UNITED STATES DISTRICT COURT NORTHERN DISTRICT OF CALIFORNIA
IN RE: BEXTRA AND CELEBREX MARKETING SALES PRACTICES AND PRODUCT LIABILITY LITIGATION
This Order Relates to:
ALL CASES CASE NO. M:05-CV-01699-CRB
MDL No. 1699

MEMORANDUM AND ORDER RE: MOTIONS TO EXCLUDE EXPERT TESTIMONY

In this Multi-District Litigation ("MDL") proceeding, over 3000 plaintiffs allege that they or their loved ones suffered a heart attack, stroke or other adverse cardiovascular event as a result of taking Celebrex, a pain medication manufactured by defendant Pfizer, Inc. ("Pfizer"). Pfizer has moved to exclude any expert testimony to the effect that Celebrex is capable of causing a heart attack or stroke when ingested at 200 milligrams a day or 400 milligrams a day. Plaintiffs have also moved to exclude certain expert testimony offered by Pfizer. The Court held three days of hearings which included direct and cross examination of certain experts. After carefully considering the parties' memoranda and evidence, and the testimony offered at the hearing, the Court concludes that plaintiffs have not presented scientifically reliable evidence that Celebrex causes heart attacks or strokes when ingested at the 200 milligram a day dose. In all other respects the parties' motions are denied.

BACKGROUND

Non-steroidal anti-inflammatory drugs ("NSAIDs") have been widely used for pain relief for several years. NSAIDs, however, have certain side effects, including gastrointestinal toxicity which results in thousands of deaths every year. The pharmaceutical company Merck & Co., Inc. ("Merck") developed Vioxx, and Pfizer (or, more precisely, its predecessors) developed Celebrex and Bextra, NSAIDs known as COX-2 inhibitors, with the expectation that they would have fewer gastrointestinal side effects than traditional NSAIDs. The Food and Drug Administration ("FDA") approved Celebrex for adult arthritis in 1998, Vioxx in 1999, and Bextra in late 2001. The recommended dose of Celebrex was and is 200 milligrams a day ("mg/d") for arthritis and 400 mg/d for rheumatoid arthritis.

In 2000 the results of a long-term randomized study of Celebrex known as CLASS ("Celecoxib Long-Term Arthritis Safety Study") were published. The study was designed to evaluate the gastrointestinal side effects of taking Celebrex at 800 mg/d. Based on investigator reported cardiovascular events, the study showed no increased risk of heart attack or stroke by taking Celebrex over diclofenac or ibuprofen. Around the same time, a similar study of Vioxx, known as VIGOR, showed a four-fold increase in cardiovascular ("cv") risk for patients taking Vioxx versus Aleve (naproxen). The FDA subsequently revised the labels of Celebrex and Vioxx to reflect the cv risk results of these studies.

Another Vioxx randomized clinical study, known as APPROVe, was published in 2004. This study demonstrated a two-fold increased risk of cv adverse events for patients taking Vioxx versus a placebo. This study contributed to Merck's voluntary removal of Vioxx from the market on September 30, 2004.

The preliminary results of APC, a randomized, placebo-controlled study of Celebrex at 200 mg twice daily (400 mg/d) and 400 mg twice daily (800 mg/d) to evaluate whether Celebrex prevents the development of colon polyps, became available in late 2004. APC showed dose-related increased cv risk for patients taking Celebrex compared to placebo: more than doubling the risk for 200 mg twice daily and tripling the risk for 400 mg twice daily. The APC steering committee discontinued the study in December 2004 because of these preliminary results.

In February 2005 the FDA convened an Advisory Committee to review the data on cv risk and NSAIDs, including COX-2 inhibitors. The Committee concluded that all COX-2 inhibitors increase cv risk versus placebo, but it did not make any findings as to what dose is required to increase the risk. It also concluded that the data was insufficient to determine if traditional NSAIDs also increase cv risk. With respect to Celebrex, the FDA found that APC is the "strongest data in support of an increased risk of serious adverse CV events." FDA Decision Memorandum, April 6, 2005, at 4, Declaration of Loren Brown ("Brown Decl.") Exh. 16. The FDA also noted that APC's results had not been replicated by preliminary data from two other randomized controlled clinical studies: (1) PreSAP, a colon polyp prevention trial of Celebrex at 400 mg/d; and (2) ADAPT, an Alzheimer's trial of Celebrex at 200 twice daily (400 mg/d). Both studies showed no increased cv risk for Celebrex versus placebo.

The FDA subsequently asked Pfizer to remove Bextra from the market, which Pfizer did in April 2005. The FDA also determined that the benefits of Celebrex outweigh its risks and therefore it allowed Celebrex to remain on the market. Celebrex is the only COX-2 inhibitor currently on the market.

The FDA also directed all NSAIDs, including Celebrex, to include a black box warning on their labels. The black box warns of cv risk as follows:
Cadiovascular Risk
- CELEBREX may cause an increased risk of serious cardiovascular thrombotic events, myocardial infarction, and stroke, which can be fatal. All NSAIDs may have a similar risk. This risk may increase with duration of use. Patients with cardiovascular disease or risk factors for cardiovascular disease may be at greater risk . . . .

Celebrex 2007 Label, Brown Decl. Exh. 3.
As a result of these developments, thousands of patients and patient representatives filed lawsuits against Merck and Pfizer alleging that the patient had suffered a serious cardiovascular injury, such as a heart attack or stroke, due to their ingestion of Vioxx, and/or Celebrex and/or Bextra. All of the federal court claims against Merck were consolidated in a MDL action in New Orleans. All of the federal court claims against Pfizer were consolidated into this MDL proceeding.

THE DAUBERT MOTIONS

Pursuant to Federal Rule of Evidence 702, Pfizer moves to exclude plaintiffs' experts from offering the following six opinions:

1. That 200 mg/d of Celebrex causes heart attacks and strokes;
2. That 400 mg/d of Celebrex causes heart attacks and strokes;
3. That Celebrex causes heart attacks or strokes more than three days after a patient stops taking it;
4. That Celebrex causes strokes; and;
5. That Celebrex causes heart attacks or stokes at durations of less than 33 months of continuous daily use.

Pfizer also asks the Court to exclude any expert opinion that Celebrex caused any individual plaintiff's heart or stroke absent epidemiology evidence that demonstrates a relative risk greater than 2.0, that is, that Celebrex doubles the risk. Plaintiffs have moved to exclude certain expert testimony offered by Pfizer; specifically, they seek to exclude admission of the meta-analyses performed by plaintiffs' experts.

In connection with these motions, the parties submitted direct written testimony of their respective experts as well as legal memoranda. The Court then held three days of hearings, which were conducted jointly with the New York Justice presiding over the New York State Celebrex and Bextra cases. Plaintiffs' experts Dr. Neil Doherty, Dr. Joel Bennett, Dr. Nicholas Jewell and Dr. Maryilyn Rymer testified on direct and cross-examination, along with defendant's expert Dr. Milton Packer. The parties also submitted post-hearing memoranda. The motions are now ripe for decision.

LEGAL STANDARD

A. Admissibility of Expert Testimony

When evaluating the admissibility of expert testimony, the trial judge "must engage in a difficult, two-part analysis." Daubert v. Merrill Dow Pharmaceuticals, Inc., 43 F.3d 1311, 1315 (9th Cir. 1995) (Daubert II). First, the court must "determine nothing less than whether the experts' testimony reflects 'scientific knowledge,' whether their findings are 'derived by the scientific method,' and whether their work product amounts to 'good science.'" Id. (quoting Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579, 589-90, 593 (1993)); see also In Re Silicone Gel Breast Impl. Prod. Liab. Lit., 318 F.Supp.2d 879, 890 (C.D. Cal. 2004) ("'[T]he trial judge in all cases of proffered expert testimony must find that it is properly grounded, well-reasoned, and not speculative before it can be admitted.'") (quoting Fed. R. Evid. 702 Advisory Committee's Notes). The trial judge's obligation "is to make certain that an expert . . . employs in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field." Kumho Tire Co. v. Carmichael, 526 U.S. 137, 152 (1999).

Many factors may be relevant to the reliability inquiry, including: (1) whether the proffered theory or technique has been tested, (2) whether the theory or technique has been subjected to peer review and publication, (3) the known or potential rate of error of

the technique or theory when applied, and (4) the "general acceptance" of the theory or technique in the scientific community. Daubert, 509 U.S. at 593-94.

[C]ourts have also found the following factors relevant in assessing the reliability of expert testimony: (1) whether the expert is proposing to testify about matters growing directly out of independent research he or she has conducted or whether the opinion was developed expressly for purposes of testifying; (2) whether the expert has unjustifiably extrapolated from an accepted premise to an unfounded conclusion; (3) whether the expert has adequately accounted for obvious alternative explanations; (4) whether the expert is being as careful as he would be in his regular professional work; and (5) whether the field of expertise claimed by the expert is known to reach reliable results for the type of opinion offered.

In Re Silicone Gel Breast Impl. Prod. Liab. Lit., 318 F.Supp.2d at 890 (citing Fed.R.Evid. 702 Advisory Committee's Notes).

In addition to determining reliability, the court "must ensure that the proposed expert testimony is 'relevant to the task at hand,' i.e., that it logically advances a material aspect of the proposing party's case." Daubert II, 43 F.3d at 1315 (quoting Daubert, 509 U.S. at 597). This is known as the "fit" requirement. Id. Here, the pertinent fit inquiry is "causation." The parties' motions address expert testimony on the causation inquiry.

B. Causation

Causation in toxic tort or pharmaceutical personal injury cases "is typically discussed in terms of generic and specific causation." In Re Hanford Nuclear Reservation Lit., 292 F.3d 1124, 1133 (9th Cir. 2002). General or generic causation means "whether the substance at issue had the capacity to cause the harm alleged." Id. In Hanford, for example, the Ninth Circuit explained that the general causation inquiry was "whether exposure to a substance for which a defendant is responsible, such as radiation at the level of exposure alleged by plaintiffs, is capable of causing a particular injury or condition in the general population." Id.

To ultimately prevail in such a lawsuit, however, a plaintiff must show both general and "individual" or "specific" causation. Id. Specific causation refers to whether a particular individual suffers from a particular ailment as a result of exposure to a substance. Id. That is, that the challenged conduct, here, the taking of Celebrex at a certain dose for a particular amount of time, was "the cause-in-fact" of the particular plaintiff's injury. Id.

The parties' motions involve the use of epidemiology to prove causation. "The field of epidemiology addresses the incidence, distribution and etiology (causation) of disease in human populations by comparing individuals exposed to a particular agent to unexposed individuals to determine whether exposure increases the risk of disease." In Re Silicone Gel Breast Implant Prod. Liab. Lit., 318 F.Supp.2d at 892. Scientists use "relative risk" to identify an association between, for example, the ingestion of a drug and a disease.

For example, if a study found that 10 out of 1000 women with breast implants were diagnosed with breast cancer and 5 out of 1000 women without implants (the "control" group) were diagnosed with breast cancer, the relative risk of implants is 2.0, or twice as great as the risk of breast cancer without implants. This is so, because the proportion of women in the implant group with breast cancer is 0.1 (10/1000) and the proportion of women in the non-implant group with breast cancer is 0.05 (5/1000). And 0.1 divided by 0.05 is 2.0.
Id. A relative risk of 1.0 suggests that there is no association between the product and the disease, that is, the same numbers of people using the product are diagnosed with the disease as those not using the product. Similarly, a relative risk of less than 1.0 suggests that the product is actually "protective" of the disease: fewer people using the product contract the disease than those not taking the product. Id. at n.5.

In general, epidemiology studies are probative of general causation: a relative risk greater than 1.0 means the product has the capacity to cause the disease. "Where the study properly accounts for potential confounding factors and concludes that exposure to the agent is what increases the probability of contracting the disease, the study has demonstrated general causation–that exposure to the agent is capable of causing [the illness at issue] in the general population." Id. at 893 (internal quotation marks and citation omitted).

Such studies can also be probative of specific causation, but only if the relative risk is greater than 2.0, that is, the product more than doubles the risk of getting the disease.

When the relative risk is 2.0, the alleged cause is responsible for an equal number of cases of the disease as all other background causes present in the control group. Thus, a relative risk of 2.0 implies a 50% probability that the agent at issue was responsible for a particular individual's disease. This means that a relative risk that is greater than 2.0 permits the conclusion that the agent was more likely than not responsible for a particular individual's disease.
Id. at 893. The issue on these motions, however, is not specific causation; there is no particular plaintiff before the Court. Rather, the primary issue is whether the Court should permit plaintiffs' experts to testify that Celebrex is capable of causing heart attacks or strokes at certain doses.

EPIDEMIOLOGY STUDIES AND TERMS

Before discussing the parties' motions, it is important to identify the different epidemiology studies relied upon by the experts. There are generally three types of clinical epidemiology studies at issue on the parties' motions: (1) randomized controlled clinical trials, (2) observational studies, and (3) meta-analyses.

The "gold standard" for determining whether a drug is related to the risk of developing an adverse health outcome is a "randomized clinical trial" in which the subjects are randomly assigned to one of two groups: one group exposed to the drug of interest and the other not exposed. After a period of time the study participants in both groups are

evaluated for an adverse health outcome. Federal Judicial Center, Reference Manual on Scientific Evidence 338 (2d ed. 2000). "Randomization minimizes the likelihood that there are differences in relevant characteristics between those exposed to the agent and those not exposed," such as smoking, obesity, aspirin use and so on that could account for any difference in outcomes between the two groups. Id.

An "observational study" evaluates causation by comparing the risk of disease between patients exposed to a given substance and patients who were not exposed. The study may be prospective, identifying patients and then following them for a period of time, or restrospective, identifying patients and then performing a medical chart review to determine what happened during the period they did or did not take the drug. The downside to observational studies is that because the investigators do not control who participates in the study, it is more difficult to control for confounding factors such as smoking, obesity and the like. The investigator attempts to address the possible role of confounding factors "by considering them in the design of the study and in the analysis and interpretation of the study results." Id. at 339.

There are two types of observational studies: a cohort study and a case control study. A cohort study identifies patients who are taking the drug (exposed) and follows them for a certain amount of time to determine if they have the alleged bad outcome, here, such outcome is heart attack or stroke. The cohort study also identifies people not taking the drug and follows them (unexposed). The study then compares the rate of the alleged bad outcomes in group one with the rate in group two to compute the "relative risk." Id. at 339-40.

A case control study identifies persons who had a bad outcome (the cases), for example, patients in the United Kingdom database that had a heart attack within the last three years, and reviews their medical records to determine how many of those persons were taking the studied drug around the time of their heart attack. The study then identifies an equal number of people who did not have a heart attack (the controls) and determines how many of them were taking the drug. Id. From those figures an "odds ratio" is computed. For example, if the percentage of people taking Celebrex in both groups is the same, the odds ratio is 1.0; that is, taking Celebrex did not increase the risk of heart attack.

Sometimes randomized controlled studies and observational studies of the same drug will have conflicting results; some will show a statistically significant association while others will not. A meta-analysis pools the results of various studies to arrive at a single figure to represent the totality of the studies reviewed. "In a meta-analysis, studies are given different weights in proportion to the sizes of their study populations and other characteristics." Id. at 380. Meta-analysis has the advantage of pooling more data so that the results are less likely to be misleading solely due to chance. On the other hand, one problem with meta-analysis, particularly in meta-analysis of observational studies, is that the pooled studies often use disparate methodologies.

When reviewing the results of a study, whether it is a randomized clinical trial, observational trial, or a meta-analysis of such trials, it is important to consider the confidence interval. The confidence interval is, in simple terms, the "margin of error." So, for example, if a given study showed a relative risk of 1.40 (a 40 percent increased risk of adverse events), but the 95 percent confidence interval is .8 to 1.9, we would say that we are 95 percent confident that the true value, that is, the actual relative risk, is between .8 and 1.9. Because the confidence interval includes results which do not show any increased risk, and indeed, show a decreased risk, that is, it includes values less than 1.0, we would say the study does not demonstrate a "statistically significant" increased risk of an adverse outcome. Confidence intervals are calculated, in part, based on the number of people and events included in the study. "The larger the sample size in a study (all other things being equal), the narrower the confidence boundaries will be (indicating greater statistical stability), thereby reflecting the decreased likelihood that the association found in the study would occur if the true association is 1.0 [no increased or decreased risk]." Id. at 361.

With these terms in mind, the Court now turns to the parties' motions.

DISCUSSION

I. Pfizer's Motion

A threshold question raised by Pfizer's motion is whether a particular dose of Celebrex is relevant to the general causation inquiry. Pfizer seeks to exclude any opinion that Celebrex is capable of causing heart attacks and strokes at 200 mg/d as well as any opinion that Celebrex is capable of causing heart attacks and strokes at 400 mg/d. It does not move to exclude expert testimony that Celebrex is capable of causing heart attacks and strokes when a patient ingests 800 mg/d, at least when taken over many months. Thus, Pfizer's motion assumes that Celebrex at different doses can have different cardiovascular effects.

The Court finds that dose matters. All of plaintiffs' experts, with perhaps a single exception, agree that there is a dose effect with Celebrex; that is, that it is more toxic, and is therefore more likely to cause an adverse side effect, when taken at greater doses. See Reference Manual on Scientific Evidence at 403 ("There are three central tenets of toxicology. First, 'the dose makes the poison'; this implies that all chemical agents are intrinsically hazardous--whether they cause harm is only a question of dose. Even water, if consumed in large quantities, can be toxic."); see also Mitchell v. Gencorp, 165 F.3d 778, 781 (10th Cir. 1999) (noting that to prevail in a toxic tort case a "a plaintiff must demonstrate 'the levels of exposure that are hazardous to human beings generally as well as the plaintiff's actual level of exposure to the defendant's toxic substance before he or she may recover") (internal quotation marks and citation omitted); Allen v. Penn. Eng'g Corp., 102 F.3d 194, 199 (5th Cir. 1996) (explaining that in toxic tort cases, "[s]cientific knowledge of the harmful level of exposure to a chemical plus knowledge that plaintiff was exposed to such quantities are minimal facts necessary to sustain the plaintiff's

burden"); see also Hanford Nuclear Reservation Lit., 292 F.3d at 1133 (explaining that the general causation inquiry is whether exposure to the challenged substance "at the level of exposure alleged by the plaintiffs" is capable of causing the alleged injuries") (emphasis added). As plaintiffs' cardiology expert, Dr. Neil Doherty, testified: it is a "fundamental principal of medicine" and "medical causality" that the risk of adverse cardiovascular events with Celebrex is dose-related. Transcript of October 10, 2007 Hearing ("Oct. 10 TR") at 328. Thus, the Court must analyze plaintiffs' experts' opinions as to causation at 200 mg/d separate from their opinions as to 400 mg/d.

A. 200 mg/d

Celebrex at 200 mg/d and the risk of adverse cv events has not been studied in published, large, long-term randomized controlled trials. Nonetheless, included in the record are approximately 30 unpublished randomized controlled trials, albeit of short duration and small size. These studies do not demonstrate any association between Celebrex and adverse cv outcomes. A meta-analysis of all available published and unpublished randomized clinical trials of all COX-2 inhibitors as well as traditional NSAIDs found that while COX-2 inhibitors as a whole are associated with a moderate increase in the risk of adverse cv events, no such association is found with the available data for Celebrex at 200 mg/d or less[fn1] [fn1: 1 Patricia Kearney, et al., Do selective cycol-oxygenase-2 inhibitors and traditional nonsteroidal ant-inflammatory drugs increase the risk of atherothrombosis? Meta-analysis ofrandomized trials, British Medical Journal 2006, June 3; 332(7553): 1302-8.]

The record also includes observational studies with Celebrex data, mostly at 200 mg/d. These observational studies together include more than 8,000 adverse cv events, and all of the studies with the most events demonstrate no statistically significant association between Celebrex at 200 mg/d and adverse cv events. A meta-analysis performed by an independent researcher unaffiliated with Pfizer ("McGettigan") concluded that while Vioxx does increase the risk of adverse cv events, "[i]n doses of around 200 mg/d, [Celebrex] was not associated with an increased risk . . . ."[fn2] [fn2=Patricia McGettigan, et al., Cardiovascular Risk and Inhibition of Cyclooxygenase: A Systematic Review of the Observational Studies of Selective and Nonselective inhibitors of Cyclooxygenase 2, JAMA 2006 Oct 4; 296(13): 633-44.] Another meta-analysis of eight observational studies showed no increased risk from Celebrex 200 mg/d compared to patients taking no medication.[fn3] [fn3=S. Hernandez-Diaz et al., Non-steroidal anti-inflammatory drugs and the risk of acute myocardial infarction, Basic Clin. Pharmacol. Toxicol. 2006 Mar; 98(3):266-274, at 270, 273.]

In sum, there are no randomized controlled trials or meta-analyses of such trials or meta-analyses of observational studies that find an association between Celebrex 200 mg/d and a risk of heart attack or stroke. And most observational studies, indeed, the observational studies that include 97 percent of the reported adverse cv events, also find no statistically significant association. It is thus unsurprising that most of plaintiffs' experts agree that the available evidence at 200 mg/d is inadequate to prove causation. See Deposition Testimony of Dr. Joel Bennett at p. 537, Brown Reply Decl. Exh. 108 ("I think that if you

look at all the evidence, I think at 200 milligrams it's hard to make a case that Celebrex has toxicity. It doesn't mean that, again, that in individual cases it couldn't, it could be lost in the big scheme of things, but, in fact, the data don't suggest that in a large population it increases the risk."); Deposition Testimony of Dr. Lemue Moye at p. 268, Brown Reply Decl. Exh. 109 ("[T]here's no study that convincingly demonstrates a signal of cardiovascular events at very low does such as 200 per day."); Deposition Testimony of Dr. Nicholas Jewell at p. 130, Brown Reply Decl. Exh. 110 (when asked whether there is reliable scientific evidence to establish that 200 mg/d causes heart attacks and strokes he responded that the evidence is not sufficient "to be definitive"); Deposition of Dr. James M. Wright at pp. 83-84, 92, Brown Decl. Reply Exh. 106 (stating that it has not been proven that at 200 mg/d Celebrex increases the risk of heart attack because "we don't have enough information").

1. Dr. Neil Doherty

Plaintiffs' cardiology expert, Dr. Neil Doherty, nonetheless asserts "to a reasonable degree of medical probability that the 200 mg dose of Celebrex can increase the risk of MI's [heart attacks]." Written Direct Examination of Dr. Neil F. Doherty III ("Doherty Written Direct") at ¶ 18. He reaches his opinion by first identifying his conclusion—causation at 200 mg/d—and then cherry-picking observational studies that support his conclusion and rejecting or ignoring the great weight of the evidence that contradicts his conclusion. Dr. Doherty's opinion does not reflect scientific knowledge, is not derived by the scientific method, and is not "good science;" it is therefore inadmissible.

First, Dr. Doherty is not qualified to favor certain observational studies over the great weight of the epidemiologic evidence to give an opinion on causation. He is a clinical cardiologist who sees patients 95 percent of his physician time. He does not have any specialized epidemiology training. He has not published any research since 1992, and his 13 publications are unrelated to the subject matter of these lawsuits. He has never participated in an observational study of any kind. He is therefore not qualified to opine that one or two observational studies are correct while all the other studies (the studies that include 97 percent of the adverse cv events) are wrong. Moreover, he only became interested in Celebrex and cv risk after he was retained by plaintiffs in this litigation; indeed, although the issue of COX-2 inhibitors and adverse cv events has been well known since at least 2005, he did not discontinue prescribing Celebrex until after plaintiffs retained him as an expert in this case. Doherty Written Direct at ¶ 2. Dr. Doherty's opinion was developed for the purpose of this litigation. See Daubert II, 43 F.3d at 1317 ("One very significant fact to be considered is whether the experts are proposing to testify about matters growing naturally and directly out of research they have conducted independent of the litigation, or whether they have developed their opinions expressly for purposes of testifying.").

Second, apart from his lack of relevant experience and training (or because of it), the foundation of his opinion--wholly rejecting the McGettigan meta-analysis and the other observational studies that do not support his opinion--is not a scientifically valid methodology. For example, while he justifies his wholesale rejection of McGettigan on

the blanket ground that meta-analysis is inappropriate for observational studies, plaintiffs' other experts rely on such studies; indeed, Dr. Bennett testified that McGettigan is a "good study." Dr. Bennett Depo. at p. 187-88, Brown Reply Decl. Exh. 108. And the American Heart Association Committee that developed a "Science Advisory" on the use of NSAIDs also relied on McGettigan. Finally, Dr. Doherty testified that he prefers the Oxford Centre for Evidence Based Medicine ranking of the levels of evidence that a scientist should consider. Doherty Written Direct at ¶ 21-22. That ranking identifies systematic review, including meta-analysis, as the highest level for each category of evidence. Oct. 10 TR at 350.

Third, Dr. Doherty testified that the "strongest evidence" for his 200 mg/d opinion "is the Andersohn study published in Circulation in 2006."[fn4] [fn4=4 Frank Andersohn, et al., Use of First-and Second-Generation Cyclooxygenase-2-
Selective Nonsteroidal Anti-inflammatory Drugs and Risk of Acute Amuyocardial Infarction, Circulation, 2006 Apr 25; 113(16): 1950-7.] Doherty Written Direct at ¶ 18. He attempts to justify his heavy reliance on Andersohn by asserting that it is the "best designed" of all the observational studies. When asked why, however, Dr. Doherty responded only that the study is derived from the United Kingdom database which is among the most complete in the world. Oct. 10 TR at 309-10. He also mentioned that Andersohn is a prospective, rather than retrospective study. Id. at 310. But many of the other studies he rejects out of hand are also prospective, and he does not cite anything in the medical literature that suggests that it is a valid scientific method to prefer one study over many that have contradictory results simply because the study that supports the expert's conclusion utilized the United Kingdom database.

Fourth, Dr. Doherty's reliance on Andersohn as "the strongest evidence" of an increased risk at 200 mg/d is undermined by his own testimony that Andersohn's results do not make "biological sense." Oct. 10 TR at 363-64. Andersohn found the increased risk of heart attack was higher at shorter durations of use (less than three months) than at higher durations; indeed, there was no statistically significant association at durations greater than three months, a finding that directly contradicts Dr. Doherty's testimony that the risk of heart attack increases with duration of use. Oct. 10 TR at 359-61. Andersohn also found that the risk of heart attack is statistically significant in patients without cv risk factors, but is not statistically significant in patients with such risk factors. Id. at 364. Again, this finding directly contradicts Dr. Doherty's testimony that the risk of heart attack from Celebrex is greater in patients with heart disease. To conclude that Celebrex 200 mg/d causes heart attacks and strokes based on a study that does not make "biological sense" is not sound science.

Fifth, Dr. Doherty's opinion is based on his fundamental misunderstanding of Andersohn. Dr. Doherty testified that Andersohn is a cohort study and he "puts a lot more weight" into cohort studies as opposed to case control studies. Oct. 10 TR at 255, 309, 350. He repeatedly testified that he relies on Andersohn out of all of the available evidence because it is a good cohort study. See, e.g., id. at 313, 315. When he was confronted with Andersohn's own description of the study, however, Dr. Doherty

conceded that Andersohn is not a cohort study, but is instead "a case-control study nested within a cohort study." Id. at 352.

Dr. Doherty also insisted that Andersohn used cox proportional hazard analysis, the analysis most commonly used for cohort studies. Oct. 10 TR at 320-21, 355. On cross-examination, however, he could not identify where in the study the authors disclose that they used cox-proportional hazard analysis and Dr. Doherty pointedly did not clarify his testimony on re-direct. The Court has reviewed Andersohn and it does not indicate that the study authors used cox-proportional hazard analysis; rather, they used logistic regression which resulted in an "odds ratio," an analysis consistent with case control tudies. Dr. Doherty's fundamental misunderstanding of the study he "relied most strongly on" to support his opinion, Doherty Written Direct at ¶ 31, is perhaps explained by his inability to explain the difference between a cohort study and case control study "off the top of his head," Oct. 10 TR at 348, and his inability to define the cox proportional hazards model or explain logistic regression analysis. Id. In any event, as Andersohn is a case control study, Dr. Doherty's heavy reliance upon it is unreliable in light of his own blanket rejection of all of the case control studies showing no association between Celebrex 200 mg/d and cv risk on the ground that case control studies are not as reliable as cohort studies. Doherty Written Direct at ¶ 37.

While Andersohn is the "strongest evidence" supporting Dr. Doherty's opinion, he also cited an additional observational study, Gislason.[fn5] [fn5=5 Gunnar H. Gislason, et al., Risk of Death or Reincarnation Associated With the Use of Selective Cyclooxygenase-2 Inhibitors and Nonselective Nonsteroidal Antiinflammatory Drugs After Acute Myocardial Infarction, Circulation, 2006 June 27; 113(25): 2906-13.] Gislason, however, had few events and merely evaluated COX-2 inhibitors and the risk of a heart attack in patients who had already had a heart attack. Moreover, the study failed to control for smoking, a well-known risk for heart attack, as well as aspirin use, even though another of plaintiffs' experts, Dr. Maryilyn Rymer, criticized another observational study for not adjusting for aspirin use. Dr. Maryilyn Rymer Written Direct Testimony ("Rymer Written Direct") at ¶ 34. In light of these limitations, and the totality of the available evidence, Gislason does not salvage Dr. Doherty's opinion that Celebrex at 200 mg/d can cause heart attacks.

Dr. Doherty also relied on the "imbalance hypothesis" as evidence that it is biologically plausible that Celebrex causes heart attacks. This hypothesis asserts that COX-2 inhibitors as a class, that is, Vioxx, Bextra and Celebrex, create an imbalance in the arteries by blocking prostacyclin (an anti-clotting agent). Under this theory, the imbalance caused by ingesting a COX-2 will lead to an adverse cv event if the patient already has a risk factor, such as high blood pressure, smoking, or high cholesterol. Dr. Doherty argues that this hypothesis means that it makes sense that Celebrex increases the risk of heart attacks and strokes. He did not explain, however, how he reconciles this theory with Andersohn—the strongest evidence of his causation opinion—which showed a greater risk of heart attacks in patients with no cv risk factors.

In any event, both Dr. Doherty and Dr. Joel Bennett—plaintiffs' imbalance hypothesis expert—agree that the only way to prove the hypothesis is to look at the data from epidemiological studies. Oct. 10 TR at 373. For example, Dr. Bennett agreed that the only method available to determine how much Celebrex is needed (that is, what dose) to create an imbalance sufficient to cause a heart attack is patient studies. Oct. 9 TR at 209, 210. As is explained above, the patient studies do not demonstrate an association between Celebrex 200 mg/d and heart attack or stroke; therefore, the imbalance hypothesis—even if true—(and it is only one of many possible explanations for the apparent increased risk of heart attacks from COX-2 inhibitors at certain doses) does not support Dr. Doherty's opinion that Celebrex is capable of causing heart attacks at 200 mg/d.

2. Dr. Maryilyn Rymer

Dr. Maryilyn Rymer's testimony does not provide the missing link. Dr. Rymer is a neurologist and plaintiffs offered her as a stroke expert, essentially to opine that Celebrex causes strokes as well as heart attacks. In her written direct testimony she opines that "the totality of the scientifically reliable evidence supports that [Celebrex] can cause strokes and other cardiovascular events at all therapeutic doses, especially in those individuals who are high risk for cardiovascular events." Rymer Written Direct at ¶ 7. She admits that there is no data from randomized controlled trials to support her conclusion at 200 mg/d; instead, she primarily relies on (1) the imbalance hypothesis, (2) the same Andersohn study upon which Dr. Doherty relies, and (3) the Wellpoint data, an unpublished observational study of unknown design. In other words, Dr. Rymer, as does Dr. Doherty, ignores the vast majority of the evidence in favor of the few studies that support her conclusion.

The Court has already addressed the imbalance hypothesis and the Andersohn study, neither of which provide scientifically valid support for her opinion in light of the great weight of the epidemiologic evidence. It is worth adding, however, that Dr. Rymer's reliance on the Andersohn heart attack study is inconsistent with her criticism of the Andersohn stroke study. The latter study, performed by the same Andersohn as the heart attack study, indeed, it is the same study but focused on stroke rather than heart attack outcomes, found no statistically significant increased risk of stroke associated with Celebrex use at 200 mg/d. Dr. Rymer criticized the stroke study for not controlling for aspirin use and having a 10 percent error rate; yet the Andersohn heart attack study suffers from the same limitations.

Dr. Rymer relies heavily on an unpublished, non-peer reviewed study from a managed care organization ("the Wellpoint Report"). Dr. Rymer attaches to her written direct testimony a letter from Wellpoint to the FDA summarizing the results of the study. The letter discloses a relative risk from Celebrex use of 1.19 when the data is analyzed to control for "age and other cardiovascular risk factors;" however, this very low risk includes all doses of Celebrex. Moreover, the letter does not identify study design, the analysis used, or even the confidence intervals. Dr. Rymer admitted on cross-examination that the study also fails to account for critical compounding factors such as smoking. This unpublished, unreviewed study, which combines all doses of Celebrex,

and fails to adjust for critical compounding factors such as smoking, is not a scientifically valid basis for Dr. Rymer's rejection of all the other observational data--including meta-analyses--that do not show a statistically significant increase in the risk of heart attack or stroke at 200 mg/d.

Finally, Dr. Rymer cited Gislason, discussed above, and Brophy,[fn6] [fn6=James M. Brophy, The coronary risk of cyclo-oxygenase-2 inhibitors in patients with a previous myocardial infarction. Plaintiffs cited this study as being available at heart.bmj.com or at www.heartjnl.com.] as support for causation at 200 mg/d. Brophy, as Gislason, evaluated the risk of heart attack in patients who had already had at least one heart attack. Brophy, however, did not find a statistically significant increased risk of heart attack at 200 mg/d, even in these high risk patients. And while it did show a greater risk in the high risk population (although not a statistically significant risk), the higher risk found in Brophy and Gislason is contradicted by the results of at least nine other studies, including Dr. Doherty's "strongest evidence" of causation, the Andersohn heart study. Such data cannot reliably form the basis for rejecting the overwhelming pattern of evidence that fails to show any statistically significant risk at 200 mg/d.

3. Extrapolation

Dr. Doherty, and to some extent Dr. Rymer, also rely on studies of Celebrex 400 mg/d to support their opinion of causation at 200 mg/d. Although Dr. Doherty acknowledges that dose matters with Celebrex, he simply takes the relative risk point estimate of APC for 400 mg/d and cuts it in half (ignoring the confidence interval) to support his opinion that Celebrex at 200 mg/d can cause a heart attack. Oct. 10 TR at 304. When the Court asked Dr. Doherty if there is anything in the scientific literature to support such primitive extrapolation, he failed to identify any scientific support for his method other than his own judgment. Id. at 342-43, 378-79. He also admitted that there is no way of knowing what the confidence interval is for 200 mg/d under his unique methodology. Id. at 340-41. Such an unscientific, untested methodology cannot support the proffered opinion of causation at 200 mg/d, especially where, as here, Dr. Doherty agrees with all the other experts that there is dose effect with Celebrex.

Plaintiffs' reliance on In re PPA Products Liab. Litig., 289 F.Supp.2d 1230 (W.D. Wa. 2003), to argue that causation at 200 mg/d can be inferred from the 400 mg/d data is misplaced. In the PPA multi-district litigation the issue was whether PPA, a drug used in cough and cold and appetite suppressant products, can cause strokes. Plaintiffs' experts' opinion that PPA can cause strokes in persons of all ages and genders was based primarily upon a study of women ages 18 to 49. Id. at 1235-36. While men were not excluded from the study, their participation was too low to draw any conclusions. Id. at 1236. The defendants argued that the evidence was therefore insufficient to support the plaintiffs' experts' opinions that PPA can cause strokes in persons of all ages and genders. Id. at 1244. The district court disagreed.

The court found that "it is scientifically acceptable to extrapolate the conclusions of the [study] to these sub-populations." Id. at 1244. As to persons older than 49, the court

noted that there are no known studies that suggest that drugs get safer as persons get older; thus, it made common scientific sense to extrapolate the results of the study to persons over 49. Id. Plaintiffs' experts also attested to the "commonplace" practice of extrapolating between the genders based on "the historical exclusion of women from scientific studies." Id.

The justification for extrapolating drug effects between biologically similar demographic groups, however, does not logically extend to the argument that all doses of a compound are harmful; accordingly, plaintiffs' experts could not cite to a single piece of evidence that suggests that their experts' extrapolation is scientifically valid. To the contrary, with nearly all compounds there is usually a threshold that must be met before there is any harm; for example, even water can be harmful if consumed at certain amounts even though there is no harm at smaller amounts. Dr. Doherty claimed that the threshold for Celebrex must be 50 mg/d because that is the dose that is effective for pain relief. That "theory," however, is nothing more than Dr. Doherty's wholly untested, unpublished, and non-peer reviewed justification for his reliance on the 400 mg/d data. Moreover, the great weight of the evidence does not support the extrapolation, that is, studies show that there is no statistically significant association between Celebrex 200 mg/d and the risk of strokes or heart attacks.

Instead of citing evidence that supports such extrapolation, plaintiffs complain that the evidence of harm at 200 mg/d does not exist because Pfizer did not initiate long term randomized trials at such dose. Such a trial, known as PRECISION, is now underway, but the results will not be available for some time. Plaintiffs cite no case, however, that suggests that they can satisfy their burden of proof based on a lack of evidence; plaintiffs filed these lawsuits and plaintiffs carry the burden of proving today based on currently available scientifically valid evidence that Celebrex can cause heart attacks or strokes at 200 mg/d.

Plaintiffs have not met their burden. In so finding, the Court is relying on the evidence presented by plaintiffs; it has not considered Pfizer's own meta-analyses. And the Court's ruling is not mandated by the lack of randomized clinical trials that show an association at 200 mg/d; plaintiffs could still meet their burden in the absence of such evidence. See Kennedy v. Collagen Corp., 161 F.3d 1226, 1228 (9th Cir. 1998). However, the opinion of Dr. Doherty and Dr. Rymer that Celebrex 200 mg/d increases the risk of heart attacks or strokes is not based on a scientific valid methodology; instead, these experts ignore the great weight of the observational studies that contradict their conclusion and instead rely on the handful that appear to support their litigation-created opinion. As the Court explained above, their reasons for doing so are not supported by scientifically valid reasons or methodology. In the words of the Supreme Court, the "analytical gap" between the data and these experts' conclusion is simply too great to make the opinion admissible. General Elect. Co. v. Joiner, 522 U.S. 136, 146 (1997).

B. 400 mg/d

Pfizer's motion to exclude expert testimony that Celebrex 400 mg/d is capable of causing heart attacks or strokes is defeated by APC, a large, long-term, randomized, placebo-controlled, double-blind, multi-center clinical trial that was halted after 33 months because it demonstrated a statistically significant risk of heart attack, stroke, and heart failure at 400 mg/d (2.6 percent hazards ratio with a confidence interval of 1.1 to 6.1) and 800 mg/d (3.4 percent hazards ratio with a confidence interval of 1.5 to 7.9).[fn7] [fn7=7 Scott D. Solomon, et al., Cardiovascular Risk Associated with Celecoxib in a Clinical Trial for Colorectal Adenoma Prevention, N. Engl. J. Med. 2005 Mar 17; 352(11): 1071-1080.] The study, co-sponsored by the National Cancer Institute and Pfizer, was designed to compare Celebrex with placebo for the prevention of colorectal adenomas (polyps). The study included a "cardiovascular safety committee" that developed guidelines to evaluate cardiovascular safety. On December 16, 2004, on the basis of the results then available as well as studies of Vioxx and Bextra, and on the recommendation of the safety committee, the APC steering committee stopped the trial. This randomized, placebo-controlled, double-blinded study with an independent committee evaluating cardiovascular endpoints is the "gold standard" of epidemiologic evidence and supports plaintiffs' experts' testimony that Celebrex at 400 mg/d is capable of causing heart attacks or strokes.

Pfizer nonetheless contends that plaintiffs' experts' opinion must be excluded because (1) APC was stopped early, and (2) its results have not been replicated by two other randomized controlled trials that evaluated Celebrex 400 mg/d: ADAPT and PreSAP.

The Court is unconvinced that plaintiffs' experts cannot base their opinions on APC because it was stopped early (after 33 months). The APC steering committee halted the trial because the evidence of harm was so significant. To exclude reliance on such studies under these circumstances would mean the more harmful the drug the more difficult it is to prove harm. While such studies must be closely scrutinized due to their early termination, Pfizer's argument goes to the study's weight; Pfizer has not shown that it is not scientifically valid for plaintiffs' experts to rely on the results. Moreover, ADAPT and PreSAP, two studies upon which Pfizer relies, were also halted early because of the APC results.

The Court is also not persuaded that the failure of ADAPT and PreSAP to replicate APC's results means plaintiffs' expert opinion on 400 mg/d is inadmissible. ADAPT was a randomized, placebo-controlled clinical trial designed to evaluate naproxen and Celebrex 400 mg/d (200 mg twice daily) and the prevention of Alzheimer's dementia.[fn8] [fn8=8 ADAPT Research Group, Cardiovascular and Cerebrovascular Events in the
Randomized, Controlled Alzheimer's Disease Anti-Inflammatory Prevention Trial (ADAPT), PLoS Clin Trials 2006; 1(7): e33.] ADAPT found a hazards ratio for Celebrex of 1.10 percent with a confidence interval of .67 to 1.79, that is, no statistically significant association. The study authors, however, cautioned that there are several limitations to their data. First, ADAPT was not designed to detect differences in cardiovascular and cerebrovascular risks and, unlike APC, it did not include a separate cardiovascular safety committee tasked solely with evaluating cardiovascular outcomes. Second, and,

according to the authors, the largest limitation of the data is the small number of cardiovascular events. Third, an editorial comment accompanying the study suggests that because study participants eligible to join the trial were required to have a family history of Alzheimer's disease, it is possible the study participants' risk factors differed from the general population. The results of ADAPT need to weighed with the APC results, but ADAPT's conclusions do not make reliance on APC scientifically invalid.

The results of PreSAP, a randomized controlled study with fewer participants than ADAPT or APC, also did not replicate the APC results. PreSAP, as APC, was designed to evaluate Celebrex's effect on the occurrence of colorectal adenomas. Preliminary results from that study did not show a statistically significant increase in cv risk for patients taking Celebrex 400 mg/d, but did not exclude the possibility of a hazards ratio similar to that demonstrated by APC. In addition, PreSAP used the same independent cardiovascular safety committee as APC to assess the risk of Celebrex on adverse cv events. Accordingly, the data from both trials were synthesized to produce a combined estimate of risk of cardiovascular death, heart attack, stroke or heart failure of 1.9 with a confidence interval of 1.1 to 3.1; in other words, combining the raw data showed a statistically significant increase in risk.[fn9] [fn9=9 Scott D. Solomon, et al., Effect of Celecoxib on Cardiovascular Events and Blood Pressure in Two Trials for the Prevention of Colorectal Adenomas, Circulation, 2006 Sep 5; 114(10): 1028-35.] The study authors combined APC 400 mg/d and 800 mg/d with PreSAP 400 mg/d because the confidence intervals for 400 mg/d and 800 mg/d substantially overlapped. While the weight to be given to this evidence can be argued, in light of this evidence, and the Kearney meta-analysis which found a relative risk greater than one with a confidence interval that barely crossed one, the Court cannot conclude that expert opinion that Celebrex 400 mg/d is capable of causing heart attacks and strokes is scientifically invalid.

C. Whether Celebrex Causes Heart Attacks or Strokes More Than Three Days After A Patient Stops Taking It

Plaintiffs do not dispute that Celebrex is not capable of causing hearts attacks or strokes more than three days after a patient stops taking it and they have offered no expert opinion to the contrary. Accordingly, there is no proposed expert testimony on this issue for the Court to exclude.

D. Remaining Issues

1. Strokes

The issue as to whether Celebrex is capable of causing strokes is close. Plaintiffs rely on the testimony of Dr. Rymer, a neurologist and the Medical Director of the Saint Luke's Brain and Stroke Institute at Saint Luke's Hospital in Kansas City, Missouri. She testified that the mechanism of and risk factors for throembolic strokes (excluding cardiogenic embolism) and heart attacks are the same; thus, if Celebrex causes an increased risk in heart attacks it also increases the risk of strokes. Rymer Written Direct ¶ 11-13. Dr. Rymer's testimony is supported by the published literature as nearly all

studies of COX-2 inhibitors and cv risk lump strokes together with heart attacks. For example, the Kearney meta-analysis of clinical trials identified the relative risk for "serious vascular events," defined as heart attack, stroke, or vascular death. Indeed, even Pfizer's expert, Dr. Packer, considers the risk of heart attacks and strokes together, and Pfizer does not dispute Dr. Rymer's testimony as to the similar mechanism of heart attacks and strokes.

Pfizer nonetheless asserts that Dr. Rymer's testimony is inadmissible because the randomized controlled trials and observational studies that do separately report strokes and heart attacks do not suggest an association between Celebrex at any dose and strokes. Dr. Rymer explains, however, that none of the randomized controlled studies was designed to look for stroke outcomes, and strokes occur far less often than heart attacks; the studies simply were not designed to find an association or not.

While there is some epidemiologic evidence to dispute her mechanism testimony, that is, evidence that suggests that even though heart attacks and certain strokes are caused by the same mechanism Celebrex does not cause both, there is also some evidence to support her testimony. On the current record the Court does not find that Dr. Rymer's testimony is scientifically invalid and inadmissible. See Daubert, 509 U.S. at 596 ("Vigorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof are the traditional and appropriate means of attacking shaky but admissible evidence.").

2. Duration

The Court also denies Pfizer's motion to exclude testimony that Celebrex is capable of causing heart attacks or strokes only after 33 months of continuous use. Because a statistically significant association did not appear in APC until after 33 months does not mean as a matter of scientific fact that none of the adverse cv events that occurred after a shorter duration were not caused by the patient's ingestion of Celebrex.

3. Specific Causation

Finally, Pfizer asks the Court to "exclude any opinion that Celebrex caused an individual plaintiff's heart attack or stroke absent a relative risk that exceeds 2.0." This is a question of specific causation as to particular plaintiffs; as the Court does not have before it evidence as to any specific plaintiff the Court declines to grant Pfizer's motion.

II. Plaintiffs' Motion to Exclude

Plaintiffs move to exclude the meta-analyses performed by Pfizer's experts. Plaintiffs' experts did not perform any of their own meta-analyses; instead, plaintiffs attack Pfizer's experts' methodologies. Plaintiffs' motion is denied. All of plaintiffs' arguments go to the weight a trier of fact gives to the meta-analyses. Plaintiffs have not shown that the methods employed by Pfizer's experts are not based on good science.

Plaintiffs also move to exclude Dr. Packer from testifying as to an alternative theory to the imbalance hypothesis. Dr. Packer's explanation, which accounts for the difference in outcomes between Vioxx and Celebrex, is based on increased blood pressure, a theory actually supported by plaintiffs' expert Dr. Rymer. In any event, Dr. Packer's testimony satisfies Daubert.

CONCLUSION

In Daubert, the Supreme Court held that federal judges perform a gatekeeping role, 509 U.S. at 597, and "to do so they must satisfy themselves that scientific evidence meets a certain standard of reliability before it is admitted." Daubert II, 43 F.3d at 1316. Plaintiffs' expert testimony that Celebrex 200 mg/d can cause heart attacks or strokes does not meet that standard. Dr. Doherty, a clinical physician with no relevant research experience and who developed his opinion for the purpose of testifying, bases his opinion on a study that he fundamentally misunderstood, is counter to the great weight of the evidence, and, by his own admission, does not make biological sense. The Court cannot find that his opinion is good science. Dr. Rymer's 200 mg/d opinion is also not good science. She ignores all the evidence that contradicts her litigation-created conclusion and instead bases her opinion on the same cherry-picked study as Dr. Doherty, even though that study suffers from the exact same limitations that caused her to reject other studies that do not support her conclusion. She also relies on an unpublished, non-peer reviewed study that does not disclose its design or confidence intervals. If the Court's gatekeeping function means anything, it must mean that these unreliable opinions are not admissible to prove general causation at 200 mg/d.

In all other respects, and for the reasons explained above, the parties' motions are denied.

IT IS SO ORDERED.
/s/
Dated: November 19, 2007 HONORABLE CHARLES R. BREYER
UNITED STATES DISTRICT JUDGE