Manufacturing the truth: From designing clinical trials to publishing trial data

MARGARET WHITSTOCK

Abstract
This paper expands on some of the points made by Deepak Natarajan on techniques used in designing clinical trials of new drugs to ensure favourable outcomes. It also considers the nexus between the manufacturers of new drugs and the publishers of medical journals in which edited versions of these favourable outcomes are presented to the medical fraternity.

The argument will be illustrated by referring to the clinical trials of rofecoxib (Vioxx®) and etoricoxib (Arcoxia®). Both these drugs are COX-2 selective non-steroidal anti-inflammatory drugs (NSAIDs) manufactured by Merck & Co. Because of the unparalleled access to Merck’s internal confidential documents, due to the subpoenaing of these documents by government and private individuals in civil and criminal actions, we are still learning about the company’s unconscionable acts. What we learn can inform our judgement concerning published reports of both new and old drugs.

The randomised controlled trial in the research process
Most national jurisdictions require that a new drug must demonstrate safety and efficacy via at least one randomised controlled trial (RCT), as this is deemed the most scientific method to evaluate a new drug. In the USA, for example, the only clinical trial format that is acceptable for demonstrating safety and efficacy is the RCT, and the US Food and Drug Administration (FDA) will not approve a new drug that has not been evaluated in this format.

However, in practice, this sought-after objectivity is shown to be an artifice. For a pharmaceutical manufacturer, conducting clinical trials of a new drug is a tactical exercise to support the approval to market that new drug. As an eminent epidemiologist, Alvan Feinstein, has said: “A randomised clinical trial is designed and analysed according to strategic policies about what questions the trial is intended to ask, what answers are to be obtained, what is to be done with the data, and who is to be convinced by the results” (1). The selection and definition of the clinical problem, the variables to be evaluated, the participating subjects, the procedures and measuring techniques, the nomination of what will be considered as an outcome, the statistical analyses to be performed, and the interpretation of those analyses – all of these are made from a position of pre-specified interest (2).

This paper discusses some of the clinical trial techniques identified by Natarajan (3) in which pre-specified interests influence how an RCT is designed and reported:
1. comparison with a treatment known to be inferior;
2. non-compliance with the recommended dosages of comparator drugs to the advantage of the trialled drug;
3. choosing to submit for publication only the selected endpoints that show the trialled drug to advantage; and
4. choosing to submit for publication only favourable subgroup analyses.

The relationship between manufacturers and medical journals
During the development of a new drug, manufacturers sponsor (or act as authors of) articles on the clinical trials of the new drug, and these articles are submitted to medical journals. Publication of these articles acts as an essential tool for advertising to the medical community who will be the future prescribers of the new drug. Richard Smith, a former editor of The BMJ, considered that medical journals are “an extension of the marketing arm of pharmaceutical companies” (4). To illustrate, at an estimated cost of up to US$ 836,000, Merck & Co. purchased 900,000 reprints of the VIGOR trial article from the NEJM to circulate to doctors to promote Vioxx® (5,6).

Wilson (7) argues that in the public interest, the potential for capture of medical journals represented by this commercial role must be acknowledged and addressed.

Sometimes articles are not submitted for publication until after a new drug is granted regulatory approval to market (8,9), as holding over publication until post-approval reduces...
the likelihood of peers finding problems that may affect the approval process (8,10). Once a new drug is approved for marketing, the number of publications on the trials of the drug usually increases (9). However, there is no effective mechanism to ensure that what is published in medical journals accurately reports the data from the clinical trials submitted for the approval of that drug.

This has had a considerable impact on the value of the peer-review process, as reviewers see only the data that pharmaceutical manufacturers have been willing to provide, and this frequently means that adverse data are omitted or played down. Smith observes that peer review has become an uncertain means to assess research papers, particularly those reporting on clinical trials (11). He notes that there is little evidence on the effectiveness of peer-review and considerable evidence on its defects (11). Smith’s viewpoint is supported by Richard Horton (editor of The Lancet) and Marcia Angell (former editor of the NEJM) (4). Brophy argues that superficial peer-reviewing in publications reporting on Vioxx® contributed to misleading inferences and conclusions that ultimately put public safety at risk (12).

**Development of the COXIBS**

Cyclooxygenase (COX) was originally understood to exist as a single enzyme, albeit with a range of expression in different tissues. Research published in 1991 (13) and 1992 (14) confirmed that there are two isoforms of cyclooxygenase, COX-1 and COX-2. It was hypothesised that COX-1 was responsible for protecting gastric mucosa, and COX-2 was responsible for inflammatory responses to tissue damage. It was further hypothesised that if a drug could be developed in which COX-2 could be selectively inhibited, while leaving the action of COX-1 unopposed, then the outcome would be that tissue inflammation levels would be reduced, and gastric mucosa protection would be maintained.

This opened up research to develop drugs that would selectively inhibit the action of COX-2, the isoform believed to be responsible in inflammatory conditions such as osteoarthritis or rheumatoid arthritis. Pfizer’s celecoxib (Celebrex®) was the first such drug on the market. Merck developed two COX-2 selective NSAIDs, Vioxx® and Arcoxia®. The FDA approved Vioxx® in May 1999. Arcoxia® had been approved for marketing in the UK in 2002, and was in use in Latin America and the Asia Pacific region, but it had not been approved by the FDA for marketing in the USA.

**Identification of risks with COXIBS – Merck’s Vioxx® is withdrawn**

Subpoenaed evidence has indicated that by 1998, Merck scientists had already recognised that COX-2 inhibitors carried a risk of cardiovascular (CV) and cerebrovascular morbidity and mortality (15). The level of risk appeared to be correlated with the level of the inhibition of COX-2 activity, which can be described by the COX-1 to COX-2 selectivity ratio, though the process underlying this risk was not clearly understood. From the perspective of COX-2 selectivity ratios, both Vioxx® and Arcoxia® could be considered to hold a significantly greater risk for patients when compared with Celebrex®.

Notwithstanding the growing awareness of the association of COX-2 inhibition with CV and vascular morbidity, Merck continued with testing Vioxx® because it wanted to overtake Celebrex® in the highly profitable COXIB market place. The intention was to offer a stronger pain-reliever for arthritis patients, while also offering protection from damage to gastric mucosa. Merck was not prepared to allow the growing understanding about the risks associated with COX-2 inhibition to derail its application to the FDA for approval to market Vioxx® in the USA. The application was made in November 1998. Vioxx® was given Fast Track approval status, and was approved for marketing in the USA in May 1999. Applications to market were also made in other jurisdictions, using the same data as submitted to the FDA.

Eventually, it was realised that while inhibition of COX-2 activity could assist in reducing inflammation, it also affected both vasoconstriction and platelet aggregation, and the subsequent development of blood clots, leading to thrombotic events such as heart attacks (myocardial infarctions [MIs]) or cerebrovascular strokes. The actual extent and complexity of the counterbalancing roles of both the COX-1 and COX-2 isoforms had not been clearly understood, and the clinical trials which had been designed specifically to prove the anti-inflammatory and gastroprotective capacity of COX-2 inhibitors were not designed to consider these complex issues.

On September 30, 2004, Merck withdrew Vioxx® from the market worldwide, citing safety reasons as shown in the data from the APPROVe clinical trial which had commenced in February 2000. By the time it was withdrawn, there had been many thousands of adverse CV and cerebrovascular events, including thousands of deaths, in patients taking Vioxx®, particularly at the highest daily dosage of 50 mg.

**1. Comparing a drug with a known inferior treatment: Merck and Arcoxia®**

**Arcoxia® with naproxen as comparator**

In the face of accumulating adverse event data concerning Vioxx®, Merck needed to show that its remaining COX-2, Arcoxia®, was safe from a CV perspective. The company had commenced trialling Arcoxia® in 1999. By 2002, it had conducted clinical trials of Arcoxia® in which naproxen (Naprosyn®, Aleve®), a traditional NSAID, was used as the comparator. Unfortunately for Merck, these studies found that patients in the Arcoxia® arms experienced a higher incidence of adverse thrombotic CV events than did patients in the naproxen arms. But rather than acknowledge the increased risk posed by Arcoxia®, it claimed, post hoc, that the lower incidence of thrombotic CV events in the naproxen arms was due to a cardioprotective capacity of naproxen. One of the trial reports described naproxen as “a potent and sustained inhibitor of platelet aggregation at therapeutic doses” (19).
Merck had previously tried this baseless and counterfactual explanation concerning the cardioprotective quality of naproxen when reporting on trials with Vioxx® versus naproxen. The FDA had sent a warning letter to Merck, in 2001, criticising the drug maker for promoting the idea that naproxen was cardioprotective without explaining that this was still hypothetical, that it had not been demonstrated by substantial evidence, and that there was another reasonable explanation, and that explanation was “that Vioxx® may have pro-thrombotic properties” (21).

**Arcoxia® with diclofenac as comparator – The MEDAL Program**

When Merck’s claim that naproxen was cardioprotective was rejected, it undertook new trials of Arcoxia® with a different comparator, diclofenac sodium (Voltaren®). Diclofenac, a traditional NSAID, had been used worldwide for many years for relief of pain associated with inflammatory conditions such as osteoarthritis. Its approval in the USA in 1988 was granted well before issues regarding the safety profile of all NSAIDs became apparent.

The MEDAL Program, consisting of three trials in which Arcoxia® was compared with diclofenac, commenced in 2002. The three trials were EDGE I, EDGE II, and the MEDAL study. By the time the MEDAL study, which compared Arcoxia® to diclofenac, commenced in the USA, much would have been known about diclofenac (22). Merck had used that drug as a comparator in at least two RCTs of Vioxx®. In one of these, in which diclofenac was administered at 33% above the manufacturer’s recommended level, the rate of thrombotic CV events in the diclofenac arm exceeded the numbers in both the Vioxx® 12.5 mg trial arm, and the Vioxx® 25 mg arm (12.5 mg – 25 mg is the recommended dosage range for Vioxx®) (23).

In a 2001 study, diclofenac had been assessed as equivalent to Celebrex® in respect of COX-2 selectivity (24), and other research available at the time of the choice of diclofenac as the comparator in the MEDAL studies would have indicated that the drug was increasingly suspected of having thrombotic CV risk (28), particularly at the 150 mg daily dosage, the dosage chosen for the MEDAL trials. Merck’s use of diclofenac as the comparator was unethical.

Merck’s continued trialling of a new COX-2 selective NSAID (Arcoxia®) after the safety withdrawal of Vioxx® is also ethically suspect, as the company was uniquely well-placed to understand the CV problems with COX-2 selectivity. Even before the Arcoxia® versus diclofenac trials had commenced, a 2001 publication focusing on two major clinical trials, one with Vioxx® and one with Celebrex®, noted the potential for adverse CV events associated with COX-2 inhibitors (26).

**Outcomes of the MEDAL non-inferiority study design**

The primary hypothesis of the MEDAL Program was that, based on confirmed thrombotic CV events, Arcoxia® would be non-inferior or, as stated, “no worse than,” diclofenac (30). The reason given for choosing the non-inferiority design was that it would not be possible to assess absolute risk as this would require a placebo arm, and this was not ethical in a long-term arthritis study (31). However, use of an active comparator is acceptable, and resorting to a non-inferiority design is not the only way to take account of the possible unacceptability of a placebo arm. The trial investigators set a pre-specified criterion for determining non-inferiority, namely an upper bound of a two-sided 95% confidence interval below 1.30 (30). The published report on the MEDAL trial, the largest study in the MEDAL Program, stated that: “Hypertension is an important risk factor for CV disease and in this study etoricoxib [Arcoxia®] was associated with a significantly greater number of discontinuations due to hypertension-related AEs [adverse events]” than occurred in the diclofenac arm (32).

The choice of the non-inferiority study design meant that the MEDAL trials could only establish the relative CV risk of Arcoxia® as compared with the CV risk of diclofenac. But what was the level of risk of diclofenac? As noted earlier, diclofenac was itself showing potential for thrombotic CV risk, and comparing the COX-2 selective Arcoxia® with another potential COX-2 selective NSAID would mean little in terms of a safety assessment.

Non-inferiority trials cannot establish absolute risk. All they can establish is that the drug described as non-inferior has an equal or lower numbers of adverse events for patients than the drug with which it is being compared. Investigating the drug which is used as the basis for comparison is an essential first step when assessing the safety of a new drug (or any drug) that relies on non-inferiority trial data. Non-inferiority trials can mask the real safety risks of the “non-inferior” drug.

As the FDA’s David Graham noted: “From the perspective of patient safety and rational therapeutics, naproxen [which had a lower CV risk], not diclofenac, should have been the reference drug in MEDAL. Had that been so, it is highly likely that etoricoxib [Arcoxia®] would have been shown to be no different than its first cousin rofecoxib [Vioxx®] with respect to cardiovascular risk” (33). Merck’s Arcoxia® trials were designed strategically to get the answers that would allow the drug to gain FDA approval to market in the USA; they were not designed to actually prove the safety of Arcoxia®. In 2007, after due consideration of the clinical trial evidence and the issues discussed above, the FDA decided not to approve Arcoxia® for marketing in the USA (28).

**2. Non-compliance with recommended dosages of comparator drugs to the advantage of the trialled drug: Merck and Vioxx®**

The early clinical trials of Vioxx® focused on comparing the clinical efficacy, safety and tolerability of this drug when compared with various traditional NSAIDs. Early RCTs showed that Vioxx® had an unexpected capacity to cause adverse thrombotic CV events (and adverse thrombotic events in other body systems, e.g. cerebrovascular accidents or strokes). These thrombotic events were in marked excess to placebo and to traditional NSAID comparators. In his testimony given at the
US Senate Hearing “FDA, Merck and Vioxx: putting patient safety first?” Gurkirpal Singh, a former Merck consultant, stated that by November 1996, Merck scientists had recognised that Vioxx® did not inhibit platelet formation, and this seemed to be associated with a risk of heart attack (MI) in studies comparing Vioxx® with other painkillers (34).

Following this realisation, Merck’s clinical trials of Vioxx® no longer focused on adjudicating CV outcomes; or, if this was done, these outcomes were not published at the same time as the other outcomes (e.g. with the VIGOR trial, where CV data were collected but were not analysed, and only briefly referred to in the VIGOR trial report in the NEJM (35). Merck’s claims for the superiority of Vioxx® then focused on the capacity of Vioxx® to reduce damage to the gastric mucosa, and reduce arthritis pain and improve joint functionality (directly related to the capacity of Vioxx® to selectively inhibit COX-2-related inflammation typical of arthritis).

**Testing the gastroprotective capacity of Vioxx®**

Traditional non-selective NSAIDs inhibit the action of both COX-1 and COX-2 isoforms of the cyclooxygenase enzyme. Traditional NSAIDs were taken by patients because of their capacity to inhibit the COX-2 isoform responsible for the inflammatory pain associated with such conditions as osteoarthritis or rheumatoid arthritis, even though, by inhibiting the COX-1 isoform at the same time, traditional NSAIDs posed the risk of varying degrees of damage to the gastric mucosa, from digestive discomfort to erosions, perforations, ulcers or bleeding.

Merck designed its clinical trials to show the benefits of Vioxx® in reducing gastric adverse events, and contrasted these with the adverse effects of traditional NSAIDs on the gastric mucosa. The higher the dosage in the trial arms using various traditional NSAID comparators, the more likely the contrast could be made evident, with the traditional NSAID comparators shown at maximum disadvantage and Vioxx® shown at maximum advantage. Vioxx® trial arms were also subject to supratherapeutic dosing levels, with the intention of showing that, even at high levels, the drug showed no adverse effects on gastric mucosa.

**Supratherapeutic dosing with the ibuprofen comparator in Vioxx® clinical trials**

The Laine et al 24-week trial of Vioxx® to treat osteoarthritis (36) was submitted to the FDA in 1997, but not published until 1999. The trial arms included a placebo arm, two trial arms with different dosages of Vioxx® (one using 25 mg, the upper limit of the recommended dosage range, and the other using 50 mg, which is twice the recommended upper limit for longer term use, as was the case with the Laine et al. trial), and one trial arm with the traditional non-selective NSAID ibuprofen (Brufen®, Neurofen®). The manufacturer’s dosage instructions for the ibuprofen comparator at that time were that the usual daily dosage was 1200 mg – 1600 mg, with provision for a maximum of 2400 mg. In the Laine et al. trial, the maximum 2400 mg dosage of ibuprofen was used. The trial report commented that Vioxx® caused “markedly less gastrointestinal (GI) ulceration than ibuprofen, with ulcer rates comparable to placebo” (36).

In the Laine et al trial, the investigators used endoscopic examinations to assess whether ibuprofen caused GI adverse events including erosions, perforations, ulcers or bleeding. However, there was an important question: would a gastric morbidity as observed via an endoscope develop into an outcome requiring clinical intervention? As Laine et al commented, “there is much controversy regarding whether endoscopic ulcers are surrogates for clinical outcomes”; it had not yet been shown that endoscopic studies could “definitively determine if the use of COX-2-specific inhibitors will indeed significantly decrease clinically important ulcers” (36).

The Day et al six-week trial of Vioxx® to treat osteoarthritis (37) focused on the clinical efficacy of Vioxx® in controlling the symptoms of arthritis. It was completed and submitted to the FDA in 1998, but not published until 2000. The trial arms were placebo, two Vioxx® arms (12.5 mg and 25 mg), and ibuprofen 2400 mg (the maximum allowable dosage). The investigators reported that clinical adverse events that led to withdrawals from the trials were most common in the ibuprofen group, and were “mostly accounted for by adverse experiences related to the GI tract”. The investigators commented that to firmly establish an improved safety profile of Vioxx® in contrast to non-selective NSAIDs, it was important that the safety profiles be compared using doses that provide equivalent efficacy; nevertheless, the investigators acknowledged that they had chosen the maximum dosage of ibuprofen rather than assess ibuprofen at the recommended usual dosage levels.

**Some discussion points on the Laine et al. and Day et al. clinical trials**

First, no data are provided either in Laine et al or Day et al trial about equivalent efficacy at anything other than the maximum dosage of ibuprofen. It could be possible that the normal dosage of ibuprofen may have provided equivalent efficacy. Second, by introducing ibuprofen at its maximum dosage, the investigators were able to take advantage of the potential for adverse GI effects known to be associated with that maximum dosage. Third, to be enrolled in either the Laine et al. or Day et al. study, patients with osteoarthritis of the knee or hip were required to have been currently treated with a traditional non-selective NSAID or paracetamol (acetaminophen) before the commencement of the trial. The investigators reported that an analysis was undertaken to see if the pre-trial medication (NSAID or paracetamol) resulted in different treatment effects. However, this analysis was not reported. If patients previously on traditional NSAIDs were randomised to the ibuprofen trial arm, and the patency of their GI mucosal lining was unknown, there is the possibility that for patients with some prior level of GI damage, a high 2400 mg per day could have precipitated the GI problems reported.

**Supratherapeutic dosing with the naproxen comparator in Vioxx® clinical trials**

The VIGOR trial (Vioxx Gastrointestinal Outcomes Research) commenced in January 1999, with the stated aim of comparing
the rates of upper GI toxicity in patients prescribed Vioxx® for rheumatoid arthritis with those for patients prescribed the traditional non-selective NSAID naproxen (Naprosyn®, Aleve®) (35). The trial started before Vioxx® was approved by the FDA to be marketed in the USA (which took place in May 1999). The VIGOR trial was completed in March 2000, and submitted to the FDA in that same year.

The investigators, Bombardier et al, used 1000 mg naproxen as the comparator, the dosage at the upper bound of the manufacturer's daily dosage instructions for long-term treatment with naproxen. There were 72 confirmed upper GI adverse events in the Vioxx® trial arm, and 148 in the naproxen arm, indicating that the GI performance of Vioxx® was superior to naproxen. However, these numbers form only a small proportion of the 1185 withdrawals from the Vioxx® arm, and the 1149 withdrawals from the naproxen arm, so there were other issues contributing to participants' withdrawals. The major issue with the VIGOR trial was the significant difference between the occurrence of thrombotic CV events (in particular, MIs) in the two trial arms. This issue will be discussed in section 3.

**Supratherapeutic dosing with the investigational drug in the Vioxx® clinical trials**

In the early trials of Vioxx®, doses of the investigational drug were often recorded as being greater than the dosage necessary to control symptoms of osteoarthritis. Examples are Laine et al. (36), in which the 50 mg Vioxx® dosage was recorded as “2–4-times the dose shown to relieve the symptoms of osteoarthritis”; or Erich et al (38), in which the 125 mg dosage was described as “5-fold higher than the 25 mg dose sufficient for meaningful clinical efficiency in this study”. While the 50 mg dosage was used regularly in Vioxx® trials, there were also a number of trials in which the dosage used was 12.5 mg and/or the 25 mg.

The FDA approval in May 1999 specified 12.5 mg and 25 mg of Vioxx® as efficacious dosages for the control of symptoms of osteoarthritis, management of acute pain and dysmenorrhoea. A 50 mg dosage was also approved for acute pain, but not for long-term use (39). Though the VIGOR trial commenced before the FDA approval of Vioxx®, Bombardier et al were aware of the dosages to be approved by the FDA, as the investigators identified the 50 mg dosage used in VIGOR as being “twice the maximal [daily] dose approved by the FDA for long-term use” (35). The dosage used for the VIGOR trial was 50 mg per day for up to 12 months (35), which counts as long-term use. There was only one Vioxx® arm.

The FDA 1999 approval for Vioxx® was for the treatment of osteoarthritis. The VIGOR trial used participants with rheumatoid arthritis, which has a different aetiology, though joint inflammation is common to both forms of arthritis. Rheumatoid arthritis is also independently associated with an increased rate of thrombosis, particularly MIs, which added a further risk for the VIGOR participants (40).

There were more MIs in the Vioxx® arm than in the naproxen arm, and the supratherapeutic dosage of Vioxx® may well have contributed to the numbers of these MIs. Because there was only one Vioxx® dosage used in the trial, there were no data about the possible occurrence of MIs at lower dosage levels. This would have been useful, particularly in view of the fact that there was a 5-fold increased risk of MIs in the VIGOR Vioxx® arm compared with the naproxen arm.

As Turner et al note, medical decisions are based on an understanding of publicly reported clinical trials, and if the evidence presented in these journals is biased, then decisions based on this evidence may not be the optimal decisions (41). The selective publication of clinical trials can lead to unrealistic estimates of drug effectiveness and alter the apparent risk–benefit ratio (41). In the case of Vioxx®, the manufacturer's purchase and circulation of 900,000 copies of the VIGOR paper published in the NEJM (4,5,6,20) played a crucial role in foregrounding the 50 mg dosage of Vioxx®. This 50 mg dosage came to be the dosage that caused the most MIs in the USA, even though the FDA approval stated that 12.5 mg and 25 mg of Vioxx® were efficacious dosages.

When assessing the value of new drugs, potential prescribers can check the approved dosage of a new drug as provided by the manufacturer in the container or in the manufacturer's online prescribing information. Other sources to check are clinicaltrials.gov, or www.fda.gov (key in drug approval package [drug name] in the search window at the upper right-hand side of the screen.) The FDA's online Drug Approval Package (DAP) database presents data from all the trials submitted to the FDA as part of the FDA's approval process – not just those trials which have been chosen by manufacturers for publication in medical journals. The FDA DAPs provide more information than clinicaltrials.gov.

**3. Publishing selected endpoints to the advantage of the trialled drug: Merck, Vioxx® and the VIGOR trial**

When the VIGOR trial data were published in the NEJM on November 23, 2000 (35), the authors, Bombardier et al, focused primarily on the GI advantage of Vioxx®, and not on its CV safety. The VIGOR paper reported that CV data had been collected, but stated that these were not going to be analysed in the published paper. The reason given was that a separate analysis of the CV data was not specified in the study design. By choosing to focus predominantly on the GI data, and by adopting a self-imposed limitation on the reporting of CV data, the investigators made decisions that proved to be greatly to the advantage of Vioxx®, as the dangers of the drug remained hidden for longer.

In their VIGOR paper, the researchers acknowledged that "because highly selective cyclooxygenase inhibitors [such as rofecoxib] do not inhibit platelet aggregation [which can lead to thrombotic CV events] ... there was a possibility that the incidence of thrombotic cardiovascular events would be lower among patients treated with non-selective cyclooxygenase inhibitors [in this trial, the traditional non-selective NSAID
na pro xen)” (35). This statement indicates that the VIGOR investigators had recognised that platelet aggregation could be behind thrombotic CV events, but they continued to administer patients in the Vioxx® arm with twice the FDA-approved maximal dose.

There were more Mls in the Vioxx® arm of VIGOR than there were in the naproxen arm, though the actual numbers of patients experiencing these adverse events were not included in the paper published in the NEJM. The VIGOR paper reported only that 0.1% of patients in the naproxen arm and 0.4% of patients in the Vioxx® arm experienced an MI. These numbers could reasonably be viewed as pointing to the potential CV toxicity of Vioxx®, but Bombardier et al were not prepared to acknowledge this toxicity. The data should have been interpreted to mean that there were more Mls in the Vioxx® trial arm because Vioxx® was cardiotoxic but they interpreted the data to mean that because Vioxx® was competing with a naproxen comparator that protected the heart, Vioxx® was competing at a disadvantage.

Though the true relative risk for the published Mls was 4.25:1 Bombardier et al calculated a counterfactual relative risk of 0.2:1. In order to do this, they called upon an implausible theory developed by the Merck public relations department concerning the cardioprotective capacity of the comparator drug, naproxen. The convoluted argument supporting the investigators’ interpretation of the CV data, together with their false relative risk calculations, meant that readers (and prescribers) were led astray. A confidential internal Merck memo, MRK-NJ0362784 (20; see endnote 5), which was subpoenaed in one of the many legal actions taken against Merck, clearly shows intent to deceive.8

Publication issues with VIGOR exposed patients to harm

The VIGOR paper was published in the NEJM. Both the authors and the publishers of the VIGOR paper acted in ways which exposed patients to considerable danger. The authors did not present all the available VIGOR data to the NEJM, and the NEJM publishers failed to properly assess the data that were withheld from the NEJM. These three factors meant that doubts about the safety profile of Vioxx® were not raised in time, and patients remained exposed to the level of harms which finally led to the withdrawal of Vioxx® three years and ten months after the VIGOR paper appeared in the NEJM.

The VIGOR paper was submitted to the NEJM in May 2000, and published in November 2000. A subpoenaed internal Merck memo shows that by July 2000, Merck knew that there were three additional Mls than had been reported in the submitted paper (43). These three additional Mls meant that the relative risk of an MI for a patient on Vioxx® versus a patient on naproxen changed to 5:1, and these changes made certain other calculations in the paper incorrect. Two further data corrections were submitted to the NEJM before publication, but these did not include the three additional Mls. On October 13, 2000, one month before the publication of the VIGOR paper, Merck submitted information on the additional three Mls in the Vioxx® arm to the FDA (5,7,43), but still did not submit that information to the NEJM.

In June 2001, the editors of the NEJM received a letter drawing attention to the three additional Mls in the Vioxx® arm which had been omitted from the published VIGOR paper, and stating that information on these additional Mls were now available on the FDA website (5,7,44). The authors of the letter warned the journal that the published VIGOR results were incomplete and made the drug appear safer than it was. Their concern was that doctors prescribing Vioxx® should be made aware that the relative risk of the drug was even higher than the published figure in the Bombardier et al. 2000 VIGOR paper. The NEJM refused to publish the letter, saying space was limited (5,7,44).

In a radio programme on August 14, 2001, in which the NEJM editor-in-chief, Jeffrey Drazen was participating, one of the authors of the letter (Hrachovec) phoned in, repeating her concerns. Drazen replied that editors "can't be in the business of policing every bit of data we put out" (5,7,44,45).

Richard Smith, a former editor of the BMJ, stated that if those three additional Mls had been reported in the original publication of the VIGOR trial in November 2000, the interpretation that naproxen was protective rather than Vioxx® was harmful "would have been much less convincing – indeed, it would probably have been untenable"(5), and that if the NEJM had corrected the VIGOR data when it was informed of their existence, then the dangers of Vioxx® "might have been highlighted much earlier"(5). If the withdrawal of Vioxx® had occurred in June 2001 rather than September 2004, the trial of devastating fatal and non-fatal heart attacks and strokes caused by Vioxx® could have been halted.

Haack notes that in a Vioxx® litigation hearing in Texas in November 2005, Curfman, an editor of the NEJM, acknowledged that neither the reviewers of the VIGOR paper nor the editors of the NEJM had questioned Merck’s theory that the higher rate of CV events among Vioxx® patients was attributable to a cardioprotective effect of naproxen, even though, as an FDA official had noted, the theory “is not supported by any controlled trials”(44,45).

The NEJM and the statements of concern

In the December 29, 2005 issue of the NEJM – published 51 months after the editors received notification of the existence of the additional MI data on the FDA website; 15 months after the safety withdrawal of Vioxx®; and one month after Curfman’s statement in the Vioxx® litigation hearing above – the NEJM editors, Curfman, Morrissey and Drazen (43), recorded their dissatisfaction on finding that the VIGOR paper had not reported all the known occurrences of Mls in the Vioxx® patient group, even though authors of the paper were aware of this information before the paper was published. The statement of concern reported that the NEJM editors had found data on three additional Mls in an update on the FDA website, and also via Merck documents subpoenaed in Vioxx®
litigation. No reference was made either to the Hrachovec letter or to her phone call alerting the NEJM that there were additional MIs uploaded to the FDA website.

In the NEJM issue of March 16, 2006, there were two responses in defence of the original VIGOR data published in the November 23, 2000 issue of the NEJM. One by Bombardier et al (46) was signed by the non-Merck contributors to the original paper, and the other was signed by Reicin and Shapiro (47), both Merck employees. According to the Bombardier et al response, they were not responsible for the decision not to publish the three additional MIs in the Vioxx® arm; they were following the only pre-specified analysis plan they knew about (46). The Reicin and Shapiro response reiterated that it was acceptable to omit the three additional MIs in the Vioxx® arm because they occurred after the pre-specified cut-off date selected by Merck (47). In a further editorial response, Curfman et al (48) maintained that the difference between the later cut-off date for GI events and that used for CV events was “an untenable feature of trial design, which inevitably skewed the results”. Use of different cut-off dates for an investigational drug and the chosen comparator drug can be seen as a warning signal that there could be data that a manufacturer does not want to be available.

The initial publication of the VIGOR trial in late 2000 occurred at a time when there was mounting evidence of the fatal and non-fatal adverse CV toxicity of Vioxx®, Merck took unethical advantage of the incomplete data in the initial NEJM article (35) in the period between its publication date in November 2000 and the September 2004 safety withdrawal of rofecoxib. The lesson from the VIGOR disputes is a warning that the existence of pre-specified protocols does not necessarily endow such protocols with the mantle of scientific disinterest. The VIGOR trial was conducted with the purpose of proving the GI superiority of Vioxx®, and thereby increasing Vioxx®’s share in the very profitable COXIBS marketplace. The pre-specified protocols reflected Merck’s pre-specified interests. Paraphrasing Feinstein (1), the VIGOR RCT was designed in accordance with Merck’s policies about what questions the VIGOR trial was intended to ask, what answers were to be obtained, what was to be done with the data, and who was to be convinced by the results.

4. Publishing only favourable subgroup analyses: Merck, rofecoxib, the ADVANTAGE trial, and the APPROVe trial

The Vioxx® ADVANTAGE trial

The ADVANTAGE trial (Assessment of Differences between Vioxx and Naproxen to Ascertaining Gastrointestinal Tolerability and Effectiveness) is an example of a seeding trial. Seeding trials are used by pharmaceutical companies to promote use of a drug that has been recently approved, or is under consideration, by the FDA. These trials are framed as science, but in truth, they are marketing ploys designed to appear as if they seek an answer to a scientific question (48). The 12-week ADVANTAGE trial, with a Vioxx® arm and a naproxen arm, was actually designed by Merck’s marketing department. It was publicly identified as a safety study, but the trial was intended to promote Vioxx® to influential doctors and their patients, and the prescribing information was then to be analysed for marketing purposes with the intent of expanding sales of Vioxx® (48). The deceptions associated with the ADVANTAGE trial were only discovered after confidential internal Merck communications were subpoenaed as part of the litigation against Merck concerning harms caused by Vioxx®.

The Merck marketing department kept the intended purpose of the ADVANTAGE trial secret from institutional review boards, participating doctors and participating patients, in a comprehensive infringement of ethical practice (48). The trial commenced in March 1999, two months before the FDA approved Vioxx® for marketing in the USA in May 1999. The intention of the ADVANTAGE trial was to allow Vioxx® sales staff to gain experience with the new drug prior to and during the critical launch phase (48), and “to get physicians in the habit of prescribing a new drug” (49). A total of 5557 participants received Vioxx®, and 600 investigators prescribed it just before it became available on the market (48,49).

The ADVANTAGE trial data were submitted to the FDA in 2000, and were submitted to the Annals of Internal Medicine and published in 2003 (50). Jeffrey Lisse, a rheumatologist at the University of Arizona, was listed as the first author (50). However, as Lisse himself later reported, Merck had designed and run the trial, the initial paper was written at Merck, and then it was sent to him for editing (51). But Merck did not supply Lisse with all the data, and so the published paper did not report on the unfavourable subgroup analysis which showed “an important and statistically significant excess risk for cardiac events, namely myocardial infarction and ‘sudden/unknown’ death” (52). The Lisse et al paper was accepted for publication just one day after its submission, with just 24 hours for peer-review (12). An independent 24-hour peer-review process is extraordinary. Had Merck “bought” space in advance in the Annals?

When assessing whether a published paper on a new drug is a report of a seeding trial, it can be helpful to check the acknowledgements at the end of the paper to see whether there are a large number of investigators at a large number of sites (the ADVANTAGE trial had 600 investigators who each managed only a few of the 5557 patients in the trial). Numbers such as these point to the possibility that the trial is actually a seeding trial designed to jump-start sales of a new drug rather than answer a scientific question about it. A further factor to consider is that large numbers of trial sites with large numbers of investigators can contribute to lower rates of consistency in compliance with trial protocols, affecting the reliability of data collected. Impossibly close receipt and acceptance dates for a published paper may also help in recognising a seeding trial paper (or at the least, a trial that is associated with some unstated relationship between author and publisher that needs explaining).
The Vioxx® APPROVe trial

The APPROVe trial (Adenomatous Polyp Prevention on Vioxx), first published under the title: Cardiovascular Events Associated with Rofecoxib in a Colorectal Adenoma Chemoprevention Trial, but subsequently referred to as the APPROVe trial, was planned to be a three-year, multicentre, randomised, placebo-controlled, double-blind trial designed to evaluate the efficacy of Vioxx® in preventing the recurrence of neoplastic colorectal polyps in patients with a history of colorectal adenomas (55). The enrolment phase of the APPROVe trial at the planned 108 centres in 29 countries started in February 2000 and continued until November 2001. However, in response to the level of adverse thrombotic CV and cerebrovascular events in the APPROVe trial, Merck withdrew Vioxx® from the market for safety reasons on September 30, 2004, before any paper on APPROVe had been published.

When details of the study were first published in the NEJM in March 2005 (55), six months after the safety withdrawal of Vioxx® from the market, the investigators presented only the thrombotic CV and cerebrovascular data. The paper reported that there was an increased risk of confirmed thrombotic events associated with the long-term use of Vioxx®, but this increase was not evident in the first 18-months of the trial (55). Event rates supporting the 18-month risk-free period were calculated using Cox proportional-hazards models, and Kaplan-Meier estimates were calculated to derive cumulative event rates over time.

This 18-month finding was regarded as an important one for Merck, as even though the APPROVe paper was not published until after Vioxx® was withdrawn, the 18-month risk-free claim could be called on in litigation cases involving patients on Vioxx® who experienced thrombotic CV or cerebrovascular events while taking the drug. The 18-month “safe” period could be used to prove that adverse events which occurred before that point were not due to the use of Vioxx®.

The APPROVe trial protocol included a rule which required the censoring of data on adverse CV events when these adverse events occurred more than 14 days after a participating patient had stopped taking Vioxx®. This limitation meant that any such adverse events were not to be considered in analyses involving patients on Vioxx® who experienced thrombotic CV or cerebrovascular events while taking the drug. The 18-month “safe” period could be used to prove that adverse events which occurred before that point were not due to the use of Vioxx®.

In July 2006, the NEJM carried a correction submitted by the APPROVe investigators. In this correction, the investigators stated that they had not used the appropriate statistical procedure in the post-hoc assessment of the data from the APPROVe trial. The investigators stated that the wording in the initial 2005 NEJM publication regarding an increase in risk after 18 months should now be removed (59). The risk that Vioxx® could precipitate an adverse thrombotic CV or cerebrovascular event was acknowledged as being possible immediately upon commencing treatment with the drug.

The formal correction of the conclusion published in the initial APPROVe paper points to the potential for trial protocols and trial analysis techniques to be used to support desired conclusions and hide those that are unwanted.

Conclusion

The essential issue underpinning the clinical trials discussed in this paper is that once a decision has been made by a pharmaceutical manufacturer to conduct a trial of a new drug, a purpose for that trial exists, and this purpose represents a pre-specified interest in the results. Inter alia, this pre-specified interest determines how trial protocols are developed, what data are to be sought, how patients are chosen to maximise the opportunity to show the new drug to advantage, how to select comparators that will show adverse performances, how to avoid monitoring and reporting of any events that could reflect adversely on the new drug, and how to manage statistical analyses to show the new drug to advantage. These decisions take into account forward projections on how the performance of the new drug is to be positioned in the market, including how the new drug will be presented in medical journal publications (which, by their very number, show that manufacturers consider them an essential part of releasing a new drug). If these forward projections are best served by publishing trial data that differ from the data supplied to the FDA to support approval of the new drug, then this, too, is considered by manufacturers as acceptable practice.

Notes

2. Celebrex® has a COX-1 to COX-2 selectivity ratio of 30 (16); for Vioxx®, the ratio is 272 (16); and for Arcoxia®, the ratio is 344 (17).
3. See Appendix 1 for chronology of events discussed in this paper.
4. In his testimony given at the November 18, 2004 US Senate Finance Committee Hearing “FDA, Merck and Vioxx: Putting Patient Safety First?”, David Graham, Associate Director for Science at the FDA’s Office of Drug Safety, said that the estimate for cases of heart attack (or sudden cardiac
death) in excess of the statistical background figures for the USA ranged from 88,000 to 139,000, of whom 30%–40% probably died. The 50 mg dosage of Vioxx® proved to be the most lethal dosage (18).

5 The invention of the counterfactual and unproven property of naproxen is documented in a confidential internal document subpoenaed as part of litigation brought against Merck in relation to Vioxx®. This document, MRK-NJ0362784 (20) shows how Merck simply crossed out the data that attested to the increase in MIs (heart attacks) in the VIGOR trial and substituted wording with the opposite meaning, e.g. where the data reported an “increase,” Merck substituted the word “decrease,” and then followed through the document changing the wording to say what Merck wanted.

6 Since then, a 2013 advisory from the European Medicines Agency recommended the same CV precautions for diclofenac as for selective COX-2 inhibitors (25); a 2014 Health Canada advisory stated that diclofenac increases heart and stroke adverse events more than other NSAIDs, and compared to COX-2 inhibitors including Celebrex® (26); and a 2014 Australian Therapeutic Goods Administration Safety Review of Diclofenac advised that there is consistent evidence that there is an increased risk of serious CV events with the use of diclofenac (27).

7 It is not necessary to have a placebo arm when there is a satisfactory alternative therapy already available. RCTs with an active comparator are acceptable to the FDA. With the MEDAL program, the issue is with the choice of diclofenac as the active comparator.

8 In 2011, Merck was fined US$ 321.6 million in a criminal case in connection with its guilty plea related to its promotion and marketing of Vioxx®. In 2012, the company was fined US$ 628.3 million in a civil case concerning allegations that Merck representatives made inaccurate, unsupported, or misleading statements about the CV safety of Vioxx® in order to increase sales of the drug (42).

9 Large numbers of participants can affect statistical significance. In a 1998 review on p values and confidence intervals, Feinstein observes that “investigators who were wise enough (or fiscally supported enough) to study large groups have been able to achieve significance and to gain editorial or regulatory approval for claiming a significant action for agents that had minor importance in science or in clinical therapy (53), Wilson-Davis confirms this position noting that “it is an unfortunate fact that as the size of the sample studied becomes larger, then the smaller is the difference required to give a statistically significant difference” (54).

References


31. Large Outcomes Study Showed Merck’s ARCOXIA® (etoricoxib) had similar rates of cardiovascular thrombotic events compared to diclofenac, the most prescribed traditional NSAID worldwide, Business Wire [Internet], November 13, 2006 [cited 2017 May 20]. Available from: www.businesswire.com/news/home/20061113005495/en/Large-Outcomes-Study-Showed-Mercks-ARCOXIA-etoricoxib
43. Curfman, GD; Morrissey S, Drazen JM. Expression of concern: Bombardier...


---

**APPEAL FOR DONATIONS TO THE FORUM FOR MEDICAL ETHICS SOCIETY**

The Forum for Medical Ethics Society (FMES), publisher of the Indian Journal of Medical Ethics (IJME), appeals to subscribers, contributors, readers, supporters and well-wishers of IJME and FMES to donate to its corpus and towards the publication of IJME. Contributions from both individuals and institutions are welcome.

FMES has published the journal on a shoestring budget, without a break, since 1993. The journal is known for being open access and spearheading discourse aimed at advancing social justice in health and ethics-centred healthcare and health research reform. It is recognised widely as an authoritative resource of knowledge in bioethics and allied areas of enquiry. FMES and IJME have not only sustained themselves but have grown over these past 25 years. Our ever-growing base of readers, contributors, and reviewers enables the publication of multiple and diverse perspectives on wide-ranging topical areas. We have also striven to upgrade the journal technically to keep up with the advances in publishing to widen our reach. This has been possible only because of your steadfast and sustained support.

We are committed to maintaining our editorial independence which is critical to safeguarding the scientific integrity of what we publish. Therefore, wider support from all those who are committed to social justice in health is essential.

FMES is registered under the Societies Registration Act, 1860 (Reg No: Mumbai-218-1995-G.B.B.D.D) and the Bombay Public Trusts Act, 1950 (Reg No: F-17441 Mumbai, 1995), and donations are eligible for tax exemption under Section 80G of the Income Tax Act, 1961.

For more details on how to support us, write to fmesmumbai@gmail.com or visit our website www.ijme.in/subscribe-support/

---

**SUPPORT OPEN-ACCESS ACADEMIC PUBLICATIONS, DONATE TO THE FMES CORPUS FOR THE PUBLICATION OF IJME!**