

# Questions to Advisory Committee – Introduction, General Discussion & Committee Statements

## 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.....	4

### Highlights

#### DR. WOOD’S VIEW:

- **CV RISK: MULTIPLE TRIALS:** There are 4-5 randomized trials that show a cardiovascular hazard, with replication of results for two drugs – rofecoxib (VIGOR and APPROVe trials) and valdecoxib (two CABG trials). For celecoxib there is one study (APC trial).
- **LARGE SAFETY SIGNAL:** This is “a far larger randomized safety signal than we have seen for any of the drugs that have been withdrawn for safety reasons”.
- **RISK IS FOR COMMON EVENT:** The present situation is unusual in that it deals with an increased incidence of a common event rather than an unusual event.
- **OBSERVATIONAL STUDIES SUPPORT DRUG RANKING:** The observational studies, with “all the caveats that we heard” tend to show the same results as the randomized trials, and can help in ranking “drugs by toxicity, and toxicity by dose”, with rofecoxib being “the most toxic”.
- **NO CLEAR GI BENEFIT FOR CELECOXIB & VALDECOXIB:** No individual trials with celecoxib and valdecoxib established a better safety profile for GI clinical events.
- **NO CONTROLLED DATA SHOWING BETTER RESPONSE IN INDIVIDUAL PATIENTS:** Despite “moving” testimony at the open public hearing, no controlled trials have been done to evaluate better response in individual patients. We need to be able to identify patients who “uniquely benefit” from a coxib.
- **EASY-TO-UNDERSTAND RISK DESCRIPTION NEEDED:** We need an understandable way to describe the possible associated risk.
- **ADAPT TRIAL STOPPED FOR OPERATIONAL REASONS:** The ADAPT trial now seems to have been stopped because of “operational reasons” and not because of “safety signals”.
- **CLEAR RECOMMENDATIONS NEEDED:** It is “wonderful” to “pontificate” but FDA needs clear recommendations to act on.

#### **DR. FITZGERALD'S VIEW:**

- **FOCUS SHOULD BE PLACEBO-CONTROLLED TRIALS:** Dr. Fitzgerald suggested that the “focus of our deliberations” should be the placebo-controlled trials – because of better quality and more biological plausibility.
- **TARGET TRIAL & PLATELET HYPOTHESIS CONSISTENT:** The lumiracoxib TARGET study (which showed no significant increase in risk) is not “inconsistent with plausibility of the mechanism”, and that the trial had low-risk patients and was underpowered to detect an increased cardiovascular risk (even then, non-aspirin users had a CV hazard ratio of 1.47).
- **TARGET TRIAL BLOOD PRESSURES UNRELIABLE:** The blood pressure in TARGET was not a pre-specified endpoint and he doubts the reliability of the 1-2 mmHg lower values on lumiracoxib since “it would indeed be amazing if an even more selective drug was less effective on blood pressure”.
- **PLATELET HYPOTHESIS EXPLAINS BOTH ACUTE AND CHRONIC RISKS:** The hypothesized platelet effects are “entirely consistent with an acute and chronic time-dependent evolution of risk”.

#### **DR. FLEMING'S VIEW:**

- **SAFETY JUSTIFIES LESS STATISTICAL RIGOR:** Dr. Fleming said that it is important to take into account, as Dr. Packer had said, multiplicity in testing individual safety parameters over time, and multiplicity in the actual safety parameters. However, “when you are looking at safety” it is less acceptable to apply conservative statistical procedures such as “monitoring boundaries” because there is a multiplicity of safety issues, and because you have to take into account both the “severity of those safety issues” and “benefit to risk”. Ultimately, the statistical procedures can “provide some guidance” but “there has got to be informed judgment”.
- **DATA ACCESS FOR DATA MONITORING COMMITTEE:** Data monitoring committees are “critical” and should have “have sole access to emerging data on safety and efficacy during the course of the trial”.
- **VIGOR WAS HYPOTHESIS GENERATION: LATER TRIALS WERE CONFIRMATORY:** The first study to show the CV safety signal was VIGOR and it is appropriate to view this as hypothesis generation that requires “confirmatory” data. He thinks that confirmation has been obtained since there are “at least a dozen trials and at least half of those trials show an indication of excess risk and the majority of those are placebo-controlled trials”.
- **INFORMATION SUFFICIENT FOR MEASURED RESPONSE:**

The information “is clearly sufficient for a measured response”.

- **SAMPLE SIZE ISSUES:** He does agree that “we need greater insight”. Drilling “down to the numbers”, to rule out a doubling in risk takes 2,500/arm versus 10,000/arm to rule out a 50% increase in risk (his best point estimate of the coxib effect is a risk of 1.4-1.5). He thinks that 10,000/arm is “conceivably doable”

in view of the 23,000 people in the METAL trial.

- **LONG DURATION WILL INCREASE EVENT RATES:** Since it is important to evaluate long-term risks, the higher event rates resulting from longer duration of therapy would give more power with a given sample size.

#### **ARTHRITIS ADVISORY COMMITTEE VIEW:**

- **ENJOYABLE INTERACTIONS BETWEEN COMMITTEES:** Dr. Gibovsky said that he, and his “colleagues on the Arthritis Advisory Committee”, “very much enjoy the interactions with our colleagues in Drug Safety”.
- **SAFETY IN APPROVED INDICATIONS, WHEN USED AS LABELED:** While “safety for patients in the absolute is important”, marketing approval for the coxibs was given by FDA on the basis that “the potential benefits of each product outweigh the potential risks, when used for the approved indications, according to the directions included in the product labeling”.
- **WITHDRAWALS USUALLY FOR AE IN POPULATION WITH INDICATION:** We may be “at the dawn of a new paradigm” and “when we leave here tonight” we

should “provide some clarity” – but it is his understanding that where “drugs have been withdrawn .... it has usually been in the context of adverse events in the group for which the drug was approved and not based on adverse events in a prevention or proposed group”.

- **FOCUS ON SAFETY IN PATIENTS WITH ARTHRITIS & PAIN:** He thinks that “we need to look at our questions both in terms of absolute safety, which is critical, as well as relative safety as we define the populations which are going to get these drugs, namely the patients with arthritis and pain”.

#### **DR. WOOD’S “CORRECTION”:**

- **TROGLITAZONE IS EXAMPLE OF AE IN NON-APPROVED INDICATION:** Dr. Wood asked to

“provide some correction” to Dr. Gibovsky’s comment about the reason for prior drug withdrawals,

and gave the example of troglitazone which was withdrawn because of adverse effects in prevention of diabetes rather than the indication (treatment of diabetes).

- **CAUTION IF EXTRAPOLATE TO OTHER POPULATIONS:** Dr.

Gibovsky agreed (“I think that is absolutely correct and it is not a uniform finding.”). He said his “concern is the extrapolation from trials of prevention to trials of treatment, and I merely indicate that we cannot be universal about that.”

#### **DRUG SAFETY AND RISK MANAGEMENT ADVISORY COMMITTEE VIEW:**

- **REVIEW REASONS FOR PRIOR DRUG WITHDRAWALS:** Dr. Gross made a suggestion on behalf of the Drug Safety and Risk Management Advisory Committee that data on drugs previously withdrawn from the market be reviewed for “commonalities and differences that could guide policy decisions in the future”:
  - The nature of the adverse events.
  - When they were first recognized.
  - The pre-marketing signals.
- The extent to which the nature of the decisions was dependent on the availability of other drugs in the class.
- **REVIEW WOULD BE GOOD FOR COMMITTEES, PUBLIC AND PRESS:** This would facilitate risk:benefit decisions by Advisory Committees, would give the public a “better perspective”, and give the Press “a better sense of relativity of all these activities”.

#### **NO NDAC VIEW:**

- Dr. Wood said that they would be “glad to hear” that he was “not going to make a statement on behalf of the

NDAC Committee” (Nonprescription Drugs Advisory Committee).

## **Discussion Text**

DR. WOOD: Thanks very much.

I thought it would be helpful if I just made a few comments about what I think we have seen over the last three days and why this has difficult. I think what I have seen, at least, is we have

seen four, maybe five, randomized controlled trials that show a significant cardiovascular hazard which was replicated for two of the drugs, Vioxx showing VIGOR and APPROVe and Bextra the early CAB study and the later

CAB study, and, for Celebrex, the APC study.

It is important to recognize, this is a far larger randomized safety signal than we have seen for any of the drugs that have been withdrawn for safety reasons. In all of these studies, as Tom Fleming pointed out a number of times, the other adverse events seem certainly to trend at least the coxibs in many of them. So you might say, well, why are we discussing this and you might also say, why has it taken us three days. I think the reason for that is that this is probably one of the first times that we or the FDA have had to deal with a drug that caused a substantial increase in the frequency of a common problem, common disease like MI or heart disease or whatever, in contrast to an increase in the frequency of a rare disease like acute liver failure, even things like torsades de pointes in which there were other issues that made it difficult. So the difficulty of struggling with that, I think, is real and has been talked about by many people.

The other question that has come up and has been raised by many people is what do we see in the observational studies. Well, from a personal level, I guess, what I saw was, which is kind of backwards, I suppose, is in some of them, at least, it seemed to show the same as the randomized trials and that is somewhat reassuring, I suppose. With all the caveats that we heard, the observational studies, may allow us to rank drugs by toxicity, and toxicity by dose, with all the caveats that we just saw with, I guess, Vioxx currently being the most toxic.

In terms of G.I. safety, although it is frequently thrown up there, we saw no

data that showed better G.I. safety at the PUBs and a hard endpoint for Celebrex or Bextra except the discredited JAMA Celebrex paper that failed to disclose the full dataset and that was now the subject of critical and apologetic comments from the Editor of JAMA, herself.

We heard testimonials from patients which I thought were both moving and important although, in fairness, it is fair to say that no one has been able to demonstrate specifically better response amongst any of these drugs in individual patients in any randomized way and, as Bob said earlier, such studies--Bob Temple said earlier--such studies would be useful.

So that brings us to the \$64 million, probably, question, what should we do? Well, first, this is a much bigger--I mean, as was said earlier, however one passes these numbers, this is a much bigger safety problem than we have seen with the 16 drugs that the FDA has withdrawn. The only reason that we have not acted, I think, or the only reason we have agonized so much is that this is a relatively common problem and it is, therefore, much harder for us to be sure that we have seen a signal.

Clearly, though, the Committee needs to act in a way that limits this hazard to patients and the public has the right to expect us, I think, to do that and I think we need to focus on that as we go through this. Although, it is interesting to discuss these issues, we really need to make sure that, before we leave here, we have provided some sort of reassurance.

If there are patients who uniquely benefit from these drugs, then we need to consider any revised marketing strategy

which could range from withdrawal to great limitations on the use of the drugs. We need to identify patients who can uniquely benefit from these drugs and work out what they need to be told and what risk they would be willing to accept for that small number of unique patients who would benefit from the drugs.

We also, I think, learned a very important thing this morning which was that, in contrast to some of the information that had been put out in the press, the ADAPT study seems to have been stopped largely for operational reasons and many of the "safety signals" that we heard about in that were not backed by the usual approach that we would take. That, I think, is an important lesson that we did get today.

So I wanted to frame our discussion to these issues and also to make clear to everybody that, when we leave here tonight, we need to have made really clear recommendations to the FDA that will help them move forward. It is wonderful to sit and discuss the issues and pontificate here, but we really need to come down to some conclusions here that they will be able to take away and act on. Now, a number of people have indicated they wanted to say something. Garret FitzGerald wanted to say something in relation to some of the comments that came up in the last session. Garret?

DR. FITZGERALD: Thanks, Alastair. I thought it might be worthwhile to reemphasize one of the points that you have made, actually, and that is that the focus of our deliberations would most appropriately be on the randomized

controlled trials particularly the placebo-controlled trials for two reasons.

One, I believe that the quality of the evidence is much greater than in the observational studies and I think everybody has said that and, two, that the biological plausibility for the issues addressed in the placebo-controlled trials of the coxibs is much greater than the biological plausibility of risk relating to the traditional non-steroidals that were the subject of the observational studies.

As far as biological plausibility is concerned, there have been several comments yesterday and today that seem to cast out the symmetry of the evidence with the plausibility of the mechanism advanced. I am only going to make comments about two of those issues.

One, the most recent one, which was the TARGET study. In the TARGET study, we had a highly selective drug which did not reveal a cardiovascular risk significantly. However, as we heard yesterday, the TARGET study was set up in a way by choosing patients at low G.I. risk to amplify the detection of a G.I. benefit and, by choosing patients at low C.V. risk to minimize the likelihood of detecting a C.V. risk. Indeed, that study was grossly underpowered to detect a signal albeit that, in the non-aspirin users, the hazard ratio for cardiovascular events was 1.47.

As far as the blood-pressure aspects of TARGET are concerned, which are, indeed, asymmetric with the mechanism, I draw your attention to the fact that blood pressure was assessed retrospectively in TARGET and the reliability of a 1- to 2-millimeter change, on average, under those conditions, to

me, is extremely questionable especially as we assume that traditional nonsteroidal comparators in TARGET were raising blood pressure through inhibition of COX-2 that it would, indeed, be amazing, if an even more selective drug was less effective on blood pressure. It certainly doesn't relate to the duration of action of lumiracoxib which is given at roughly 30-fold greater than the concentration necessary to completely inhibit COX-2 so that, although it has a short half-life, its pharmacodynamic half-life is extended and we were shown that it is an impact on prostacyclin by a synthesis. It is sustained throughout the 24 hours and corresponds to the other drugs in the class, yesterday by Paola Patrigniani.

So I think I would not view the TARGET experience as inconsistent with the plausibility of the mechanism.

Finally, the other point I would make is that Bob alluded to the platelet activation issue as being the manifestation of the mechanism. As I described, this mechanism has acute and unfolding chronic manifestations and, indeed, the data that we have seen in the controlled trials are entirely consistent with an acute and chronic time-dependent evolution of risk. Thank you..

DR. WOOD: Thanks. Tom, I put you off from before the break, so feel free.

DR. FLEMING: It's fine. Basically, I wanted to quickly comment on the essence of what I see from the Packer, Temple and O'Neill presentations. Clearly, when judging strength of evidence, it is important to take into account multiplicity, as Milt Packer was indicating. When you are looking within

the context of a single trial, that multiplicity can arise as multiple testing over time as well as multiple endpoints. Clearly, as he notes, with safety issues, there is a wide array of different measures and we have to take that into account when considering strength of evidence; monitoring boundaries, give us a guideline.

Yet many of us have struggled with trying to formulate monitoring boundaries when you are looking at safety because of the multiplicity of safety issues and the fact that you have to take into account severity of those safety issues and you have to take into account benefit to risk. Ultimately, while those statistical procedures that Milt was talking about can provide some guidance, there has got to be informed judgment.

Data monitoring committees are critical and I think we see, from the ADAPT trial, just another example of why it is also critical for the data monitoring committee to have sole access to emerging data on safety and efficacy during the course of the trial. What does this tell us, though, about where we are today now that we are looking at a wide array of studies.

The VIGOR trial was the first study out. That study, as Milt would say, needs to be viewed in the context of confirmatory and exploratory. There is much less confidence that you have in the reliability of a result that was suggested by the data as opposed to a prespecified hypothesis. There is also regression to the mean. So, when you are seeing an estimate of the two-and-a-half-fold increase, there is a reason to expect that

that single trial might be overestimating that overall strength of evidence.

But we now have considerable insight beyond that first trial. We have got, by my count, at least a dozen trials and at least half of those trials show an indication of excess risk and the majority of those are placebo-controlled trials. So, in my own sense, the issues that Milt is raising are very relevant but we are now in a context of having an extensive amount of information.

In my own view, it is clearly sufficient for a measured response and yet, at the same time, I would agree with Bob Temple, that we need greater insight. What he has put forward is one strategy for that insight, to get at a better sense of the extent to which this excess is specific to indication, to the dose, to the duration, to whether or not there is ancillary care.

Just to kind of get it drilled down on the numbers here, if you were trying to rule out a doubling, it would take about 2,500 people per arm, or 88 events in a pair-wise comparison. It would be more, in this case, because my own sense is I think VIGOR is overestimating the true risk. I don't think it is a two-and-a-half-fold increase. My best sense is, in a general aggregate sense, it is more on the order of a 1.4 to 1.5 relative risk.

To rule out a 1.5 relative risk would take 10,000 people per arm or, in Bob's study, about 50,000 people, a big trial. But METAL has 23,000 people so this does seem conceivably doable. Bob O'Neill makes the key point that duration--that the events, the risks, can be different over time. So this trial, if it were to be done, should be done in a way to get at longer-term effects as well,

which does, also, allow us to somewhat reduce the size of the study.

So, bottom line, is we know a lot, enough to certainly take measured responses, but it is also going to be important for us to get additional insights that are necessary.

DR. WOOD: Dr. Gibofsky?

DR. GIBOFSKY: Mr. Chairman, we very much enjoy the interactions with our colleagues in Drug Safety, speaking for my colleagues on the Arthritis Advisory Committee. But I think I speak for most of them in suggesting that, while safety for patients in the absolute is important, the important language for us is the standard language of the introduction to the questions; namely, the notion that the original approvals and subsequent supplemental approvals were based on a determination by FDA that the potential benefits of each product outweigh the potential risks when used for the approved indications according to the directions included in the product labeling.

I think that is important because, depending upon whether that clause is inserted into Questions 1 through 3, quite possibly, there could be different answers for both the absolute and the relative answers depending upon whether or not we consider that clause.

My colleague and friend Dr. Abramson has suggested that we may be at the dawn of a new paradigm here. If so, I agree with our Chairman that, when we leave here tonight, we provide some clarity. But I would earnestly implore my colleagues to remember that the last temptation and the greatest treason is,

perhaps, to do the right thing for the wrong reason.

Where drugs have been withdrawn, whether it has been because of their numbers or because of the increased incidence of risk, it is my understanding that it has usually been in the context of adverse events in the group for which the drug was approved and not based on adverse events in a prevention or proposed group.

So I think these comments need to be considered somewhat carefully and that we need to look at our questions both in terms of absolute safety, which is critical, as well as relative safety as we define the populations which are going to get these drugs, namely the patients with arthritis and pain. Thank you.

DR. WOOD: Well, let me just provide some correction to that. I am not sure that last comment is correct, the one about drugs being withdrawn because of adverse events in the indication for which they were approved.

DR. GIBOFSKY: Not all of them; that's correct.

DR. WOOD: Hang on. Rezulin produced acute liver failure in two studies in which it was being used to prevent onset of diabetes.

DR. GIBOFSKY: I think that is absolutely correct and it is not a uniform finding.

DR. WOOD: Now, these were not--

DR. GIBOFSKY: My concern is the extrapolation from trials of prevention to

trials of treatment and I merely indicate that we cannot be universal about that.

DR. WOOD: All right. I think we are ready, probably, to start--sorry; go ahead.

DR. GROSS: I would like to make a comment for the Drug Safety and Risk Management Advisory Committee and it is a perspective for the future. The question is: Is there something we can do to avoid the confusion that comes up every time adverse events arise after marketing the new drug, particularly when the signal for the adverse event was not totally clear before the drug was approved?

I suggest we consider an approach that our committee had discussed in the past, and that approach is to review the drugs that have been pulled from the market and look for commonalities and differences that could guide policy decisions in the future, questions such as: What were the adverse events? When were they recognized? What were the signals before marketing? and What decisions were made when other drugs that were available in the same class, such as the statins, were done, and what were the decisions made when there were no other drugs in the class such as occurred with alosetron or Lotronex?

If this were done, lessons could be drawn. Advisory Committees would be better informed to make benefit/risk decisions and the public would be better informed because they would be able to acquire a better perspective and the Press, along with the public, would have a better sense of relativity of all of these activities.

DR. WOOD: Okay. You will be glad to hear I am not going to make a statement on behalf of the NDAC Committee. Let me read the first part:

Three COX-2 selective non-steroidals are currently available for marketing in the United States, Celebrex, Vioxx and Bextra. The original approvals and subsequent supplemental approvals were based on a determination by the FDA that the potential benefits of each product outweighed the potential risks when used for the approved indications according to the directions included in the product labeling.

Since approval, additional data regarding the safety and effectiveness of these products has accumulated, in particular, new information regarding the potential cardiovascular risks of these products.

FDA must consider the impact of these new data on the benefit-versus-risk profile of each product in making decisions about appropriate regulatory actions.

Although--and this is important--although Merck voluntarily withdrew Vioxx from marketing worldwide on September 30, 2004, questions relating to Vioxx are included below since it will be necessary for FDA to determine the appropriate regulatory action regarding the approval status of this product.

Based on the data presented in the background package during the committee meeting, please address the following questions.