decision a decision by the committee that the increased cardiovascular risk is a class effect. At the two minute Sponsors’

summary presentations on Day 3, Merck appeared to be on the point of clarifying this point, but the two minutes allotted were up and the microphone went dead – to the accompaniment of considerable laughter from the audience.

- **DIFFERENTIAL EFFICACY IN INDIVIDUALS & CHOICE:** Dr. Farrar said that the discussion so far had not really addressed benefit (a relative lack of GI toxicity is not an absolute benefit but a reduced risk). Clinically, he has patients who say one drug is effective whereas another drug is not – although no controlled studies have substantiated such an effect. He thinks choice is important. Dr. Dworkin commented that there is “a really solid basis for their needing to be a choice amongst several drugs, because you have the variability in the pain benefit amongst patients and the variability in their tolerability”.
- **GI MORTALITY RISK IS SUBSTANTIAL:** Dr. Gibovsky pointed out that, although myocardial infarction can be fatal, Dr. Singh’s data show that 16% of GI bleeds are also fatal. In addition, COX-2 drugs are preferable in patients on anticoagulants post-surgery because of their lack of platelet inhibition. He also mentioned the ACDA and PACES trials in which patients preferred celecoxib or diclofenac over acetaminophen. Another point is that tolerance to NSAIDs is common (the latest data set shows that patients switch therapy 3-4 times within 18 months) and that patients may not tolerate a given drug because of allergy or idiosyncrasy. This argues for maintaining a choice of drugs. Dr. Wood expressed concern that the

most recent estimates of GI complication rates indicated that this was less of a problem than was being suggested. Dr. Singh said that the more recent figures did show that the death rate and duodenal ulcer rates have dropped. However, the hospitalization rates had been underestimated in the earlier studies and “there are a lot more than that”, and the “gastric ulcer rates and the gastric ulcer hemorrhage rate have not gone down in the same fashion”.

- **INVOLVE PATIENT IN RISK ASSESSMENT:** Dr. Morris said that it was difficult to weigh the relative importance of events such as TIA or a gastric ulcer but “we need to understand patients' evaluation of these outcomes”. Dr. Platt said that “we could do a very much better job than we do by using the existing data that FDA already has to provide good information to patients about the risk stratum that they inhabit.” Dr. Bathon commented “that is, in fact, what most of us rheumatologists have been doing for the past four months with every single clinic visit”.
- **NEED FOR AGGRESSIVE MANAGEMENT OF CARDIOVASCULAR RISK:** Dr. Cush said that using only traditional agents such as “Tylenol and aspirin and ibuprofen .... would be a gigantic step backwards” and “what we need” is “a strategy for risk modification”. Dr. Hennekens said that arthritis patients are at increased cardiovascular risk and that their cardiovascular risk factors should be “managed aggressively”.
- **NO PROOF THAT ACETAMINOPHEN IS SAFER OR AS EFFECTIVE:** Dr.

Hennekens expressed concern that it was not clear that acetaminophen “is either sufficiently efficacious or much safer”.

- **NEED FOR DRUGS IN PEDIATRICS:** Dr. Ilowite expressed concern about the limited choice of agents in pediatrics. Also “the risk of cardiovascular disease ... is very low in pediatrics” but that because duration of therapy is “longer even than adults”, it is important to “to get some insight into the pathogenesis of this ... so that early markers could be explored in children...”
- **FDA INTERVENTION OR DOCTOR/PATIENT RELATIONSHIP FOR RISK MANAGEMENT:** Mr. Levin said that we should not “just sort of slide this all off on patients and physicians supposed in this Nirvana good, up-to-date information” and that we “can't abrogate our responsibility, and we can't pretend the Government, through the FDA, doesn't have a statutory responsibility here to protect the public health.” Ms. Malone

disagreed, saying that the rheumatologists on the committee “do have the ability to form a relationship with” patients and “I applaud them for that”.

- **NON-SELECTIVE NSAIDS MAY NOT BE SAFE ALTERNATIVES:** Dr. Manzi expressed concern about those suggesting there are “safe alternatives” to the COX-2 drugs since “I think we have signals actually to the opposite .... patients are going to have to turn to something, and do you feel comfortable saying that the alternatives are safe?” Dr. Wood said that on the other hand “I think it's highly improbable that the committee would have approved any of these drugs given the safety signal we have got right now”. Dr. Temple agreed that “some of them I think probably would not have made it”. Dr. Manzi responded that such a decision would depend on the need for the drug and what alternatives were available rather than looking “at it in isolation”.

## Discussion Text

### *Question 1 Discussion*

DR. GROSS: On the first question, I would like to propose a construct to deal with the issue: Is the increase in cardiovascular risk a class effect. My proposal is to say yes, it is, but the degree of difference and the time of difference varies and is different enough that one or more of the drugs that we have discussed should be marketable with a precaution and/or warning, and one or more of the drugs we have

discussed should not. A reasonable analogy is statins. As we know, they all have potential for liver toxicity and myopathy. That is a class effect, but the degree of this difference and the time when it occurs varies and is different enough that one or more of the drugs have been marketed with a precaution or warning, and one or more have not. Tomorrow, we will discuss specifically the recommendations on celecoxib,

valdecoxib, and rofecoxib, but I thought I would start off the discussion with this question about a class effect.

DR. WOOD: Okay. Dr. Nissen.

DR. NISSEN: Did you mean class effect for the COX-2s, or are you talking about NSAIDs, as well, because the question is asked for both here. So, I want to know which of those you mean.

DR. WOOD: Let me make a suggestion. I think we should start with COX-2s. The data we have seen is by far the most convincing for that. Then, let's move on to any other issues.

DR. NISSEN: So, let me agree that is what we are talking about then.

DR. WOOD: Let's have a discussion around the COX-2s first and whether the available data support a conclusion that cardiovascular risk is a class effect for all--

DR. GIBOFSKY: Could I just interject and ask then that we discuss it in the context of patients with arthritis versus patients with other conditions?

DR. WOOD: Okay, that's fine; the committee can do that, but remember we are not discussing the relative risk-benefit at this point. We are discussing whether there is an effect, a signal, in other words.

DR. GIBOFSKY: I understand, but I think it is relevant to look at the populations in which the signal has been detected.

DR. WOOD: Do you want to comment on that and save us going to back? Do

you think that the arthritis population will be likely to have a lower risk than the other populations?

DR. GIBOFSKY: I am merely saying that I think that one looks at populations. As we have heard, there is variability in the population, and just as we wouldn't automatically extrapolate efficacy data from one population to another, I am not certain we can automatically extrapolate safety data from one population to another, and I think we need to discuss it in the context of the population studied.

DR. WOOD: Any other comments on this question?

DR. ILOWITE: I think you made the point that this was merely a discussion of safety, but I think the way the proposal was worded, there is implications about cost-benefit with regards to whether they should be approved or not. Could you repeat the--

DR. WOOD: You have the question in front of you.

DR. ILOWITE: The proposal.

DR. WOOD: Dr. Gross made a proposal, but the question we have got in front of us is to discuss the available data supporting a conclusion of increased cardiovascular risk for COX-2 selective non-steroidals. I think we need to discuss that before we get to risk-benefit frankly. Dr. Nissen.

DR. NISSEN: I think your proposal is an appropriate one and I would point out that we have at least one randomized trial for every drug that has been marketed in the class that shows an effect.

DR. WOOD: You mean a risk.

DR. NISSEN: That makes the grade in terms of calling it a class effect, but I think that there is clearly evidence of a gradient in risk, and that gradient is not only by drug, but also by dose. So, saying it is a class effect means--let me tell you what it means to me. It means that if you give a high enough dose of one of these drugs to a risky enough patient, you can produce an increased risk of adverse cardiovascular outcomes. But it doesn't mean that a particular dose in a particular population is risky.

DR. WOOD: Dr. Fleming.

DR. FLEMING: I think there is a great deal of data that is giving us a general sense, but there is an inadequate amount of information to really get at the specifics, and what I mean by that is certainly the indication, the dose, the duration of therapy, the nature of ancillary care, for example, aspirin use, these are all factors that obviously could influence the answer.

The approach that I took was to try to summarize the essence of what I think we have been presented in the randomized trials, and I focused in particular on those that were the major trials, many of them looking at somewhat longer term exposure and longer term follow-up.

There are about 15, and just to quickly run through them:

- In the Vioxx setting, there are 23,000 patients from four major trials. Those studies indicate something on the order of 1.4 to 1.5 relative risk, and

are driven heavily by VIGOR and APPROVe, and neutralized somewhat by the Alzheimer's 078-091 trial although that trial had surprisingly considerable excess deaths.

- In the Bextra setting, the Nessmeier 071 trial, the 035, and 069 studies give about a 2 1/2 relative risk even though it is certainly heavily driven by this CABG setting.
- In the Celebrex trials, the CLASS, the Alzheimer's 001, the APC, the PreSAP, now, we know there is the ADAPT, but we haven't been shown that, so I did the first four, and we are looking about a relative risk of 1.3, driven heavily by the APC trial and the 001 study, and neutralized by the CLASS study and the PreSAP that were more neutral.
- The etoricoxib, the EDGE trial, and the other three that we were presented give us a relative risk of about 1.625.
- And in the lumiracoxib, it is about 1.18 relative risk from the TARGET trial.

Now, to put these into context, if we were trying to show--I am just going to give you four scenarios--a doubling. By the way, I am working off a 1 percent background rate, and that is just about what these data show in the aggregate, in 73,000 patients, about a 1 percent aggregate rate for the primary cardiovascular endpoint of cardiovascular death, stroke, and MI. If you were trying to show a doubling, it takes 88 events or about 5,000 people. If you were trying to show a 50 percent

increase, it's 256 events, 20,000 people. If you are trying to show a 33 percent increase, it's 508 events, 40,000 people, and if you are trying to show just a 20 percent relative increase, from 1 to 1.2, it's 1,265 events or 115,000 people. Where we are, if you ignore all those factors that I was arguing we can't ignore because the answer isn't the same, but if you put all this into a single pool, these add up to something in the neighborhood of a relative risk of about 1.4 to 1.45.

So, essentially, what it would have taken to discern that is an aggregate data of about 70,000, although the observed results that you would have to have, you would have to have a study on the order of about 5,000 people, because an observed result of 1.55 or 1.45 is statistically significant when you have about 6,000 or 7,000 people.

So, the point is when you look at the aggregate, we have substantial data to say there is conclusive evidence here in the aggregate that there is the cardiovascular risk.

Now, what can we say individually? In the Vioxx setting, where the risk is about 1.43, one would need to have, with that observed rate, you would have needed to have data on about 6-8,000 people. We have data on 23,000. That is why the evidence there is very clear. In the Bextra setting, we only have data on 3,000 in the Nessmeier 035 and 069 trials, but the relative risk is 2.58, and for a 2.58, you need less than 2,000 people, hence, that is why it is statistically significant in that setting although it is only in the CABG setting. It is also in the etoricoxib setting with the relative risk of 1.625, we would have needed less than 5,000 people. We have

17,000, so it is statistically significant in that category. In the lumiracoxib setting, we have a relative risk of 1.18. That would have taken over 40,000 people for that relative risk to be detectable. We only have 18,000. So, it is suggestive of a modest or moderate excess, but it is not proven because of the smaller sample size or because of the smaller effect. In the Celebrex, where it is about 1.29, it would have taken 20,000 people, if you observed that in 20,000, it would have been marginally significant. We observed it in approximately 12,000, so it is suggested, but not established.

Now, a lot of this, this is looking at things in a first pass. It is suggestive that there is something going on in all of these cases, but at very different levels is what at least the data show, but the data aren't conclusive for us to be able to say in a reliable way: What is the indication? What is the dose? What is the duration? Is it in aspirin, not in aspirin, but globally?

There certainly is an effect that is going on here, and for three of these five, it seems to be conclusively established, and for the other two, more modest effects that are suggestive.

DR. WOOD: Thanks. Dr. Shafer.

DR. SHAFER: I don't know that this will help our discussions at all, but I think one of the things we need to address is what is a COX-2. It has been assumed, I think, that we are going to go with the company's definitions when they way we have a COX-2 drug, COX-2 selective, but, in fact, if you go to Warner's review in FSAAB here, from 2004, we see that meloxicam, Sulindac, and as we have heard, even diclofenac is

potentially considered a COX-2. Should we include these drugs in the discussions? We certainly won't have the evidentiary evidence that we have.

DR. WOOD: Let's stick to the drugs for which we have got evidence, otherwise, we will be here until midnight.

DR. WOOD: Charlie.

DR. HENNEKENS: I want to support the very crucial statements of Steve and Tom here. It does appear to be a class effect, which varies by drug and by dose, but the magnitude of that risk, which I also estimate to be 1.4 to 1.5, is lower than one would have guessed based on the early data-dependent stopping of some of these trials, based on the reported research, and the media coverage of all of this, so I think it is important to get Tom's quantitation and Steve's caveats about dose and drug and magnitude in there.

DR. WOOD: Dr. Abramson.

DR. ABRAMSON: I want to go back to this issue of definitions because I don't think that it is that simple to say that the coxibs that we are talking about are the only drugs that we need to discuss. The concept of "diclofenac lite", that Garret proposed, or has stuck, but I think that term could be applied to meloxicam, nemesulide. I think what we are stuck with is that assuming there is a class effect, we haven't excluded the fact that that class includes those other COX-2 preferential drugs, that we might agree that COX-2 inhibition is at fault here, and that is giving rise to some of these side effects, but it isn't precisely due only to those drugs.

Now, those drugs happen to have done the long-term studies, they have done them frankly at 2X dose compared to their comparators, and frankly, when you look at the randomized clinical trial development program with a relatively few placebo arms, the drugs look relatively comfortable. So, if you conclude that there is an increased risk because of COX-2 effects and hypertension perhaps, then, I don't think it's really fair to restrict the discussion simply to those drugs that got marketed as coxibs, the randomized clinical trials, especially if you do agree that perhaps Naprosyn has a modest protective effect, I think don't give a bye to these other drugs.

I think we have to look at certainly the case of celecoxib, that drug is relatively comparable pharmacodynamically to the other several drugs, so I think it is a much more complicated question than simply saying the COX-2 coxibs in this discussion, and we have to do apples to apples if we are going to make recommendations.

DR. WOOD: Dr. D'Agostino.

DR. D'AGOSTINO: My comment is very much the same. I am concerned about taking all these individual studies. I think, you know, sort of the potential is clear, but are we really lumping just because there is a direction on these here. Tom, for example, the question about splitting the arthritis populations versus the other populations, the arthritis populations come basically from the old clinical trials where adjudication was a problem and things of that nature. So, how much do we believe that data and how much do we want to draw this inference? So, I don't have a problem

with sort of coming up with some sort of global statement that we are concerned, but I am concerned at this point about quantification in a very heavy way, and we just may be overdoing it in terms of how we are sort of answering this question.

DR. WOOD: Just to make sure I understand your point, you are concerned about putting a number on it?

DR. D'AGOSTINO: I am concerned about this global, I mean for us to say that it's all COX-2.

DR. WOOD: But you would be comfortable naming names?

DR. D'AGOSTINO: I don't know what I would be comfortable with. I am uncomfortable with the sort of global statement that we have seen a number of studies--

DR. WOOD: I am just trying to draw out what you are saying. You would be more comfortable--

DR. D'AGOSTINO: The only thing we said in terms of separating was the arthritis studies. Are we comfortable with the arthritis studies, do we have enough information, do we feel comfortable enough with the adjudication process, the recognition of the cardiovascular events in those studies? I mean I think I am much more comfortable when we come to these new studies.

DR. WOOD: So, what you are saying is that the dilutional effect of these old studies may be substantial.

DR. D'AGOSTINO: Exactly, and I don't know how we are actually dealing with that.

DR. WOOD: So, that is important for people to understand. Do you want to develop that a bit?

DR. D'AGOSTINO: What is that?

DR. WOOD: So, what you are saying is that the studies that didn't have a cardiovascular endpoint--

DR. D'AGOSTINO: And trying to get adjudication. They were showing a signal. We already there is a signal.

DR. WOOD: So, they may be diluting the effects from when Tom adds on the back of the envelope--is that reasonable, Tom?

DR. FLEMING: Well, I fully agree that the best analysis of this is one that we don't have the time to summarize here right now, but it is one that will drill down in all of these dimensions as best we can. Almost certainly, the answer is here, if we had unlimited data. We have about 75,000 people. That is a lot of insight, although we need far more than that. We do have another 30,000 coming along shortly. In essence, though, what we really need, if we had the ideal, is that ability to drill down, as Ralph says, by indication, and by dose, by ancillary care, looking at whether or not it is an aspirin or not an aspirin, by duration of therapy. These are all things that we have seen the data suggesting that there is very likely these factors are influential. So, essentially, my attempt was to say, in a very crude way, what do you see from 10,000 feet here, but then

acknowledge exactly, as Ralph said, that you really do need to drill down.

DR. WOOD: It is not likely to be less than the numbers you gave. It is likely to be more, right?

DR. FLEMING: Well, my own sense about this is this is the weighted average of the compilation of all these different settings, and so in all likelihood, in fact, with certainty: There will be settings where it is less. There will be settings where it is more. Can you, for example, give Celebrex at a low dose with a short enough duration that in wide settings, it would be safe. That is still entirely possible within the context of what we have said. Those are issues that we really need to understand.

DR. D'AGOSTINO: And I am concerned somewhat that if we give this global statement, that we sort of can't get back to the question you just raised: Can we look at Celebrex at a low dose? Because somehow or other, we are saying it's in all the COX-2s. So, I want to be just careful in how, the answer to this question, how it comes out quantitatively and what it locks us into in terms of further discussion.

DR. FLEMING: What I would say is that--and I do agree that we need to look beyond these five or six products--but what I would say in several of these products, not just Vioxx, in the certain settings that we have looked at, in my view, there is evidence that establishes there is an excess risk. There are other products where there is a suggestion, and we are underpowered, though, to discern whether or not that suggestion--it is not a suggestion of nothing, though, it is a suggestion of something, but it is more

modest in size than the other agents although it could be a dose issue, it could be an indication issue that explains it.

DR. WOOD: Dr. Domanski.

DR. DOMANSKI: I think if one looks, there is really an attempt here to look very carefully in these quantitatives that we can. I think Dr. Fleming provided a remarkable compilation just now for us. But I think if one backs off to sort of high altitude and looks at these drugs, the signal, as people are using the term, it is pretty clear that there is an excess risk conferred by some or all of these drugs. It seems to me the process is sort of turning around. We are trying very hard, you know, the idea is to demonstrate--and it's the sponsor who has to do it--to demonstrate safety and efficacy, and not necessarily the purpose of the FDA or its advisory committees to somehow demonstrate that the thing is unsafe. It does look like they are unsafe, but the problem is that the studies presented really are not very good studies, and, in fact, one of the reasons that we probably didn't learn sooner that there is probably a real problem with these drugs is because of the relatively poor studies that were presented for approval. So, I think it is important to remember who has got what role. It is theirs to demonstrate safety and effectiveness; it is not ours to demonstrate it's unsafe.

DR. WOOD: Right, but I think the FDA is looking to us to give them some guidance here, right?

DR. DOMANSKI: Well, that is important for the future particularly, that

is: What studies should be done next?  
And that is a legitimate concern.

DR. WOOD: Bob, do you want to say something?

DR. TEMPLE: Just that it is the company's job to show that it is safe, but we sort of have to say what would constitute adequate evidence, what sort of level of risk do you have to rule out, how long, and things like that. Of course, we are in the process of learning about those things as these data come in. As I said before, I don't think anybody would have thought you need a four-year study, but that is sort of on the table now, and it wouldn't have been before. So, it is helpful to know what kind of risk is plausible to rule out and all the things that Steve said before, I mean you have got to worry about what doses to study.

We encouraged everybody to study high doses to rule out GI distress. Whether that was wise in retrospect, I am not sure, and I think I probably had something to do with it a long time ago. I am not sure that was the best thing. We want to be really sure you couldn't make an ulcer. I, at least, wasn't thinking about maybe making something else. So, all of those questions are things we need help with even though, yes, it's the company's job to bring the data forth.

DR. WOOD: Dr. Farrar.

DR. FARRAR: In terms of the specific question that we are addressing, I also just want to point out that there is a second part to that, which it says, also, discuss the possible mechanisms of action for an increased cardiovascular risk with these agents. I think that has

bearing related to the fact that if we accept that the in vitro selectivity of the COX-1/COX-2 analyses at least have some bearing on at least their metabolic process.

What I am struck by is the variability of the agents with regards to other factors indicating that simply COX-1/COX-2 inhibition is not their only action, and that the ones with some of the higher risks are not necessarily the most COX-2 selective. What that suggests to me is that we really don't understand the process yet well enough to be able to say that it is a group selective, because I am not sure what a COX-2 selective one is. Where do you draw the line? I think a more appropriate way to say this is to say that clearly, the role of COX-1 and COX-2 inhibition are important in the process of both anti-platelet and perhaps platelet aggregation, and that those with a more predominant COX-2 component need to be studied carefully for the potential excess cardiovascular risk. I have a great deal of difficulty, though, saying it is the group of drugs that have been called that by the pharmaceutical industry. I think that that is being very short-sighted about this. In fact, the data that we have seen in these presentations make me want to go back and look at ibuprofen with regards to a whole host of issues that we hadn't thought of before.

So, I think it is very important that we keep in mind that there is not a distinct relationship between those numbers specifically and that we need to be a little bit broader in terms of that look. The other point I would like to make is that we clearly need to differentiate in terms of what we are considering between the placebo trials, which have

been done primarily in cancer prevention, and the comparative trials with other agents. As has been brought up many times, none of the agents are the same, and so the comparisons there need to be carefully considered.

As such, I am in favor of a statement that says that we are consciously aware that COX-2 is an important component of this issue, but that all agents that claim to have, really all agents that are developed in the future and all the current agents need to be carefully looked at for the balance between the cardiovascular, GI, and other risk factors including the hypertension, including the pulmonary edema that we have heard so much about.

DR. WOOD: Right. One pragmatic way perhaps that we could handle this is there are two drugs of whatever the class we are talking about is that are left on the market, and for which we have a number of randomized trials recently, and we could consider them as a sort of present tense evaluation, and for future tense, other drugs that may have signals that we don't really understand, and certainly drugs that were likely to be marketed in this area, this space, and whatever that means, and would need some sort of evaluation. So, that would sort of divide up our work, so that we would be considering what to do about the ones that are out there, what to do about the ones that are potentially out there, and I guess a third group is what to do about other drugs that may or may not fit into this class or may not be at some extreme of this class. Is that sort of capturing the essence of what you are saying?

DR. FARRAR: That is certainly one way of dividing up the work.

DR. WOOD: Let's think of it in terms of that as we move forward. Dr. Holmboe.

DR. HOLMBOE: Actually, a number of the things that I was going to say have been said. I would just add one caveat, Tom, to what you said, that if you look at these trials, over 40 percent of the patients never made it to the end of the study, which means that we probably don't have over 70,000 patient observations. We probably have about 20- to 30,000 less who actually made it to the end of the trial. That, I am very concerned about, and you look at these trials also, although they look similar when they are first randomized, that is the purpose of randomization, the populations that get to the end don't, so I think that we are also lacking some very important information, what happened to a fairly large number of individuals who started but never got to the end of the trial.

DR. FLEMING: Just to respond to that: That is a key point. Now, the analysis that I did yielded approximately 7- to 800 events, so we are getting the total number of events that we would have needed from a 70,000 person trial, but your point is still well taken. We are not underpowered because of the lost to follow-up, but there may be a bias here that we all talked about earlier, that if you really wanted to get the most insightful, reliable assessment, you need to have high quality follow-up, so that is something that is still a relevant point. They are unequal in their quality of study conduct in the area of follow-up.

DR. D'AGOSTINO: This was part of my concern in terms of what I was trying to raise, that we have studies, but we don't have studies, there is a lot of problems with it.

DR. WOOD: Dr. Dworkin.

DR. DWORKIN: Ralph, I have a question to follow up on what you were saying earlier. If I understood you, you were saying that you are uncomfortable with a global statement of the sort that Dr. Gross was making because you feel there is some kind of heterogeneity amongst the data, of the type that Tom summarized. But then it seems to me you are between a rock and a hard place, because if you believe that there is a great deal of variability in the results with respect to risk, how could we possibly discriminate amongst the different drugs.

DR. WOOD: We have done that lots of times before.

DR. D'AGOSTINO: The qualities of studies are different. I think the arthritis studies is where we get the CV information, they weren't designed to get the CV information, cardiovascular information, so I think there is a signal there, but I don't know how to interpret it. I think there is the problem of lost to follow-up and things of this nature, and all of those things make me very uncomfortable, and sort of making a global statement and then living with that global statement. I think it is clear or hopefully it is clear what I am concerned about. We don't want to be locked into, by making a global statement, later on saying that no matter what drug we look at, we have an answer for, and it may be a low dose of Celebrex, that may be

viable, and not unsafe. We really need to worry about the studies that we are going to suggest, and if we absolutely thought there were safety problems, why are we suggesting them, why aren't we just saying stop the studies and get the drugs off the market. I think there is a lot of room for maybe there is something going on that is safe, and we want to really pin the issues down in good clinical trials.

DR. WOOD: Well, we have dealt with drugs within classes before. I mean a statin was removed, but the other statin is on the market, and troglitazone was removed, but the other drugs stayed on the market. I guess the difference here, which is only fair to point out, is that this is thought to be producing toxicity through the primary mechanism of action. At least that is one of the postulates, but we certainly should deal with them as individual drugs, I think, rather than as a class of drugs. Dr. Abramson.

DR. ABRAMSON: I guess what I am thinking, it is possible to accept the fact that many of these toxicities are via the COX-2 mechanism, but recognizing that all of the class of NSAIDs, by definition, when they are effective, are inhibiting COX-2, and I am still troubled by the population data which shows signals with indomethacin and meloxicam, and by older data which shows congestive heart failure particularly with the non-selective drugs. So, I think that we have to look at again the entire class, and particularly if you look at the CLASS and the TARGET trials, why is ibuprofen and diclofenac behaving pretty much like Celebrex and lumiracoxib, so if there is an assumption on our part that this class of drugs, even the highly

selective COX-2s, increase by 1.4, 1.5 the relative risk, why is ibuprofen and diclofenac looking pretty comparable in those large population trials. One answer is that they, in themselves, whether diclofenac is rofecoxib lite or not, but they themselves are imparting a risk, but they themselves have not been subject to these long-term placebo-controlled trials that we see in APPROVe and ADAPT. So, therefore, I think the COX-2 mechanism may pertain, but it cuts across all degrees of relative selectivity.

DR. WOOD: Dr. Furberg.

DR. FURBERG: Well, I spent about five years looking for a definition of class effect, and so far I have been unsuccessful. There is in the literature no definition of class effect. The closest I came was an FDA definition of class labeling, and that was not a good one. So, I think the working definition of a class effect would be that members of a particular group or class share common actions in the broad sense, and I think that would apply to the COX-2s in my reading. They provide pain relief, GI protection, raise blood pressure, cause fluid retention, have the undesired effects on cardiovascular risk, so in my mind this is a class and sharing a lot of actions, and that would include the increased cardiovascular risk.

DR. WOOD: I am going to take two more questions on this topic and then I would like to move us to, I guess, considering which drugs, to answer Question 1, which drugs, rather than a class, which drugs we see a cardiovascular signal with, which is one way to approach the problem. You are the next question, Steve.

DR. NISSEN: What I wanted to make sure we got to is this issue of mechanism, which is actually in the question here, and the reason it's important is that I am not quite ready to accept the hypothesis that one can predict from the COX-2 selectivity and duration what is going to happen in these drugs. Let me see if I can explain that because I think it is very important as we think about how to go forward here. I see a broad spectrum of blood pressure changes that don't seem to be as tightly linked to the COX selectivity as one would guess. Lumiracoxib, for example, which is very COX selective, doesn't appear to have much effect on blood pressure. Rofecoxib has the largest effect on blood pressure by far and is relatively COX selective, and they are very different. So, for the FDA, I think if you want to characterize the drugs, not only do we need clinical trials around looking for GI safety and cardiovascular safety, we need a standardized method to look at the effect of these drugs and their intended doses on blood pressure, and they ought to all be subjected to similar scrutiny, so we can compare apples to apples and wherever possible with active comparators, let us understand that. Now, why do I say that? Because Bob and I have sat at many a meeting and looked at blood pressure drugs, and I can tell you the data on the relationship between relatively small differences in blood and cardiovascular morbidity and mortality is rock solid across a huge number of drugs and interventions, and you almost can predict what will happen. So, we need to know--and as I sit here, I can't tell you that drug X in this class has Y blood pressure effect and drug A has B blood pressure effect--and so we don't know, and we can't inform physicians about that unless we have better data on

blood pressure. So, I am making an appeal that we get to that level of specificity, and that is not a very big trial to do that. Bob, what do you usually ask for in the blood pressure study?

DR. TEMPLE: Well, if you use automated pressure monitoring, I think you can get a decent answer with 20 or 30 per group, maybe 40. It's very easy.

DR. WOOD: Let's move on. I am going to give you the last word in a second. After we get Dr. Gross's comment, we are going to divide this first question into three things, which drugs do we see a cardiovascular effect of, and the secondly, we can ask whether we see a class effect, whatever we understand that at, and I am not sure we do, and then the third question that is in Question 1 is what do we see as a mechanism. So, let's divide them into these three things and let's move to an answer. Peter, last word.

DR. GROSS: What we say here today about these is going to have a significant impact in molding public perception, and if we conclude that there is a class effect for the selective COX-2 inhibitors, and don't say the same thing about non-selective NSAIDs in general, then people are not going to want to use COX-2 inhibitors at all and they will be using the non-selective NSAIDs, which from the data presented, doesn't look as though many of them are better from a risk point of view. So, I just issue that note of caution.

DR. WOOD: Does the FDA want us to go around the table asking people for an answer to each of these questions, the subsets of these, John?

DR. JENKINS: I think we really viewed these questions as things to stimulate your discussion, not necessarily things that are amenable to yes/no answer. The yes/no answers come tomorrow.

DR. WOOD: So, can we move on?

DR. JENKINS: If you think you are done with No. 1.

DR. WOOD: We are done with No. 1.

DR. TEMPLE: I just wondered if people could come to grips a little bit with some of what Steve said and some of what other people said. I absolutely don't want to put words in anybody's mouth, but what I heard people say was they think the class has at least the potential for having this problem because of the imbalance and because of the stuff we have heard about before, and that you need to look at each drug to see whether that is manifested at a particular dose-dose interval and all the rest. I just wondered whether that is getting close to what people are saying or not, and I absolutely am not giving my view on it, I am just suggesting it.

DR. WOOD: Let me try and answer that and then we will go around and ask other people. I think what we are saying is almost the same as the GI effect before the GI effect or not was worked out for the so-called COX-2 inhibitors, that I see an effect, a cardiac effect from valdecoxib, certainly from Vioxx, and from celecoxib, and there is a dearth of data on the non-steroidals, the other non-steroidals at this stage in terms of cardiac safety, and we are not going to be able to decide that even on Friday, it seems to me. In the presence of that signal, the prudent activity would be to

go look at it sometime in the future, but we can't do that between now and Friday night. So, that is sort of where I come down. Dr. Abramson.

DR. ABRAMSON: I might have a slightly different view, because I mean I think of the class more broadly as it is defined now to include both the COX selective and non-selective drugs. I think that there is a signal probably for all of these drugs, maybe by different mechanisms perhaps. I think we have under-recognized that in the population, I think physicians have not been concerned enough about blood pressure changes. So, my view is that maybe there will be different mechanisms, but that each of these drugs is suspect as having an increased relative risk when used chronically, whether it is ibuprofen or the most selective COX-2. My own view is, as I said earlier, is this is not dissimilar to the late '90s, and until you prove otherwise, this is GI warning that these drugs may cause cardiovascular risk or GI warning it may cause serious adverse events, and I think each of them should be held to that right now until someone proves otherwise, because I think it would be wrong based on the evidence to assume that three drugs have a cardiovascular risk, and several of the others don't, simply because we don't have the evidence. I am also concerned about some of the research that was talked about by one of the public speakers. At most of our universities, these studies have actually stopped because of concern that these drugs are not as safe than the non-selective drugs. I think, particularly in cancer and others, we are doing the public a disservice by prematurely picking out these drugs as being unsafe and stopping some very important research where the risk-benefit

might even be more important than in arthritis.

DR. WOOD: Dr. Shapiro, I missed you, I am sorry.

DR. SHAPIRO: That's okay. I think I agree with you, and I think it is hard to properly answer this question unless we ask ourselves why it is being asked, and if it's being asked because the FDA wants some broad-brush, uniform regulatory approach to this, what I hear people saying around the table is that that would not be appropriate for each and every one of the drugs that are in this possible class. But if we are saying that we think that drugs that are related in composition, structure, this, that, and the other thing, should raise a red flag, which is what I think you are saying, that is what I think we want to say, and I think we are getting hung up on this class effect definition because we haven't gotten behind and asked why we are being asked the question.

DR. WOOD: Dr. Furberg.

DR. FURBERG: I think I disagree with Steve Nissen, and I think it is a mistake to focus on one mechanism of action. Members of a drug class, they don't have to share all mechanisms of action. In fact, I don't know of any drug class where all the members share all mechanisms of actions, so the term is more loose and relative.

DR. WOOD: Dr. Shafer.

DR. SHAFER: Actually, I think Dr. Nissen said that it's not all one mechanism. I think that is exactly your point.

DR. NISSEN: Exactly my point. My point, Curt, was that these drugs do differ by some of them have much more of pressor effect than others, and that seems to be dissociated at least somewhat from their COX selectivity, and so I want to characterize the drugs individually, not necessarily collectively.

DR. SHAFER: Continuing that same line of argument, Bob, in answer to your question that you had raised, if this was all the FitzGerald hypothesis, then, the class effect makes a ton of sense, because you would say okay, you look at the COX-2 selectivity, we kind of go on the list, and we do our cutoffs. We have the blood pressure data. I point out once again we do have the aspirin data in some very big trials. The effect should have gone away in the presence of aspirin particularly I point out again to the APPROVe trial, the thrombotic risk was 3.25. We have talked about this on and off, and you haven't been feeling well, so we haven't had a chance to really get together and discuss, at risk of my health, despite having lunch at Chuck E Cheese, but I am concerned because we haven't explained the aspirin effect, and aspirin, unlike the other drugs, doesn't go away, it doesn't have a pharmacokinetic component. I mean that should have clearly made a statement if the aspirin effect had reversed these prothrombotic effects. Steve, I think that argues to your point that there are several mechanisms. One is certainly in part the FitzGerald hypothesis although there is partly a class effect, but the aspirin also shows that there is something else going on.

DR. TEMPLE: Then, how do you characterize the class? I mean it sounds like most people think you are

characterizing the class as one with a preference for the COX-2 receptor, for that one, and if you can't do that, it is hard to know how to go forward.

DR. SHAFER: Can I answer that?

DR. TEMPLE: But I agree with you about the aspirin, it's a fly in the ointment.

DR. SHAFER: It seems to me that you can look at where we have data that is consistent with the FitzGerald hypothesis, that is consistent with it, that you can say these drugs are behaving as in class, and certainly for the coxibs, as Dr. Fleming presented the data, it appears that they are all behaving in a way that is consistent with a class effect. Where we don't have more specific data that would say they are behaving in this fashion, and I would point out these are the COX-2s that we don't have data because they are older drugs, but they appear to be COX-2 selective, I am reluctant to include those in the class and sort of damn them because of where they show up on some table. At the same point in time, I am reluctant just to give them a get-out-of-jail-free card, if you will. I think that something needs to be noted that they are potentially at risk for this effect.

DR. WOOD: Dr. Cush, then Dr. Hennekens.

DR. CUSH: I would support what Steve said and that I think that we came here with the spotlight focused on Class II specific agents, but we become more curious as we have seen all of them fall, but then seen all the other drugs, the non-selective drugs also seem to have some of the same failings, we don't want

to focus solely upon the COX-2 specifics, but I think that we can start there and then extend our concerns to the other agents, as well. It doesn't have to be, it can be linked to COX-2, and that may be where we start, but it obviously needs other investigation to look for a mechanism of action.

DR. WOOD: If Raymond Lipicky were here, he would say: Show me the data that tells you that these other drugs have this effect in published trials.

DR. CUSH: Well, one would be I guess some of the observational data.

DR. WOOD: Randomized, published randomized trials.

DR. CUSH: Well, I think the only one we really have is the Norwegian study.

DR. WOOD: Is that a randomized trial?

DR. CUSH: I believe it was. Well, they were randomized to--

DR. WOOD: That showed aspirin also had a negative effect.

DR. SHAFER: Alastair, that is the reason the Challenger blew up, the sort of show me it's safe, prove to me it's safe or I am not going to make a statement.

DR. WOOD: That is not the issue at all. I mean we have got to be careful, I think, rushing ahead of credible data on the basis of rumors of war that are brought in from outside. I mean we have got four randomized and controlled trials for three drugs, and we have got some news of other drugs, it seems to me, that are not--and documented very well. That is not giving anyone a get-out-of-jail-free

card, but I think we have got to sort of go through this in an orderly fashion. Otherwise, we will be regulating on rumor forever, and I think that is a very dangerous step to take.

DR. TEMPLE: You do have some diclofenac data in comparison to some of the drugs that are of interest, so you have some. It's not the placebo-controlled trial you are dreaming of, but you do have that, and you have naproxen and several comparisons, as well.

DR. WOOD: And that looks pretty good.

DR. TEMPLE: Naproxen looks good, ibuprofen looks the same as--there are, I didn't count them up, three or four control groups of the older ones scattered around.

DR. FLEMING: Well, we can be specific because Bob is right, we do have--I mean basically, because of all of these other studies that were done for the COX-2 inhibitors, there is a lot of data on naproxen and a lot of data on diclofenac, and diclofenac in the etoricoxib trial and in the CLASS trial more or less came out looking like the COX-2 inhibitors, while the winner is naproxen. Basically, in the VIGOR trial, in the etoricoxib trial, very much in the lumiracoxib trial, it came out positive. Now, we are going to hear something tomorrow about the ADAPT, but looking at these others, it sure looks like naproxen is a winner, and it does look like the theory that was put forward that diclofenac is COX-2-like is at least supported by the trials where it was studied.

DR. WOOD: Right, but all we can say is they look the same as another drug

where we are not absolutely certain of the effect of that other drug.

DR. FLEMING: That is true although we have a lot of other studies on the other drug, and it is always you have got to be careful when you say A is better than B, and then B is the same as C, is C worse than A, but there is that kind of evidence.

DR. WOOD: Which is what I am concerned about. Dr. Hennekens.

DR. HENNEKENS: I would say I am struggling with trying to gain this clarity, but as I view the drugs that either have been or are marketed with regard to cardiovascular risk, the picture that emerges, begins to emerge to me is that rofecoxib, ibuprofen, and possibly valdecoxib are in one bin, diclofenac and celecoxib in another bin, naproxen in a third bin, and then aspirin in the fourth bin, going from concerns about hazard to neutrality to benefit.

DR. WOOD: Other comments? Dr. Farrar.

DR. FARRAR: Two quick--well, I guess every time we mention aspirin, it never ends up being quick--but two quick comments, one of which is that I am not as concerned about aspirin knocking out the issue of the COX-1/COX-2 problem primarily because, in fact, aspirin is a surrogate marker for people with cardiovascular disease. If you look at the actual rates in all of the aspirin groups, they are at least, at least 2 to 3 times the rates in the non-aspirin groups to start with. So, I think that there is an issue there. I think the second issue has to do with what was just discussed in terms of the comparison of drugs, and just to

emphasize the fact that what we are talking about is we have data for there being a risk factor in the placebo-controlled trials, primarily the best data, which we will have a whole lot more of in two months or three months, of the cancer prevention trials to tell us what the level of risk is. Then, we have the comparison data that Bob Temple was just talking about in terms of the non-selective versus the selective that say that they have very similar levels of risk. The third point just to make is that all of this discussion about risk, I don't want to imply that I think that this risk is big enough to actually warrant the continued hold on all the trials that we have going, and I think, in fact, what it suggests to me is that we need to continue with trials to understand better what the data is telling us.

DR. SHAFER: May I respond to the aspirin point? This confusion that you raise came up when I first raised it, I guess it was just yesterday, but the risk that we are talking about is not aspirin versus non-aspirin, because clearly, aspirin will be a marker for increased risk. What we are talking about is the risk of rofecoxib in the case of APPROVe, the risk of rofecoxib versus the comparator in those patients taking aspirin, so that the increased risk of cardiovascular events has been evenly distributed between the two groups, because that is the blinded comparator variable. So, we are talking about the risk of COX-2 versus non-COX-2 in those patients on aspirin. It is different from the risk of aspirin versus non-aspirin, which as you say is, of course, that risk is confounded. But in this case, that risk is evenly distributed between the two groups.

DR. WOOD: Go ahead, Dr. Farrar.

DR. FARRAR: I think the problem is that what you are saying is that aspirin is somehow only a COX-1 inhibitor and therefore it has a role there that somehow should balance the COX-2 or there should be some other process going on. There is no question that aspirin and its indication of increased cardiovascular disease has an effect on the relationship of the COX problem. We have seen multiple examples in the cardiovascular risk, in the group who have the high cardiovascular risk, there is a different response to the COX-2 problem than in the lower group, so there is no question about that. But I would argue that aspirin is as very different drug in terms of how it works, in terms of its binding to the sites, so all I am saying is that I am not sure that that obviates the need to say that there is an issue there with COX-1/COX-2 that we need to look at more thoroughly.

DR. WOOD: Dr. Manzi.

DR. MANZI: I actually have a problem making inferences about diclofenac and naproxen in studies where I think we have a difficult time feeling comfortable with the results in relationship to the COX-2s. I mean the trials that are really

driving the signal here are the placebo-controlled trials of long duration. So, to feel that we don't have enough information to really feel comfortable with COX-2s, and then to try and extrapolate to the comparators in those, I think is dangerous.

DR. WOOD: That is what I was saying, too. You know, it's ten past 5:00, just to draw everybody's attention to that. John, you are saying that you don't necessarily want a vote on this, is that right? So, I guess the question is, is there further discussion on this specific question that the committee feels they can't hold until tomorrow? Tom?

DR. FLEMING: I share the caution in that last comment, but I will just note that methodologically, it is the exact problem we run into or situation we run into in non-inferiority trial designs, because you have placebo-controlled trial of agent A, and now you want to look at whether B is adequately safe, and you are looking at B against C, the new agent, and if C is the same as B, that was shown to be non-inferior, or you knew what its relationship was to no treatment, it is that non-inferiority issue. Nevertheless, many of us have concerns with non-inferiority settings, but that is the methodologic challenge.

## ***Question 2 Discussion***

DR. WOOD: That is my concern, as well. Let's move on Question No. 2. We may have discussed this a lot already, but this really addresses the contributions and limitations of the currently available observational studies to the assessment of cardiovascular risk for the non-selective and COX-2

selective--and let's not bog down in what we mean by that. In particular, discuss the role of such observational studies in informing regulatory decisions about postmarketing safety issues. Now, let me ask a clarification question. Does this mean we just sort of ignore the randomized trials here or take them as a

given, or how do you want us to handle that?

DR. JENKINS: I think the idea here was to get your thoughts on how we should consider and weigh these studies in a mixture where we have some control trials, we have the observational trials. Sometimes they don't agree with one another. Sometimes the observational data come at a time when we don't have the control data. We are trying to get your take on what weight should we place on these data as we are trying to make regulatory decisions.

DR. WOOD: So that we could modify the question to sort of include the randomized trials and say how do we relatively assess these and weigh them up?

DR. JENKINS: Sure.

DR. WOOD: All right. So, that is a helpful clarification. Comments on that question? Yes. Dr. Stemhagen.

DR. STEMHAGEN: A couple things. I think I want to make sure that it is understood, in my view, that they are definitely supplementary to the randomized clinical trials. I think we all recognize that the value of randomized clinical trials is the randomization, that we don't have the selection bias that certainly takes place in observational studies, but nevertheless, when we think about the magnitude of the studies that we have, we have over many hundreds of thousands of patient years of exposure, we have in the cohort studies.

In the case-controlled studies we have more than 25,000 cases. We do have a very rich data set. I think we have talked

a lot about the fact that we have got a number of studies and we see a lot of consistency in the results between those studies. There was an issue of maybe they are all biased in the same direction. I think they were conducted in very different ways, many of them, and many very different databases.

We also see some data on dose response, which is another suggestion that there is something going on and that the data should be believed.

I think if we talk about lost to follow-up in some of the randomized trials, in some of the very stable populations that we have in some of the databases, we actually do have long follow-up, although ideally, we would like these studies to go on longer. None of them are really as long as we would like, and part of that I think is the data being on the market or available within those databases at the times that the studies were done.

Another thing that really is different with these studies is we are not just talking about volunteers. When we do our clinical trials, we are talking about volunteers. In our databases, we really have the totality of patients, of cases, of exposures. So, I think we have got somewhat different groups of patients. The clinical trial patients are essentially a subset to that.

We also are looking at actual use doses, which are somewhat different doses perhaps than in a lot of these clinical trials where we have talked about high doses are pushing the dose.

So, I think they are different pieces of information. The endpoints that we are

looking for are very hard endpoints, and I think we have talked about, and there was some evidence, that in some of these studies, there are adjudications, the same way there are in clinical trials when the medical records are collected. There have been some validation studies looking at the ascertainment of MI and feeling that it is very complete. So, I think we can feel reassured that in these closed populations, we probably have identified the cases that we are interested in, and we also have a lot of data, not necessarily exclusive, on the confounders, and there have been adjustment for confounders. So, I really want to urge that when we look at the data, we don't just dismiss the randomized clinical trials, but they are telling us something. They do have some patterns, and they do show some differences between the products.

DR. WOOD: Dr. Cush.

DR. CUSH: I think there is obviously a value for observational studies, but one thing I keep hearing is that the FDA is not properly empowered to mandate that postmarketing trials be done until maybe a significant issue like this comes up. This kind of public health issue sort of underscores some of the weaknesses of the current MedWatch system where common events like this are not going to get reported on new drugs, because people get heart attacks and heart failure and uncontrolled hypertension, and I think that one thing I would like to see come out of this is that Congress and others empower the FDA, so they can do postmarketing trials that need to be done, either mandate it or as they need to occur, and if they can, mandate registries as they need to be done, as well. That is certainly right now what I think is a big

hole in our current safety system. We heard today from the patients, they want to know that we are going to help them. That mainly means they want to know that we are going to give them medicines that are safe.

DR. WOOD: Dr. Bathon.

DR. BATHON: I would like to take the example of naproxen for a minute where it seems like from observational studies, it has a neutral effect on cardiovascular risk, at least that was the overwhelming notion, whereas, in randomized trials it seems to be more protective. I would like to explore for a minute why that discrepancy, if it is true, why it might be true. I would like to posit that in the randomized trials, we have people taking drug every day or at least we think they are taking it, and they are taking it in the appropriate dose to have consistent COX-2 or whatever, COX-1 and COX-2 inhibition. In observational data, those are driven, NSAID drug use is driven primarily by acute pain syndromes and osteoarthritis, where people, if we go to the acute, somebody has back pain for a few months, a lot of the people using those drugs might be on them for a few weeks or a few months. The proportion of patients like the rheumatoids or the bad OA patients who might be taking them every day is probably relatively very small in that group. Even within the OA group, I think a lot of us probably have OA in here, some of us who have gray hair or getting gray, even the OA patients do not take the drug every day on average. The rheumatoids tend to, the OA patients don't, and then the acute pain syndrome people or the back pain are more intermittent. So, I wonder if the difference between observational data and the clinical trial is driven by the fact

that we are looking at very different treatment regimens, treatment durations, and so forth. So, I think the randomized trials are more valuable here than the observational data.

DR. WOOD: Dr. Holmboe.

DR. HOLMBOE: I would just make a couple of points. If we agree on No. 1 that there actually is harm, then, I think yes, you are going to have to do observational studies. I mean it is going to be hard to randomize somebody to study harm. I think that we can take some comfort even though that the effects are different, that the observational trials were reasonably consistent with a lot of the randomized controlled trials that were presented today. Second, I think a poor randomized controlled trial actually may be worse than a good observational study. As I mentioned earlier, a number of these studies had over 40 percent dropout with these patients not being followed, and I think that that is an opportunity for the FDA to follow these people out to see if there is something inherently different about those populations who aren't continuing on the study drug. The third point I would make, that with regard to meta-analysis, it is very important that the trials be fairly homogeneous in the way they were done. In all the stuff reported, I did not see anybody talk about a test of heterogeneity to see if they really truly could be combined. While I understand that because the events are so low, you are trying to pool risk, there is some danger in pooling studies that are quite disparate. So, I think that is something that needs to be taken into consideration. The last thing I would say is that I think there is a real lesson here potentially for

the FDA. The comparator drugs were approved before we truly understood the biologic mechanism of these drugs. Our understanding of COX-1 and COX-2 occurred long after the original comparator drugs were approved. So, it is a real challenge I think for the FDA to go back and say wait a minute, could these comparator drugs potentially be a lot like the drugs that we are now studying, that we think are being proposed as different, but, in fact, may not. So, I think that that is real lesson, it has created a lot of the confusion we are now having to deal with, because a lot of the comparator drugs it turns out actually are very similar to the COX-2s that we are evaluating.

DR. WOOD: Dr. Day.

DR. DAY: Concerning the 40 percent dropout rate in the randomized trials, we have all the sponsors here, and they have lots of data and computers, and so on. Would it be useful to get the percentage dropout for each of the target drugs and the comparators and/or placebos in a giant chart before tomorrow to see, and then try to get a breakdown of what the reasons were for dropout? Do they retain that information when a patient drops out, what the reason is, or is that on file somewhere?

DR. TEMPLE: They always provide it. The question is how reliable it is. A lot of them say administrative reasons, and it really requires people to pursue that question, interview the patient, and while that is properly done sometimes, it isn't by any means always properly done.

DR. DAY: So, the breakdown isn't possible. What about the percentages for

each of the groups that we have seen just in these studies?

DR. TEMPLE: Pretty much all studies know how many people stopped and completed and when.

DR. DAY: Do we know? Have we been given those data?

DR. WOOD: Well, I guess the Kaplan-Meier curves, and under each Kaplan-Meier curve, I think there is a number of patients at each point.

DR. D'AGOSTINO: Part of that was dropout, but part of that was the way they planned, you know, follow-up on the individual. The individual could, for some reason or another, say they are not going to take the drug anymore, and they only follow them 14 days, so that was legitimate in the study. A dropout that just disappears was sort of illegitimate, that was not split up.

DR. WOOD: It is still a dropout, I mean the person didn't complete the study.

DR. D'AGOSTINO: Well, it followed the protocol. I mean you can't now go back and say they should have done something.

DR. WOOD: Dr. Paganini.

DR. PAGANINI: One of the things that I was surprised at here was the lack of information on the older NSAIDs, and that is one of the things that we are trying to deal with is: What is the difference? That then speaks to continued observational studies in the post-market venue where if we had had that, we would have at least had some sort of observational anchor to put some

of the newer drugs on. Let me also add that while we always look at prospective randomized controlled studies as being the be-all and end-all, there is now an emerging--and I will ask some of the biostat folks to comment on this--a developing thought process of having a wild arm, the wild arm being what is usually and customary done when doing something. For example, if you do a dose of a drug, or you do an amount of O2 delivery or some sort of a respiratory issue in the ICUs, frequently, when you enter into a study which is randomly controlled, you have one arm versus the other arm, and they are fixed arms, but there is now a third arm that people are starting to ask for. It's a wild arm, what do people usually do outside of the study, and I think that is a very important issue for when you are using drugs in a common, out-of-the-box way where everybody is using the drug. So, postmarketing observational studies might be considered the wild arm for some prospective randomized controlled trials in that same era.

DR. WOOD: Dr. Nissen.

DR. NISSEN: It is interesting. We like our observational studies when they show us what we want to see, and we just hate them when they show us what we don't want to see. I have lived through this with the estrogen business. I had people tell me that it was absolutely unethical to do a trial of postmenopausal estrogens because everybody knew they were beneficial, every observational study had shown it. So, it is important that we use observational studies as hypothesis-generating studies. If you see a signal in an observational study, it is an indicator that you need to do a randomized controlled trial, and that is

how we ought to use them. If we get too far beyond that, we are going to get into the women's health initiative kind of problem again. It comes up every generation as another example of this, where every observational study tells us one thing until we do a randomized trial, we find exactly the opposite.

DR. D'AGOSTINO: I want you to recall that the Framingham studies said just the opposite, it was the observational study that didn't agree.

DR. NISSEN: Thank you, Ralph, you are usually right.

DR. WOOD: You were down next to speak, Ralph, is that your question?

DR. D'AGOSTINO: Oh, is it my turn for my question?

DR. WOOD: Yes.

DR. D'AGOSTINO: When we have placebo-controlled trials, randomized controlled trials, I mean in some sense it is I think the gold standard, and when you have positive comparators, randomized controlled, it's the next level, I think that we have a lot of data that is well developed in terms of the studies. We have questions about the dropout, and so forth, and I raised them also, but I think the randomized controlled trials have put us in the situation where we can minimize in some sense the observational studies. Yesterday, I made my comment about torturing the data. We can torture the observational studies forever and ever, but I think our weight should shift on the placebo-controlled trials.

DR. WOOD: I agree with that. Dr. Fleming.

DR. FLEMING: Maybe just to be specific here about different kinds of observational studies, there is passive surveillance and active surveillance. Passive surveillance has been widely used, for example, in vaccines, childhood vaccines, and with the Veer system. Essentially, it worked really well when you are trying to detect rare events and events that are proximal to the time of the intervention. So, intussusception with rotovirus and encephalitis, and anaphylaxis, et cetera, have been assessed fairly well.

The problem with those, and we heard naproxen experiences in what I would call passive surveillance, the problem is if you have events that occur with more regular frequency in the background, it is going to be almost impossible. There is under-reporting, you don't have denominators. So, a step up is the large-linked databases or the active surveillance systems, and I think this is what a lot of what we have been talking about with these observational studies. They give us numerators and denominators, they give us more complete ascertainment, but they still have unavailability often of confounder information on aspirin use, smoking, outcome specificity and sensitivities are less reliable.

We have talked earlier today about how it is extremely difficult in that context to do a valid ITT type analysis and have a time zero cohort and minimize lost to follow-up, and ultimately, you are not randomizing, and randomizing doesn't solve all problems, but it does, in essence, eliminate the systematic

occurrence of imbalance. It doesn't eliminate randomly occurring imbalances until you have large numbers, but you cannot, with covariates, go back and adjust for what is different in an observational study, because I always say the known and recorded covariates are just the tip of the iceberg, so you are left with a great deal of uncertainty about bias.

Where they are very effective is understanding natural history, understanding event rates, understanding covariates, understanding how people are treated, but we really want to use them to understand causality, does intervention have an effect. Essentially, if it is a very large effect, you can get some reasonable senses, but in most cases, I think they serve a very useful purpose, but it's hypothesis generation, it's development of clues. So, if we look at the overview that David Graham gave, my sense is he was able to give us insights about a wide array of issues that we have not yet got adequate randomized trials, so specifically, the nonspecific NSAIDs, what does it look like there, and issues about dose, but I would call those hypothesis generation or clues. I would be very reluctant for the majority of what we saw from those analyses to take those results as established. It rather gives us a guide because we can't do randomized trials in every setting. It gives us a guide for how to design those trials and where the most pressing questions are. So, the observational studies go hand in hand, but the ultimate answers in most cases really come from the randomized trials.

DR. WOOD: Right, and the estrogen studies shouldn't be forgotten, right? Dr. Morris.

DR. MORRIS: I think Tom said a lot of what I wanted to say, but a lot better. In terms of causality assessment, living through what the Agency of Healthcare Policy and Research went through for outcomes, I think the conclusion is unless you randomize, you are never really sure. In terms of observational studies, I think it is interesting that like event rates or something like that, where we think it is so much better, yet, I was struck in the discussion today of some of these drugs is how much the event rates varied by center or study or country. What isn't done in observational studies, what could be done, is more of a population-based sampling, so we have a better understanding of how much or how well that particular database is representative of the broader population of the U.S., so we can do some kind of sampling or extrapolation and get much better event rates, where I think observational studies can really do a much better job than clinical trials because they can measure naturally occurring events much better.

DR. WOOD: Dr. Domanski.

DR. DOMANSKI: You know, one always hates to admit ignorance, but I want to pursue this business of a wild arm. I mean I have seen some pretty wild arms in clinical trials, but never as a third one. I don't understand where that is, I have not heard of that one, and I would like to learn more about it. Can you explain that?

DR. PAGANINI: I will give you an example of an NIH-VA study that is now ongoing looking at dose of dialysis delivered in which there is a high dose delivered and then there is a low dose

delivered. Then, there is the thought process of putting a third arm on there is what is everybody delivering anyway, so it is whatever the wild type is, to see if, in fact, people are artificially placed into one dose versus a second dose, and that, in and of itself, is an artificial placement of patients as opposed to what people usually do. So, therefore, what is the comparison between one dose versus a second dose versus what is usually and customarily done.

DR. DOMANSKI: But don't you usually use a registry for that kind of question, that is, how well does it represent practice I guess?

DR. PAGANINI: It could be retrospective, but in effect now what they are doing is a prospective collection of data of what is normally done in that particular institution when people are off study.

DR. DOMANSKI: Again, registries can be prospective, of course. I am having trouble seeing the difference. I mean are those people randomized, as well?

DR. PAGANINI: No.

DR. DOMANSKI: Okay, so it's a registry.

DR. PAGANINI: It's just a registry.

DR. WOOD: Dr. Hennekens.

DR. HENNEKENS: I would view the strengths and limitations of observational studies to be a function of the effect size. For the moderate to large effects, we can make safe clinical and policy decisions based on consistency of the data from the observational studies.

As the effect sizes get smaller, however, it's a two-fold problem because now the effect sizes we are seeking are as big as the amount of uncontrolled and uncontrollable confounding that is inherent in the designs. There is a certain seduction from these large-scale databases because you have a large number of data you control confounding on, you could get very robust p values, so you begin to believe that you have really discovered something, but I agree strongly with Tom that for small to moderate effects, they are useful to formulate, not test, hypotheses, so what Dr. Graham told us this morning are useful to formulate hypotheses. If people took them as serious evidence that this indicated harm, he might be right, but it would have nothing to do with the data that we have seen. I conclude with the statement, I had the privilege to know Sir Austin Bradford Hill who, on this question, and I think Rich would agree with this, he said, "Don't let the glitter of the tea table detract from the quality of the fare."

DR. WOOD: Dr. Elashoff.

DR. ELASHOFF: Two comments. One, in this situation, especially when there is very specific evidence that the relative risk may vary over time, looking at the standard way that observational studies lump it all into patient years is bound to be misleading. A second point has to do with the fact that in a randomized trial, when you are comparing events, the analysis per se tends to be pretty transparent, but in an observational study, in order to understand it in detail, there are many covariates, pretty fancy footwork in the statistical realm, and it may not be very easy to tell exactly what was done or to think of reproducing it.

So, the observational study tends to be a lot less transparent in terms of the way it has been analyzed.

DR. WOOD: Dr. Friedman.

DR. FRIEDMAN: Two points. One, if I can follow up a little bit on this “wild arm”, if you will. As Dr. Wood knows well, this whole issue came up, to my dismay, if you will, about a year ago when we were dealing with the ARDSNet issue, and I think the general conclusion there was that it, in general, is not a very good way of answering a specific question. It might contribute in some fashion, but in general, it is not all that helpful. Second, I am looking at the specific question here, and it says discuss the role of observational studies in informing regulatory decisions about postmarketing safety. It seems to me that one of the things we might do is suggest ways that the FDA can improve some of the postmarketing surveillance issues. For example, we have talked about all of the difficulties in using observational studies, and I don't disagree at all with any of them, but if some of them are planned ahead of time, with good ways of collecting data in consistent ways, we won't completely eliminate all of the problems, but we can reduce them, and I think we ought to at least consider that approach.

DR. WOOD: Dr. Platt, last comment on this.

DR. PLATT: To emphasize that point, taking everyone's thoughtful comments into account, it seems to me we have to be careful not to let the best be the enemy of the very good. I think that Tom Fleming's reference to the CDC's large databank for vaccines is quite on

point. It seems to me that there is every reason for FDA to require, as part of the approval process, that there be a substantial and organized observational set of studies that give at least a sense that generates hypotheses that would allow us to recognize the possibility that there is a signal of events that never be seen in clinical trials, events on the order of 1 or 2 or 3 per 1,000. It is possible to do that with what in the scheme of these discussions we are having would be a relative small investment, and we wouldn't have to rely on the occasional observational trial or the clinical trial that shows up to start a discussion like this. It seems to me that that is a very easy, relatively small step for CDC to take, to have every manufacturer of a new drug commit to doing a reasonable observational study.

DR. WOOD: But, Richard, isn't that the problem that Tom highlighted ages ago, that that sort of registry approach will pick up events that are relatively rare in the background, like devastating encephalitis or something like that relatively easily. But where you have got a background noise that is as high as MI, it is going to be extraordinarily difficult to pick that up from that kind of study.

DR. PLATT: Well, in the vaccine field, the large-linked database has been extraordinarily useful for things like febrile seizures after a DPT immunization, and that is a relatively common event. So, I don't take the point that you can't make reasonable observations about even relatively common events.

DR. WOOD: Bob, do you agree?

DR. TEMPLE: Well, just to make the same distinction you were making. You can look for intussusception or something that basically is very unusual, but how to find an increase in the rate in the rate of MIs requires a structured study and a plan to do it, and you sort of have to have a hypothesis or you don't know what to look for. It is totally different from liver, you know, from gross hepatotoxicity, which comes in through the AERS pretty well actually, maybe you could stimulate those, but it is totally different when you are looking at a change in something that has a high background rate.

DR. PLATT: The fact that it's challenging doesn't mean that you can't learn something useful, and it is pretty clear from the observational studies we have that we can learn something useful about that.

DR. TEMPLE: I was reacting to what you said, should we have the capacity or have the ability to get people to do studies once something emerges or once a question arises, or once you know something about the drug class, I am not challenging that at all, that's fine, but to have it in place as a mechanism for sort of automatically putting stuff up, I guess I don't know what that mechanism is. There has been talk about encouraging places to report, and we have an arrangement with some liver centers, and those things are fine. Those might be ways to find hepatotoxins maybe faster than we do now, but that still doesn't answer the question of a change in the rate of a common event, which is a fundamentally different problem, requires a study, not a report.

DR. PLATT: Well, the model of the large linked databases I think gets around the idea of having to have active reporting. I think that there is a lot of ability to capture the outcomes that are of interest. Obviously, you don't look for every outcome for every drug, but you can make up the list of things that you care about for certain classes of drugs, and it is possible to use automated systems to take you a long part of the way in understanding whether there is a problem that needs serious analysis.

DR. TEMPLE: Can I propose an alternative? I think what you are really saying is the thing you are worried about with drugs, where there is a high background rate of something, is always cardiovascular outcomes. So, I think what you are saying is you might want to look for any chronically used drug at cardiovascular outcomes, and that you could probably put in place.

DR. WOOD: Wait a minute. Are you suggesting that we insist on a cardiovascular study for every drug that we get approved? I mean that would make it prohibitive to approve any drug.

DR. TEMPLE: No, no. We used to fund more of them than we do now, that's a problem that other people will discuss, and certainly I won't, but we have access to databases, whether it's California Medicaid or whatever, and one can do that. It doesn't seem inconceivable to me--and I am talking about something that other people know more about than I do, so I should probably shut up, but I won't--I can imagine that a couple of years into the approval of a drug that is widely used, you could ask the question at certain sites, can we see an increase in cardiovascular risk. I am not sure how

many other high background events it is that are common in the population that we are really that worried about. Maybe that is something that we could think about.

DR. WOOD: So, if we could just sum up where we are, what the committee is saying, I think, is that we are impressed as the primary data source, and that the primary data source should be randomized and controlled trials, and observational studies may be good for hypothesis generation, and I guess the third point is that the AERS database is of almost no value in detecting adverse events that are common in the background in a situation. Is that sort of fair for what we have sort of got out of this? Do people disagree with that? Yes, Dr. Farrar.

DR. FARRAR: There is one specific point to this question, which is that all of the non-experimental studies that have been presented here, I would certainly suggest, and I would hope people would agree, are hypothesis generating at best. Every single one of them is confounded by indication. The best example is the indomethacin one where it is only used in people who are sicker than people who aren't. So, I think there are clearly examples. What I was hearing before was a discussion about what we might do, and I just wanted to be clear that what we might do is very different than what we have right now.

DR. WOOD: You put it much better than I did. That is what I was trying to say. Yes, Dr. Jenkins.

DR. JENKINS: I found this discussion to be very interesting because I think you all know there has been a lot of

Monday morning quarterbacking about what FDA has or has not done in this class, and a lot of that has been based on observational study results, many of which fall into the range of what we have been calling small to moderate, I think, at best. I don't think we need to revisit that here, but I think the questions we have going forward, first of all, we have to look at the data set we have today, and you have to look at the data set you have tomorrow on answering the questions about what do we do now, and observational studies are part of that data set. We have controlled trials that are part of that data set also.

I think we are also interested in hearing your thoughts on going forward. I suspect that this going to be a mining exercise for everyone who does observational studies in the world probably. They are going to be looking to do another COX-2 or another NSAID observational study. We are going to see more and more studies published, and as I think someone said, it often becomes attractive to say, "Oh, look at that, you have got a very small p value, yeah, the relative risk is only 20 or 30 percent or 40 percent, the p value is very small, the study was very large, FDA, you should take regulatory action, you should take this drug off the market, you should restrict its use, whatever."

You are telling us you view them primarily as hypothesis-generating, and that they should lead to controlled clinical trials. The reality is even if we have the authority that we might like to mandate those trials, it is going to take years to get those controlled clinical trial data, and there is the pressure between people wanting you to act based on the observational data versus the scientific

desire to wait until you get better controlled clinical trials. I would be interested in having the committee say a little bit more of your thoughts about, you know, what do we do in the future in this class when we get the next observational study that is touted as wow, this really shows something, FDA, you should take action.

DR. NISSEN: Can I suggest some courses of action? One of them is that as people have pointed out, the strength of the association, I mean the hazard ratio is really important, and if somebody comes up with something which suggests 2 or 3, that is very different from a 1.3. The other obviously is to have a rigorous process for looking at the quality of it. One of the things I have learned from several of you is that there is observational studies and then there is observational studies, and some of them are done very well, and some of them are not done so well. The FDA has the expertise to evaluate that. Now, the problem is, of course, if it gets into the political arena, you get a lot of political pressure, but what we would want you to do in the public interest is look at the strength of association, look at the quality of the study, and make a decision on whether there is enough there to put a warning out. We have seen some strange things go on, like the warning around naproxen, that was clearly based upon pretty weak evidence. So, I think having a good standard is where you have to kind of hold your ground.

DR. WOOD: The other thing, in response to your question, is if we walk through the scenario here, the first signal was from a randomized clinical trial, and the question I guess then is what would we need to strengthen that observation

because it wasn't against placebo and all the problems there were with it. It would seem to me that what we don't need is a bunch of observational trials. That hardly is going to convince anyone, it seems to me. What we do need is an appropriately powered randomized trial that looks at the issue directly, and I am not so sure how long that would necessarily take. It only took 2 1/2 thousand people and approved to get the data. The question to which we don't know the answer, in fairness, is would it have taken less time if we had done a larger study, and I don't know the answer to that, no one knows the answer to that, but it is certainly potentially possible that we could have gotten the data quicker if we had done a larger study and the effect appeared faster. We don't know the answer to that, but that is one approach. I guess, responding to your question, it certainly seems to be in the public interest that you should have the power to ensure that that kind of a study gets done, and that is something certainly people should hear and hear loudly, I think.

DR. O'NEIL: Could I say something relative to a point that Janet Elashoff had brought up? The general process for the review of randomized controlled trials, such as the ones we have been reviewing, is we have the data in, there has been a strong movement for prospective specification of events, even blind adjudication, we look at the protocol very seriously. We actually have the data in hand. We actually can re-analyze, regroup, adjust, stratify, do many things. We are normally not in a position to do that on observational studies. We don't have the same level of process review for an observational study. In fact, it is not even clear what

the prespecified hypotheses were, even if you wanted to say the best that the observational study could do is to generate a hypothesis. However, there are many of them that have confirmed important things for us, the last of which was a protocol that we played a heavy role in, and that was propranolamine and its association with CVAs. That was a five- or six-year prospective case-controlled study that was done, that we reviewed the protocol. We had a heavy hand. In fact, David had a heavy hand in how that was designed, and that turned out to essentially support a regulatory conclusion. The point I am making here is that if we do open the door for observational studies, we have to have a different way of actually having access to the data, the quality of the data, and give it the same level of attention that we do in the review of randomized trials but for the fact that it's not randomized. Right now that is not in place, so we are talking about trials being balanced against observational studies where the standards for the trials are dramatically higher than the standards for the observational studies, not that they couldn't be better balanced, but I think that is an important issue. There has been a society, ISPE, the International Society for Pharmaco-Epidemiology has tried to put good principles in place to sort of say these are how you would do these studies, but we really don't have a process that would require that along the same ways that we would in these IND type studies or the larger randomized trials that we are seeing for the safety.

DR. WOOD: Tom.

DR. FLEMING: Just to reinforce some comments that Bob was just making, and Larry Friedman was making earlier, and

Steve Nissen, as well. Not all studies are the same, we know that is true of randomized trials in terms of their quality, it is certainly true in observational studies. Stuart Pocock more than 20 years ago put forward criteria for what you would want to do if you were doing an observational study that would be as reliable as possible. Essentially, it is just like a randomized trial, it is very complicated and takes considerable effort to ensure that you are putting in the structures. You can then have the sensitivity and specificity issues assessed or addressed by independent committees. You can do your best to try to define time zero cohorts. You still don't have randomization, though, and ultimately, the level of reliability is increased, but it still doesn't match the reliability of a randomized trial, as Charlie Hennekens was saying, until you are persuaded that the signal exceeds the potential magnitude of the bias, you can't be confident that the result is reliable. So often what we are looking at are effect sizes that aren't, in fact, larger than the magnitude of the bias, so that leads us down the pathway of needing randomized trials. John, getting back to your point, if you have a profoundly low p value, this may be obvious, but it doesn't mean we know the truth. There are two fundamental aspects around the truth. One is variability and one is bias, and I can have 100 trials put together and give me a highly precise estimate. I mentioned yesterday, you just end up with a precisely biased estimate, and that is my concern in the absence of randomized trials. I believe these are very useful clues, we need these results, but just because you have profoundly low p value doesn't mean we got at the truth.

DR. WOOD: That is what happened with the estrogen studies, of course.

DR. HOLMBOE: I just want to make one point because we keep hearing about the estrogen study. There is one very important fundamental difference here. Estrogen had been posited to have a positive effect on cardiovascular

mortality in observational trials, so it made a lot of sense to use randomized controlled trials to prove that hypothesis. The hypothesis here is that COX-2 inhibitors are harmful, therefore, you are doing a randomized controlled trial that in investigating harm, not benefit, and I think we have to keep that in mind.

### ***Question 3 Discussion***

DR. WOOD: Good point. Now, we are going to move on, Steve, to the next question. The next question is discuss the available data regarding the potential benefits of COX-2 selective non-steroidals versus non-selective non-steroidals, whatever they are, and how any such benefits should be weighed in assessing the potential benefits versus the potential risks of COX-2 selective agents from a regulatory perspective.

DR. JENKINS: Dr. Wood, could I make a comment about that as you get started about this particular discussion point? We put this in here for a reason, because clearly, we didn't want a three-day meeting to just focus entirely on risk, because the decisions you need to give us advice on have to be balancing risk and benefit. I think here we are particularly interested in hearing your views about benefit in a wide range of categories. You know, this class of drugs was developed for the GI effect, so we are interested in hearing your conclusions about the benefit of these drugs on the GI toxicity, but there is also other areas. Any input you have on their efficacy for pain relief for the treatment of inflammatory conditions will be useful. I am also interested in hearing your comments about the value of

choice. We heard that from some of the people in the public hearing today, that, you know, don't limit my choices, and we hear that a lot from physicians, we hear that a lot from patients, but we often are also hearing a competing view that if you have got one that looks like it is safer than the others, then, you don't need the others, but that is at odds with the idea that people like to have choice, because people don't respond the same to every drug, they may be allergic to one drug or whatever. So, in this context of benefit, I would like you to cover a lot of different areas, and not just to gastrointestinal benefit, but that is clearly one of the major focuses of benefit here.

DR. WOOD: Okay. Dr. Nissen.

DR. NISSEN: A couple of things. One is I haven't seen any compelling evidence that in terms of pain relief, that the drugs are actually more effective, and if such data is available, I would love to see it, but I don't find it there, so I think that is a little easier for me. I don't think we can minimize the importance of the GI aspect. There is actually two things, one of which was talked about interestingly by the public, but not necessarily by us or the companies, and that is, you know,

patient quality of life and patient preference. Any of us, I have certainly taken NSAIDs and gotten gastritis from them, and it is not fun, you know, having your stomachache, and people who have that every day, you know, there is a suffering related to that, that we heard from the public, and that has to be taken into account as we think about these drugs. In addition, I would be the first to say that a GI bleed is not a benign event. If these drugs were drugs that were better for treating acne, and they caused cardiovascular harm, that would be one thing, but the events, the GI events here are serious events. They are not as life-threatening as a stroke or a heart attack, but they can be, and they don't produce the permanent disability that a stroke or an MI does. You know, I take care of people with heart failure, and if you have had a big MI, and your pump doesn't work, your life is changed, the rest of your life is going to be different. Most people with a GI bleed recover, and so as I weigh these events, I don't discount GI benefits, but I have to give them less credence than the kind of hard, permanently disabling effects of MI and stroke, and I also think we absolutely have to factor in here the sort of suffering of patients who just don't tolerate the conventional NSAIDs, and I think that compassion has to come into our decisions.

DR. WOOD: Dr. Fleming.

DR. FLEMING: A great deal of the focus on the data we have had presented to us relates to the cardiovascular risks and relates to the confirmed complicated upper GI events, so if I start by focusing on that, it looks as though in a crude estimate that we might be having the rate of these events using the COX-2

inhibitors rather than the non-selective NSAIDs. It looks as though that might be, in 1,000 people, preventing 5, 6, 7, 8 events, something on that order. If we took a relative risk of 1.4 as the relative risk for the increase in cardiovascular events, that would be about 4 events.

So, coming back to what Steve is saying, when you look at it in that context, yes, these ulcerations are important events, but 7 per 1,000, how is that up against 4 events that are strokes, MIs, or cardiovascular deaths? I don't think it adds up.

If that is the whole picture, I would have a concern, but in a number of settings, it isn't the whole picture. We have heard about the oncology setting. We have talked about, truly, we haven't talked about efficacy. We have only had a number of comments stated that the pain relief seems to be about the same. Well, if it is the same, then that balance that I was saying concerns me as not being a favorable balance, but we heard a lot of people testifying, and I will be the first to say open sessions at these meetings are not random samples of the entire public, but we still heard a lot of comments that reflected the fact that there seems to be some differential protection or pain relief in certain patients. Can we quantitate that?

Can we, in fact, more scientifically, rigorously establish certain subpopulations where there really is a differential relief? Then, the benefit to risk shifts, or in the oncology setting, the benefit to risk shifts. The bottom line here, though, is to me the issue isn't so simple as choices. The issue is informed choices, and it takes the kind of scientific studies to reliably identify

what are the true benefits and risks, so that patients are in a position to make an informed choice, and part of the challenge to this, as one of the speakers at the public session pointed out, is it is not always the case that what might be learned by those people doing the studies is being effectively transmitted to the bedside or to the patients and their caregivers, and that is the other aspect, as well. So, it is critical to follow a strategy here that allows us to reliably address benefit to risk and allow patients to make an informed choice.

DR. WOOD: Dr. Hoffman.

DR. HOFFMAN: I think for the last two days we have been hearing appropriate angst about damning a class of agents for which there is a measure of efficacy, both in regards to pain and GI events because of newly-discovered adverse events, but I feel like we are walking on eggs in trying to get away from a consistent observation that is the dose-response effects, relative risks that we are looking at in terms of cardiovascular endpoints. We have heard this from experts at the FDA, independent investigators. We have even heard it from the thought leaders of industry, there seems to be a consensus to the effect that there is a class effect. I do take Steve Abramson's point that all of these drugs are not pure in their effects in terms of COX-1 or COX-2, but this is the data that we have, and it seems like there is a consensus about a class effect, and there also is a consensus in acknowledging that the patients that we enter into randomized controlled studies are probably the people least at risk that we may not see in our practices, who come in with 3 or 4 comorbidities that may have excluded them from being in

this trial and actually having seen even a clear signal.

The data, of course, that we would like to have is something that we don't have, and that is, the old standards of treatment for pain, whether it's the arthritis pain of OA/RA or postoperative pain, with NSAIDs plus PPIs over a long, extended period of time.

We would all like to know the data for that over 2 or 3 years compared to the COX-2s, which I don't think any of us are saying should, as a class, be taken off the market, but certainly should be used at the lowest safest dose. Now, at the lowest safest dose we don't even know their efficacy qualities. We don't know whether at the lowest safest dose we have the same benefits in terms of preventing peptic ulcer disease, treating pain effectively, decreasing inflammation effectively, and that it seems is the data that we need to have.

I am a little concerned, as a footnote to that, about the issue of choice. I think it is our obligation to provide patients choice within the realm of relatively safe medications, but most of us would not give as a choice a narcotic analgesic to a patient with, say, fibromyalgia. I don't think we should keep drugs on the market because of public pressure if we have a signal that we feel is a very strong one. We shouldn't give people a choice if we think that choice is uninformed and potentially does harm. Now, I am not saying that for the class of COX-2 inhibitors, I am just saying that we need more data to be able to provide for ourselves adequate information to make that choice and give our patients informed choice.

DR. WOOD: Dr. Cryer.

DR. CRYER: We were asked to kind more widely consider the potential benefits. As I see it clearly, one of the benefits is GI, and I will comment on it, but I do want to reiterate some of the comments that I personally don't see the benefit with respect to efficacy. I think the clinical trial experience to date has pretty consistently indicated that the efficacy is similar to the traditional NSAIDs. We did see some provocative data with etoricoxib today suggesting greater efficacy in one trial than naproxen, but that wasn't replicated. So, overall, I have to think that the efficacy is the same as we have with the traditional NSAIDs. I appreciated the testimonials of the patients about their individual efficacy responses, but my conclusion about that is those are anecdotes and it is consistent with the clinical experience that we have with efficacy of NSAIDs, which is that there is variable and idiosyncratic, unpredictable responses between patients, and it is very common that you will have one patient who responds to one NSAID and does not respond to another. I do think that we would still be giving these patients a wide range of choices given that there are 20 other NSAIDs available in the U.S. among which they can choose.

The benefits clearly I think are in the GI tract, but I will say that my conclusion is that the GI benefits are less than previously speculated. If you look at the three outcome trials which we have, that looked at GI benefits, we have VIGOR, CLASS, and TARGET. The results in the VIGOR are clear, but I think that was clearly also of a manifestation of the comparator, and one of the things that I

would like to be remembered is that the comparator NSAID matters. One sees a greater degree of GI benefit when one compares against naproxen than when one compares against diclofenac, so I do think there is value from the CLASS trial. I know that there was a GI benefit shown against ibuprofen. In the TARGET trial, those GI estimates are overestimated primarily because they enroll a low risk group of individuals and in a lower risk. We have consistently seen in trials that when you have low risk GI group, the relative risk is higher although the absolute risk in a low risk population is very low. So, the benefit is going to depend on the comparator. It is probably less than the 50 percent that you suggested it to be, because that 50 percent is based upon the VIGOR trial. It is probably closer to maybe a 30 percent benefit that I would estimate. It also narrows when you consider low dose aspirin.

In the face of low-dose aspirin, there is no apparent GI benefit. So, I think we also need to modify our estimates based upon the population that would be using or not using low-dose aspirin. So, my conclusion about the GI events is that, yes, there is a benefit, it is not as large as we thought, the appropriate target population is smaller with respect to the target group. It could be low risk people not taking low dose aspirin, but this event doesn't happen very commonly in low risk, and when you look at the high risk people in whom these drugs were originally targeted, several data sets suggest that the high risk people do not, in fact, have any appreciable benefit of GI risk reduction from a COX-2 specific inhibitor.

Final comments about other areas of benefit. Dyspepsia isn't one that is very convincing. When you look at the dyspepsia data from the clinical trial experience, it is only a few percentage points reduced. Dyspepsia, I consider mostly a nuisance symptom for which we have other very safe therapies to effectively deal with this. Finally, from the GI perspective, the polyp story could be another potential benefit, but with regard to the polyps, we have to remember in every trial we have seen, we are only modestly reducing the polyps and ultimately, we don't reduce cancer risk unless we eliminate adenomatous polyps, so it doesn't really change our algorithm in terms of how we would manage these patients, which would be colonoscopy and polypectomy.

DR. WOOD: Before you finish that, there are only two drugs on the U.S. market now, celecoxib and valdecoxib, so let's review the upper GI safety for them first. Is there a study that you are aware of with valdecoxib looking at complicated ulcers that showed in randomized fashion that there was a safety signal?

DR. CRYER: No. Wait, what do you mean by safety signal?

DR. WOOD: GI benefit. Is there a VIGOR trial for valdecoxib?

DR. CRYER: No.

DR. WOOD: So, confining our discussion to the two drugs that are on the U.S. market, there is no VIGOR equivalent, if you will, in valdecoxib, right?

DR. CRYER: Correct.

DR. WOOD: Now, for the other drug that is on the U.S. market, celecoxib, the published study didn't show the full data set. For the full data set for that, there wasn't benefit either.

DR. CRYER: Correct, but we did have the benefit of--

DR. WOOD: I understand, but there is always a benefit--I mean there is mortality problems halfway through, too, that disappear, that we ignore when we get to the end of the trial. So, for the two drugs that are on the U.S. market now, we have no clear randomized data that show GI benefit given the endpoints that were predefined and the end of the trial, not the trial that was published without the complete data set. The TARGET trial looks at a drug that is not on the U.S. market. So, our job is to evaluate the two drugs that are on the U.S. market, it seems to me.

DR. CRYER: So, I agree with your comments about the fully published results in JAMA for the class, however, we did have the benefit of reviewing the full CLASS results in the FDA hearing four years ago, and it is based upon that evaluation that I am deriving my conclusions of the full data set in which there did appear to be a demonstrable GI benefit when compared to ibuprofen in people who were not taking aspirin, certainly not when compared to diclofenac.

DR. WOOD: But the trial was not--that was a subsequent analysis taking out the aspirin. That wasn't the predefined endpoint.

DR. CRYER: Point well taken.

DR. WOOD: So, I mean just summarizing the point again, we have a benefit in a trial for a drug that is not on the U.S market, but we are not prepared to extend a class effect to cardiovascular risk necessarily, so I don't think we can just sort of step back and say that we are going to give a class benefit to GI benefit either extrapolating from studies of drugs that are not on the U.S. market.

DR. CRYER: Just because they are not on the U.S. market does not reduce the validity of the observation, for example, with lumiracoxib, and just because this was not absolutely predefined, and the benefit was recognized in, let's say, a post-hoc perspective, I still think there is recognized benefit in the data that we see in terms of assessing the GI benefits of celecoxib versus ibuprofen, and lumiracoxib versus its comparators.

DR. WOOD: But the non-aspirin group also had a cardiovascular risk, right?

DR. CRYER: Absolutely.

DR. WOOD: I mean as we are doing Tom's sort of analysis, when we take out that aspirin group and say, wow, there is a GI benefit there, when we take out that aspirin group we find there is a cardiovascular risk. So, you know, we can't have it both ways.

DR. CRYER: Well, I would say that the cardiovascular risks extend to both groups, aspirin and non-aspirin.

DR. WOOD: Right, but it was clear in-- okay, Dr. Fleming.

DR. FLEMING: Just to pursue a bit further, Alastair, what Byron is saying,

there are two aspects that I hear you saying that are really critical to the comments that I had made earlier. One is that I might be overestimating the actual GI benefit when I say you are having maybe it's a 30 percent.

DR. CRYER: It depends on the comparator.

DR. FLEMING: But the other, even more important thing to me that you are saying is that in spite of what might appear in the open session, which we know is anecdotal, the scientific data you are saying repeatedly are showing in the RA, OA, CABG settings where we have done studies, that there is not a difference in the pain relief and the efficacy. I would like to get more sense about that. If that is even close to true, then, there should be an incredibly low threshold for what you would accept in additional cardiovascular events, because the only thing you are getting relative to nonspecific NSAIDs then would be a very small GI. So, it seems like the efficacy here about the pain relief is a key issue.

DR. WOOD: I think the company wants to say something.

DR. KIM: Mr. Chairman, if I could, I will just make a comment, please. As I said yesterday, at the time that Merck withdrew Vioxx from the market, we based that decision on the available data that was available to us at that time, and we also stated that we thought that it would be possible to continue to market Vioxx with a labeling change that incorporated the results of the APPROVe trial. But we decided and we concluded that the most responsible course of action to take, given the

information that we had at that time, and the availability of alternative therapies, was to voluntarily withdraw the drug from the market. We have heard over the past two days new data and we have seen in the New England Journal new data on some of these alternative therapies.

Merck's interpretation, as you have heard, of these data are that we are dealing with a class effect, and the major question on the table right now is how large is that class. We are a data-driven company. If this committee and the FDA agree that what we are dealing with here is a class effect, then, I think it would be important for us to take the implications of that conclusion into consideration with regard to Vioxx, particularly given the unique benefits that Vioxx provides, one of which you are alluding to. So, I just wanted to make that point.

DR. WOOD: So, just to understand, what you are saying is that if we think the cardiovascular effect is a class effect, you would consider putting Vioxx back on the market.

DR. KIM: What I am saying is that at the time we withdrew the drug from the market, we did so because of the availability of alternative therapies and the science that was available at the time. That science has progressed. We are now engaged in a discussion around that science. There are unique benefits to Vioxx, one of which is it is the only COX-2 inhibitor with proven reductions in gastrointestinal events, another one of which it is the only coxib which is not contraindicated for patients with allergies to sulfonamides, and the third is that we have heard numerous reports, and you have heard a few today, from

patients, including patients with chronic debilitating pain that Vioxx was the only drug that relieved that pain.

DR. WOOD: Okay, good. Dr. Farrar.

DR. FARRAR: I wonder if I could just be very clear that so far I don't think we have talked about benefits. The point I want to make is that what we are talking about with the GI, quote "benefit" is, in fact, a reduction of risk. No one that I know of takes coxibs of any kind for an upset stomach. I think what we need to do is focus on the benefit to the patients, and we heard some of that in the public forum today, and I want to be as clear as possible about the issue of that benefit.

There are two ways of measuring benefit, and, in fact, in outcome trials, there really are only two summary statistics that are possible. One is a mean or a median or some central tendency with a spread, standard deviation. The second is a proportion, and it is a proportion of responders, it a proportion of people who die, which is the easiest, and in pain management, we get into all kinds of arguments about how much improvement you have to have to be a responder. If you look at the data, we are used in most of our clinical trials to looking at means and standard deviations, and if you look at means and standard deviations, it is very hard to find a difference between any of the NSAIDs and acetaminophen, any of them. If you ask patients about what works for them, in clinical practice, every patient will tell you that one works and that one doesn't. "I get sick with that one, I don't get sick with the other one." That is not something that we measure typically in our clinical trials. If you look at what level of drug is effective, with

almost any NSAID, it is never, it is never above 50 percent in terms of patients who actually go on using the drug in a chronic process.

What we are talking about is trying to identify less than 50 percent of a population who respond to a drug, and I can tell you from clinical practice, as any of you who have treated patients with rheumatoid arthritis know, people like specific drugs because they don't cause side effects and because they do have an effect. I think choice actually is a very important issue. Granted, we don't want to provide choice if there is an absolutely huge risk associated with that choice, but I think it is really important to understand that pain kills in the same way that the drug potentially can kill. I think it is very important to understand those two principles, the principles of the difference between a proportion and a mean value. Now, I am obviously talking to the converted here, but I think the issue really is looking at those issues.

We don't have any good trials, any that look at switching behavior within our patient populations, so there is no data that I know of that will help inform us about the need to and exactly how to go about this process, but I do know that in spite of all of our understanding of what goes on with the COX-1/COX-2 pathways and the inflammatory pathway, that when it gets down to using it in the patient, the issue is, is it absorbed, does it cause local effects, does it get to the active site, once it's at the active site, are there enough receptors for it to then cause the effect that we are looking for, a whole host of factors that we really can't measure and haven't measured yet in terms of metabolic process. My honest sense

from the data that we have heard here is that the drugs that we are considering today, the two, perhaps three, has to do with the relative benefit of those drugs.

What is very clear is that there are people, and a large portion of people, who have trouble with the current list of what we call non-selective COX inhibitors, and that there is a very important role for the more selective COX-2 group, however we want to define that. I think it is also, however, very important to understand that not everybody should be on a COX-2 predominant agent, and one of the problems that we are struggling with right now is the fact that because they were marketed as being safer, there was a very large push to switch people over who may not have needed to be switched.

So, I think that the issues that we need to consider are there is very good data that these drugs are effective at least in some segment of the patients in whom they are tried. There is I think reasonable data to suggest that the potential risks is not clearly very different between them, at least not the data that we have to date, and that from that perspective it is going to be important that we carefully think about how we then go about controlling those drugs. I would end with just saying that I agree absolutely it is about informed choice, and that I think that there needs to be a fairly large amount of information in the label and information conveyed to patients and physicians to help them make those choices.

DR. WOOD: Dr. Gibofsky.

DR. GIBOFSKY: I am particularly pleased about the nature of the

conversation because as a student of medical history, it reminds me that the first treatment we had for arthritis was, of course, willow bark, and we told our patients to ingest willow bark in order to get salicylates, which, of course, have an anti-inflammatory effect. So, if only our patients could take aspirin, perhaps we wouldn't need the whole class of non-selective and selective COX-2s, but, of course, they can't. There are problems just with aspirin in the treatment of arthritis at the doses they need it.

I am intrigued by the comments that, well, you know, an MI is an MI and you are dead, but a GI bleed, you get up, you get over it with no long lasting effect, and that may be true for the people who survive, but as Dr. Cryer showed us yesterday, and the best data set we have from Dr. Singh, 16 percent of patients who have a GI bleed die, so for them, it's a fatal event and one that they are not going to get up and continue on. I don't want to get into a discussion of the GI benefit and whether, in fact, it was achieved with one agent versus another.

But what is clear is something that hasn't been remarked yet, and that is for patients going to surgery, who are going to require anticoagulation following their surgery, and that is particularly in large part patients who have arthritis and are undergoing joint replacement surgery, the risk of a traditional nonsteroidal with an anticoagulant appears to be far worse in terms of bleeding later on than the risk of being on a COX-2 because of the lack of platelet inhibition. So, certainly there is a benefit for patients in that group who are going to go to surgery and require concurrent anticoagulant.

With regard to the issue of patient choice, there is several sets of data--and we heard one--showing that when you give a patient two different medications, in one study, the ACDA study, looked at acetaminophen versus diclofenac, another one, the PACES study, looked at celecoxib versus acetaminophen, and you asked patients without knowing which drug they were getting, in which arm, patients expressed a preference for either diclofenac or celecoxib over acetaminophen in the treatment of their arthritis.

The other issue with regard to choice is that we have also recognized, even in the pre-COX-2 days, that not infrequently, patients develop what is called a tolerance to the agent that they were on, that the latest data set we had suggested that inside of 18 months, patient who were taking medication for their arthritis chronically had to be rotated among agents three to four times in that period of time. So, the necessity for multiple agents in our armamentarium, the necessity for agents that allows for this individual idiosyncrasy that we have heard of is quite important. As was alluded to, there can be two patients in the waiting room on the same drug, one will swear by it, one will swear at it, and so it is for that reason that we need to have, not just one agent in a class, whatever we define that class to be, but sometimes several. Sometimes they are agents of allergy or idiosyncrasy which necessitate having more than one agent available. I think it is for all those reasons that we have to consider that in the benefit part as long as we are discussing benefit in the last part of the day.

DR. WOOD: I think we have to be really careful accepting this data, this 15-year-old data from Dr. Freis. I mean he has published, he published multiple updates on that, and people keep showing that same data, and that data isn't what is in his latest revision.

DR. GIBOFSKY: Accepted. Dr. Cryer?

DR. CRYER: I would like to comment on that, and I think your point is well taken. While I showed the 16,500 data yesterday, at the same time I said that that estimate, based upon more recent evaluations, is probably an overstatement of the actual mortality risk, GI risk attributable to NSAIDs. Dr. Singh has showed me more recent data which he has conducted in the U.S., which has shown that the risk has dramatically decreased in the U.S. That is probably related to several factors included in which is the eradication of HP, the introduction of PPIs into the U.S. marketplace, as well as the introduction of COX-2 specific inhibitors. The most recent estimates that I have seen would suggest that the mortality is about half of what Dr. Singh previously suggested it to be.

DR. GIBOFSKY: Accepted, but even the mortality rate of 8 percent in a population is unacceptable.

DR. CRYER: It is not 8 percent, it would be 8,000.

DR. WOOD: It is much lower than that, and if you look at the curve, the fall occurred long before COX-2s were on the market.

DR. CRYER: You are correct.

DR. WOOD: The data are out to 2000 on his paper, and that fall had occurred by 1998, so that is before any of these drugs were on the market. My point is that we keep throwing this 100,000 number around, including from the industry people, when the data is 15 years old, and the author has updated it multiple times, and that is not reasonable, guys. Dr. Singh.

DR. SINGH: As the author of the papers that you are discussing--

DR. WOOD: I am talking about Dr. Freis's paper, which was actually published. Yours is an abstract, I think.

DR. SINGH: Also, the 16,500 was from my paper that we estimated with the Aramis data set, and that, you are right, it is not 15 years old, but that is about '94, '95 data, and now that we have newer data sets, that was an estimate from the Aramis data. The latest work now is actually on real hospitalizations based on the nationwide inpatient sample, which is a much better estimate of what is really happening than an estimate from a small patient population. When we go back and look in '93, '94, of what the total number of deaths that the Federal Government said occurred in the United States, we were off by 32, that's it. It was like 16,486. That is how far we were off by, just to let you know in terms of an estimate. This is also true that now, today, the latest data set that we have available from 2002, that has dropped significantly, and the death rates are more like 8,000. But the other place where we underestimated was the hospitalizations.

We underestimated the hospitalizations, they are not 108,000, there are a lot more

than that. The mortality rates today have gone down tremendously, and the mortality rates today are probably more in the 5 to 6 percent range, and that is where Byron is correct, as well. Then, as far as the trend is concerned, the data that I showed you today is based on 483 million hospitalizations. We are not counting about 50 hospitalizations and then extrapolating it to the country. There are 483 million hospitalizations and 3.68 billion patient years. Yes, the trend line started going down way before the COX-2s were introduced, but then there are two sharp years of decline. The trend line actually, if you look at my slide, is very interestingly correlated with PPI use, and I showed data to Byron from the same data set, that it also explains it very nicely because the duodenal ulcer rates have gone steadily downward, which would be attributed primarily to PPI use and H. pylori eradication therapy.

The gastric ulcer rates and the gastric ulcer hemorrhage rate have not gone down in the same fashion. They went down when the '94-'95 H. pylori eradication campaign started. Then, they plateaued off pretty much, and PPIs haven't done very much to gastric ulcers until 1999, when the gastric ulcer rate dropped dramatically. In 1999, there is a 22 percent drop per 100,000 prescriptions sold in this country. I don't know what it is because of. Coincidentally, in 1999, January 1, celecoxib was introduced. I don't know what it is because of.

DR. WOOD: Let's move on. Dr. Dworkin.

DR. DWORKIN: Much of what I wanted to say has already been said, but

I just want to emphasize that while there are no differences on average in pain relief amongst these drugs, certainly none that are replicated, as Byron pointed out, that there is a great deal of variability in response, and I think there is every reason to believe that some patients respond better to one drug than another, so you have variability in the pain benefit, and you have to consider at the same time there is variability in the tolerability of the drug. So, there are two sources of variability in patient response, which at least to my way of thinking provides a really solid basis for their needing to be a choice amongst several drugs, because you have the variability in the pain benefit amongst patients and the variability in their tolerability.

DR. WOOD: Dr. Cush.

DR. CUSH: I prefer to say that these drugs are equally potent between the COX-2 specific and the non-selective drugs. I think there is a variability, but that speaks to the need for choice. Every rheumatologist at this table will tell you they cannot manage in any effective or compassionate way osteoarthritis or rheumatoid arthritis using just Tylenol and aspirin and ibuprofen. That would be a gigantic step backwards. So, they are equally potent. I think when it comes, however, to the risk, thankfully, this risk is incredibly low, but we would like to make it lower, and what we need to put forward is that we need a strategy for risk modification that is going to extend to all these drugs that we are examining here, much in the same that occurred with GGI, I think that we can start with some recommendations and then make it the responsibility of the manufacturers to come up with studies that will further define how we can best

reduce the risk in people who may need to receive these medicines.

DR. WOOD: Dr. Morris.

DR. MORRIS: Let me focus on the question that asks about the weighting, because what we have is--I guess everybody interprets this question differently, but what I interpret it as is how do you look at these non-comparable outcomes and how you trade off a TIA from a gastric ulcer or something. I think what we can do is we can describe the effect and we can describe the probability of the effect, but what we don't know is what is the right way to weigh those things, and I would make a plea that probably the right way is to try to involve in some way the views of patients in that decision-making. I don't mean that qualitatively, I mean that quantitatively, is in quality of life type data where people have looked at various outcomes, looked at it on a single scale, and apply some of those ways, so we understand how patients view it, and go beyond just medically what we think patients should evaluate it, but how they actually do evaluate it, and try to use some of the input of those data. That literature suggests that we get it wrong, that there is things worse than death, and we always think of death as the worst thing to happen in a medical outcome, but yet from a patient's perspective, being paralyzed by a stroke is perceived as worse, and we need to understand patients' evaluation of these outcomes, so we can make those weightings better for them.

DR. WOOD: Ms. Malone.

MS. MALONE: Obviously, this is complicated. I agree with most of what

the previous speakers have said especially Dr. Cush, Dr. Gibofsky. A big problem is like Dr. Gibofsky had said about having choice and trying different drugs, and having a period of time when they would work, and then they wouldn't be as effective and you would have to try something else. That is why the need for choice is there. I have spent the last 35 years probably on each of the NSAIDs that are still available, and went through that, and the frustration and the pain, and just--it's very difficult, so when people give this anecdotal information and say that they have found something that works for them, they are going to fight for that. We have to be able to prove to them that the risk far exceeds the benefit, and we have to be able to show that, and we can decry anecdotal evidence as not being sufficient enough, but, in reality, it all comes down to anecdotal evidence. It all comes down to the personalization of it, what happens to me when I take this drug, what happens to me when this drug is not available. But I think behind everything is the whole element of trust, and they place their trust in us, in FDA, and we can't give in to pressure, okay, but we can't give in to pressure either way. We have to keep an open mind about it and realize what they are going through and try to put yourselves in their shoes.

DR. WOOD: Dr. Platt.

DR. PLATT: In the spirit of supporting informed choice, it seems to me we could do a very much better job than we do by using the existing data that FDA already has to provide good information to patients about the risk stratum that they inhabit. Saying that there is an overall 1 1/2 or 2 percent difference in the risk of a GI complication or

myocardial infarction is not doing the best service to most people who take those drugs. I would imagine that those data can be used to support predictive modeling that would allow a fair amount of discrimination so that individuals could be told that people like them can expect a risk of 1 in 1,000 or 1 in 100 or 10 in 100, and that would make it a lot easier, I think, for individuals and their doctors to make thoughtful decisions about the tradeoffs of the benefits and the risks. It seems to me those data are there and it would be a straightforward thing to make them available. We do that with breast cancer all the time. The NIH did a tremendous service I think to the public by providing good predictive models that let women know what their risk of breast cancer is to help them decide whether to take preventive action. I think we could do it with these drugs.

DR. WOOD: Dr. Bathon.

DR. BATHON: It is interesting that you would say that because that is, in fact, what most of us rheumatologists have been doing for the past four months with every single clinic visit, is weighing the benefits and the risks based on the data that exist right now, and it is a difficult endeavor. I think that we are really hearing from our patients, and we heard this today, we are in a different era of patient-doctor relationships, and patients want to be a collaborator in these decisions, and they want to know the information. I think that the way I am thinking about this problem right now is that these drugs, whether they are selective or non-selective, are another risk factor in the GI complications and the cardiovascular complications that we have to weigh along with their blood pressures, their diabetes status, their

BMI, their family history, and everything else to come to a final decision about what we recommend with their input. Until we see an unequivocal cardiovascular risk that outweighs all those other factors, I think that is the appropriate approach with the patient is to put the drug in with all the other risk factors and try to come up with the best benefit-risk ratio that exists for that individual.

DR. WOOD: Dr. Hennekens.

DR. HENNEKENS: I find Question 3 extremely complicated in a number of dimensions. I am attracted to Tom's formulation of benefit to risk, but I think we also have to consider these arthritis patients with regard to the use of selective coxibs. As a group, they are at maybe a double the risk of heart disease of their non-arthritis counterparts. They are also suffering terribly with pain. From that perspective, the data we saw over the last two days on naproxen was somewhat reassuring to me, but for the patient who has gastroesophageal reflux disease or an allergy to aspirin or non-selective NSAIDs, I think there the benefit-to-risk obviously shift although even here, I think they have to have their cardiovascular risk factors managed aggressively, and I would add three more dimensions. One is I am not reassured at all by the data that are available on the short-acting non-selective NSAIDs with regard to risks and benefits, and I think we need a lot more data there. I am also not reassured by data we haven't reviewed that acetaminophen is either sufficiently efficacious or much safer, and then finally, the problems with high doses of aspirin are real. I do point out, though, the UK TIA trial of 2,400 people that

gave aspirin 1,200 mg in a placebo-controlled design for 5 years, the rate of GI side effects attributable to the aspirin was 14 percent, significant bleeding was 3.3 percent, but this flies in the face that 25 percent of the people on placebo had GI side effects and 1.6 percent of them had a significant GI bleed, so I think nothing is straightforward here.

DR. WOOD: Dr. Nissen.

DR. NISSEN: Just one brief comment, and that is, one of the things I am struggling with for all of you, and maybe some of those that either deal with these diseases can help me with this, is that the people at greatest risk for GI bleeding are the older and more frail individuals who are also at the greatest risk for cardiovascular disease, and so finding the sweet spot for the drugs becomes a little bit harder. There obviously are certain populations where it is obvious, but the big populations where there is risk, is it not true--I think I heard from Byron that older people are at greater risk for GI bleeding, and I can assure you they are at greater risk for coronary disease, so the question is how does it tilt in any given patient. It is not so easy to figure it out.

DR. PLATT: But you can quantitate it. I mean it seems to me you could tell the patients what individually, approximately what they could expect on both dimensions, and for a lot of patients, they would be high on both, but at least they could make an informed decision about that.

DR. CUSH: But it's the same situation as the GI problem. We know what the risk factors are, and age is a risk factor, and we counsel patients, and we probably

should tell the ones who might be willing to accept some small risk, because they don't seem like they are at risk just because of their age, but they don't have any other factors, and the same thing can happen here with regard to the cardiovascular risk if we have some appropriate guidelines.

DR. CRYER: Steve Nissen, I think you have got it exactly right and that there seems to be a great degree of overlap in those who are at GI risk tend to be, not uncommonly, the same patients who are cardiac risk. They are older, they may have a previous history of cardiovascular disease, and other risk factors which are common to both risk considerations, GI, and cardiovascular.

DR. WOOD: Ms. Malone, do you want to say something?

MS. MALONE: Yes, I do. Just what Byron has said, all of that brings in the importance of the doctor-patient relationship, and today, with the health care climate that we have, I have heard patients say how difficult it is to go in and get an amount of time when you can talk to your doctor, have a relationship with him, and especially, as people become older, and where I live in South Florida, there are many elderly people who do not have family around, so they are going to their doctor by themselves, and they are dependent on that doctor's viewpoint. They will say, "Well, what do you think?" I used to say if I were your child, and then it was if I were your wife, now it is getting to be if I were your grandmother, you know, with the age of everyone, and I hope I live to say if I were your granddaughter. But that is very true, and again it is not a simple situation, and whether we need some

sort of health educator to assist the doctor to be able to explain this to the patients, so that they are not taking valuable doctor-patient time, but something needs to be done.

DR. WOOD: Thanks. Dr. Ilowite.

DR. ILOWITE: I wanted to talk to a few pediatric issues about these agents, the granddaughter:

First of all, about choice, there are far fewer choices in pediatrics. There is only three NSAIDs approved, only two liquids and none any longer that are available as once-a-day dosing regimens.

The second issue is about tolerability. Certainly, children have fewer serious gastropathic events, but they do have a lot of symptoms, and it is often difficult to get children to take medications that give them even bellyaches.

Third, is the risk of cardiovascular disease, which is very low in pediatrics. A new clinical research network called CARRA, Childhood Arthritis and Rheumatology Research Alliance, organization polled its 130 members of whom 92 or 71 percent responded, and there were no events of myocardial infarction or stroke that couldn't otherwise be accounted for easily that were attributable to these agents.

Lastly, is the issue of exposure. It is likely that children with chronic rheumatic diseases are going to be on these agents longer even than adults, and the cumulative risk is of great concern. I think it would be very important to try to get some insight into the pathogenesis of this, not just the frequency, so that early markers could be explored in children

who are exposed before they exhibit the clinical endpoint.

MR. LEVIN: I haven't spoken for two days, so now I may go on. A couple of thoughts. One is I am all about informed choice, but the question is how informed is the choice, I think, as others have raised, and I want to point out that I think we have this sort of mythology of a changing environment which is patient-centered in which there is this sort of partnership. With all due respect to the clinicians around the table in the room, I don't think that characterizes most people's experience in the health care system today. I think it is totally unrealistic. We have 45- to 50,000 people who are uninsured, who have very haphazard access to care, certainly don't have an ongoing relationship probably with a practitioner who is going to sit down and run through the benefits and risks in the alternative therapies and help them make an informed decision.

We know from studies of how much time physicians have with patients and what they convey when they prescribe a drug, that is far from the role of the learned intermediary that is sort of I think mythic, and we need to get over. I agree with Lou that we need to ask patients what they want and what their experience is, but on the other hand, we have a regulatory context here. We have 1906, we have 1938, we have 1962. For better or worse, the Congress has decided that there is a role for government to play in protecting the public from harm.

So, I don't think we can just sort of slide this all off on patients and physicians supposed in this Nirvana good, up-to-

date information, making intelligent choices through this very difficult, complex issue. The Government does have a responsibility, and that is why we are here. We are being asked for, I think, advice on how government can best meet its responsibilities under statute to protect the public health and to do what it has to do. We all recognize that there are lots of things that need to be improved, I believe, in the way new drugs come to market, because I have sat through this before when we are chasing the train. The train is out of the station, folks, it is going down the track very fast, and we are trying to catch up to it and figure out what do we do. You know, it is heading for the crossing, there is a car on the track, how do we stop the train. It is too late. We are always going to hear from patients no matter what the drug, "This drug worked for me, it's wonderful, it changed my life." I believe them, I certainly empathize with them. There will always be that appeal.

So, I guess we have a complex task, the train has left the station, but we can't abrogate our responsibility, and we can't pretend the Government, through the FDA, doesn't have a statutory responsibility here to protect the public health. We can't just say put information out there, make it transparent, let this mythical doctor-patient relationship sort of bubble up and make things all right, because it's not going to happen that way.

DR. WOOD: Helpful comments from our consumer representative. Dr. Manzi.

DR. MANZI: First, I would like to congratulate the members of the panel who I thought have brought some very

relevant points to the table, and I agree with most of them, but it is interesting to me how many times I have heard the term "safe alternatives" used. I look at our first question about weighing the benefits of the COX-2s versus the non-selectives, and I think the assumption, as we are trying to deal with the coxibs, is that there is, quote, "safe alternatives" in the non-selective agents that we would feel comfortable having our patients turn to in the event that these other COX-2s were not available. My question would be, or I guess my challenge to my other panel members would be to provide data that has been obtained with the same rigor and had to undergo the same scrutiny as the drugs that we have just looked at to prove that the other non-selectives are safe alternatives. I don't think we have it. I think we have signals actually to the opposite potentially. So, I just think we have to keep that in mind as we are making decisions that patients are going to have to turn to something, and do you feel comfortable saying that the alternatives are safe?

DR. WOOD: Another way to think of the same thing, though, is that if we were sitting here thinking about approving these drugs right now, would we approve drugs with a clear cardiac risk in randomized clinical trials. I think that is an important question for the committee to address because if we don't address that, we will either not be able to address it for drugs coming up in the future and/or we are going to apply a different standard to drugs that are on the market, and I understand all these points, but I think it's--maybe I am wrong--I think it's highly improbable that the committee would have approved any of these drugs given the safety signal we have got right now. I think it is

highly improbable that the FDA--I am talking about from randomized clinical trials--I think it is highly improbable the FDA would have approved drugs if they had had all the randomized studies they have right now. That doesn't mean they wouldn't have approved them eventually perhaps, but they certainly wouldn't have approved them on that basis. Is that fair, Bob?

DR. TEMPLE: I think it varies depending on how you view various collections of data, but some of them I think probably would not have made it.

DR. WOOD: All right, some of them we would not, but that is a fair comment.

DR. MANZI: Could I just comment?

DR. WOOD: Sure.

DR. MANZI: I would argue that that would depend on the need for the drug, and it would also depend on the alternatives available, and so I think it is hard to look at it in isolation.

DR. WOOD: Fair point. Dr. D'Agostino.

DR. D'AGOSTINO: The Framingham study has generated many risk assessments. They are in the cholesterol guidelines. Cardiac risk assessment tools do exist. Would the physicians use them? I am not sure that cardiologists use them, nor other classes of physicians to automatically use them and sit with the patient and go through that, but they do exist, and if you could build a scenario for that, it would be possibly very useful. But one of the things I wanted to really mention isn't just the existence of these tools, but there seems to be something synergistic about taking

the drug and your cardiac risk, so it is not just a matter of telling you you are diabetic and how likely you are to have a heart attack. This drug seems to double that or triple that, and so forth, so you will be presenting very high risk to the subjects, and I am not so sure how easy that is to do, but it should be kept in mind that there is an elevated risk beyond the normal cardiac risk.

DR. WOOD: Unless someone else has a burning question, I am going to give Ms. Malone the last word.

MS. MALONE: I feel the need to speak up for rheumatologists. I had been on this panel I believe starting in 1995, and as a consumer rep. I filled someone's term, and then I had my own term. So, I was on it for five years, and then I came on as a patient rep intermittently. From my 35 years dealing with rheumatologists and being on the panel, I have to say that rheumatologists, on a whole, are a unique set of doctors. They are in there for the long haul and I have always felt that when I was on this committee, if I were not here, that the voice of the patient would still be heard. I find that I don't think there is one rheumatologist on here who would not spend time with their patient, who would not spend time educating them and listening to them albeit it it's not a half-hour, but I think they do have the ability to form a relationship with them, and I applaud them for that, and I disagree with Arthur on that point.

DR. WOOD: Stephanie, I will give you the last word and then we are stopping.

DR. CRAWFORD: Thank you so much, Mr. Chairman. I simply can't quite leave without at least attempting to address this stunning near cliffhanger that we

were given about 40 minutes ago. I am going to ask, if I may--and please forgive me if I get your name wrong, it's not listed on my papers--I think it was Dr. Kim from Merck. Thank you. Yesterday, I asked the question to Dr. Braunstein about what was or were the deciding factors in the extraordinary step that Merck made in deciding to voluntarily withdraw rofecoxib. I am not sure I heard a clear-cut answer, so I am going to ask you something very related to this last question we have been addressing from the opposite side. Tonight, what considerations would you weigh or would you ask this committee to consider when we deliberate tonight or tomorrow in determining the benefit of potential re-introduction of rofecoxib, or if you wish to say this class, where the benefits would far outweigh any issues of safety concerns?

DR. KIM: Thank you for that question, and I will say that it has certainly been a very educational and informative day, two days actually, listening to these discussions. I think the issues are complex, and I think that all of the complex issues are being brought up. As I said, Merck believes, based on the new data that has just become available, that what we are dealing with here in terms of cardiovascular risk is a class effect. The thing that we are struggling with, which you are all struggling with, is what does that mean in terms of the size of the class, and, in particular, is it limited to just inhibitors of COX-2 or does it include inhibitors of COX-2 that also now have an effect on COX-1. The only point that I was trying to make was that at the time that we decided to withdraw Vioxx from the market, we did so based on the information that was available to us at that time, knowing that

there were alternative therapies and that there were questions that were raised by the APPROVe trial.

Now, where the science has progressed to, where we see, we think, and we look forward to your decisions, but we think we are dealing with a class effect, then, I think we are no longer dealing with a situation where Vioxx is unique in its cardiovascular risk, but instead is a member of a class. Then, I think it is important for us--again, we are looking to you, this committee and the FDA, for your evaluation of whether or not you agree with our interpretation that this is a class-specific effect, but if that is the case, then, I think we need to take a look at the unique benefits that Vioxx provides, which I mentioned, and actually a fourth benefit which was already mentioned, that is, that Vioxx is the only COX-2 inhibitor which has been proven to reduce the events, serious GI events, as compared to naproxen. Vioxx is the only COX-2 inhibitor that was approved that is not contraindicated in patients with allergies to sulfonamides, and Vioxx was the only COX-2 inhibitor with approval for juvenile rheumatoid arthritis in addition to the fact that we have heard numerous reports from patients, some with very chronic debilitating pain, that Vioxx was the only drug that worked for them. With that, I will leave it to the committee. We really await your decision on this issue.

DR. WOOD: Okay. It's never the last word, is it.

DR. STRAND: May I finish the answer to a question that I was asked yesterday?

DR. WOOD: Who are you?

DR. STRAND: I am Dr. Strand and I responded to you yesterday about the use of COX-2s in patients, the benefit-risk profile. I simply want to say that with Dr. Hochberg we authored an editorial in 2002 after the introduction of the data from CLASS and VIGOR to point out that there is benefit with these COX-2s, which is at least numerically preserved from a GI point of view, both from TARGET and CLASS data, with a baby aspirin, and, in fact, most of the

cardiovascular risk may be abrogated by co-administration, and we certainly don't have to then worry about the potential interaction as has been demonstrated with ibuprofen. So, I think it is important in your deliberations to consider that point. Thank you very much.

DR. WOOD: Kimberly tells me the committee has to meet in the lobby in 15 minutes. I think that is pretty optimistic, but good luck.