

B. Analyses and Presentation of the APPROVe Trial Cardiovascular Data.

1. Introduction.

After Vioxx was withdrawn from the market, MRL scientists worked to present the APPROVe Trial cardiovascular data in several fora. First, MRL scientists, in collaboration with external scientists who had worked on the APPROVe Trial Administrative Committee or External Safety Monitoring Board, prepared an article about the trial's cardiovascular results that was published on the website of the New England Journal of Medicine on February 15, 2005 and in its print edition on March 17, 2005. Second, MRL scientists prepared and submitted to the FDA on June 6, 2005 an abbreviated Clinical Study Report providing the APPROVe Trial results.⁵⁶

The APPROVe article and the abbreviated Clinical Study Report stated that there was an overall increased relative risk of confirmed thrombotic events⁵⁷ in the Vioxx group of the trial as compared to the placebo group (relative risk 1.92; 95% confidence interval, 1.19 to 3.11; p=0.008) and that the relative risk of experiencing thrombotic events on Vioxx versus placebo was not constant over time but rather increased after extended exposure to Vioxx.⁵⁸ Specifically, the article stated that thrombotic event rates

⁵⁶ In addition, prior to the publication of the APPROVe article, preliminary cardiovascular data from the APPROVe Trial were presented at the October 2004 American College of Rheumatology Annual Meeting. 10/18/04 presentation, "Preliminary Cardiovascular Safety Data from the APPROVe Study," MRK-AQG0002001-32.

⁵⁷ The confirmed thrombotic event endpoint was one of the two composite endpoints of confirmed events that MRL used to evaluate the results of the APPROVe Trial. It is discussed in more detail in Section B.3.c of this Appendix.

⁵⁸ Bresalier* RS, Sandler* RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med. 2005;352:1092-102, at 092, 097, 100; see

were similar in the Vioxx and placebo groups during patients' first 18 months in the study and that the increased relative risk in the incidence of thrombotic events in the Vioxx group as compared to the placebo group became "apparent" or "evident" after 18 months of continuous treatment.⁵⁹ Merck made similar statements in connection with the February 16 to 18, 2005 meeting of the FDA's Arthritis Advisory Committee and Drug Safety and Risk Management Advisory Committee convened to discuss the safety and efficacy of selective and non-selective Cox-2 inhibitors.⁶⁰

On May 30, 2006, Merck announced that it recently had discovered that the February 2005 APPROVe article and the APPROVe Trial abbreviated Clinical Study Report described incorrectly a component of one of the statistical tests underlying the Company's conclusion that the hazard ratio – i.e., the ratio of the cardiovascular hazard

also 3/15/05 APPROVe Trial abbreviated Clinical Study Report, MRK-I8940100731, at 742 (stating that, in the APPROVe Trial, "[t]here [was] a significantly increased risk for confirmed thrombotic cardiovascular events with rofecoxib 25 mg daily compared with placebo, beginning only after 18 months of continuous treatment").

⁵⁹ Bresalier* RS, Sandler* RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med. 2005;352:1092-102, at 092, 097, 100.

⁶⁰ MRL Background Package for 2/16/05 – 2/18/05 Joint Meeting of the FDA Arthritis Advisory Committee and the Drug Safety and Risk Management Advisory Committee, MRK-AAD0408404, at 439 ("Data from the APPROVe study confirm the findings in our other clinical trial databases that there is no evidence for an increase in the relative risk of sustaining a thrombotic CV event for the rofecoxib group versus placebo over the first 18 months of treatment. However, in APPROVe, the risk of thrombotic CV events in patients taking rofecoxib 25 mg began to diverge from placebo beginning after 18 months of chronic therapy; over time the difference became significant. The mechanism(s) of the CV safety findings from APPROVe were uncertain."); see also Transcript of 2/16/05 Joint Meeting of the FDA Arthritis Advisory Committee and the Drug Safety and Risk Management Advisory Committee, at 162, MRK-AIU0185869, at 6033 (Dr. Braunstein stating: "Overall, there was an approximately two-fold increase in risk with rofecoxib compared to placebo. However, there was a statistically significant change in relative risk over time. Event rates were similar to placebo over the first approximately 18 months, consistent with our previous data. Starting after 18 months of treatment the curves began to separate with the difference becoming significant.").

rate on Vioxx to the hazard rate on placebo – was not constant throughout the course of the trial but rather increased over time. The Company announced that the error, which is discussed below, did not “change the results of the APPROVe study, in which an increased relative risk for confirmed thrombotic cardiovascular events for VIOXX compared to placebo was observed beginning after 18 months of continuous daily treatment.”⁶¹

This section discusses (i) the error generally; (ii) the drafting of the APPROVe article, including the source of the error and other issues; (iii) the drafting of the APPROVe Trial abbreviated Clinical Study Report; (iv) the Company’s additional analyses of the hazard rates in the APPROVe Trial; and (v) the Company’s discovery of and reaction to the error.

2. Overview of the Error.

a. Summary of the error.

As discussed above, MRL scientists conducted numerous statistical tests concerning the cardiovascular hazard ratio in the APPROVe Trial. The most formal of these tests – and the source of the APPROVe article error that Merck announced on May 30, 2006 – was a test of the proportionality of the hazard rates (also known as the “proportionality of hazards” test). The test of the proportionality of hazard rates presumes that the hazard ratio for the events at issue is constant over time, such that, for example, the risk of experiencing a cardiovascular event on Vioxx relative to the risk of

⁶¹ 5/30/06 Merck press release, “Merck Corrects Description of a Statistical Method Used in APPROVe Study,” MRK-ARQ0006011, at 11.

experiencing a cardiovascular event on placebo at any given point in time is roughly the same throughout the duration of the trial. The result of the test of the proportionality of the hazard rates is a p-value (a number between 0.0 and 1.0) indicating the strength of the evidence for rejecting the assumption that the hazard ratio is constant.⁶² The lower the p-value, the more evidence there is for rejecting the assumption of constant hazard ratio and concluding that the hazard ratio changes over time. Tests of proportionality of hazards, however, generally have low power, which means that they may not detect small changes in the hazard ratio. As a result, many biostatisticians consider any p-value equal to or lower than 0.10 (as opposed to the 0.05 p-value discussed throughout this Report), in conjunction with graphical analysis, sufficient to reject the assumption of proportional hazards.

While there are a variety of ways to test the proportionality of hazards in a given trial, biostatisticians commonly use a Cox proportional-hazards model fitted with one covariate for the assigned treatment (here, Vioxx or placebo) and a second covariate for either (i) the interaction between treatment and “logarithm of time” (the “Logarithm of Time Test”) or (ii) the interaction between treatment and “linear time” (the “Linear Time Test”). Whether one chooses the Logarithm of Time Test or the Linear Time Test in the first instance is a matter of personal preference, and the choice, once made, can be easily switched by plugging the other covariate into the relevant computer program.

⁶² P-values are discussed in detail in Exhibit 3 to the Report.

The February 2005 APPROVe article and the abbreviated Clinical Study Report submitted to the FDA on June 6, 2005 stated that the Logarithm of Time Test had been performed for the confirmed thrombotic event endpoint and had yielded a p-value of 0.01, which “confirmed” the “changing pattern of the treatment effect over time.”⁶³

On May 30, 2006, however, Merck announced that:⁶⁴

- The APPROVe article and the abbreviated Clinical Study Report described incorrectly the test that had been used to assess the proportionality of the hazard rates;
- The test that had been used and yielded the reported $p=0.01$ was the Linear Time Test (not the Logarithm of Time Test, as the article and the abbreviated Clinical Study Report had reported);
- “[T]he result using logarithm of time has a p-value = 0.07,” as compared to the p-value = 0.01 reported in the article and the abbreviated Clinical Study Report.

In addition, as discussed more fully below, Merck also announced that the Statistical Data Analysis Plan for Protocol 203 – a pre-planned combined analysis of thrombotic events from the APPROVe, VICTOR and ViP Trials (discussed in detail in Appendix M) intended to serve as a cardiovascular outcomes trial – “called for numerous statistical and graphical methods to be used to assess whether the relative risk of VIOXX compared to placebo was constant over time or if it changed over time” and that “[t]he use of the variable, logarithm of time, was an element in the primary method specified [in that data

⁶³ Bresalier* RS, Sandler* RS, Quan H, *et al.* Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. *N Engl J Med.* 2005;352:1092-102, at 97, MRK-ARQ0000659; see 3/15/05 APPROVe Trial abbreviated Clinical Study Report, Cardiovascular Safety Report, MRK-I8940100962, at 976, 986.

⁶⁴ 5/30/06 Merck press release, “Merck Corrects Description of a Statistical Method Used in APPROVe Study: Study Results Unchanged,” MRK-ARQ006011, at 11.

analysis plan for Protocol 203].”⁶⁵ Finally, the press release stated that “[t]he results of diagnostic steps specified in the data analysis plan [for Protocol 203] indicate[d] that [the Linear Time Test] is an appropriate method to assess changes in relative risk over time.”⁶⁶

b. The statistical test at issue and relevant data analysis plans.

The error in the APPROVe article raised questions about whether the precise method to be used to test proportionality of the hazard ratio was pre-specified. As discussed in Appendix A, prior to unblinding the data at issue, scientists often pre-specify the statistical analyses to be performed on the unblinded data to avoid bias in the way the data are analyzed or the conclusions are drawn. In general, the results of pre-specified analyses drive a study’s conclusions and must be reported.

The APPROVe article stated that the proportionality of hazards assumption was tested using the Logarithm of Time Test. The article also stated that this test “was specified in the cardiovascular-analysis plan.”⁶⁷ The article did not define the “cardiovascular-analysis plan” or indicate whether the “cardiovascular-analysis plan” was part of the Data Analysis Plan for the APPROVe Trial.

⁶⁵ 5/30/06 Merck press release, “Merck Corrects Description of a Statistical Method Used in APPROVe Study: Study Results Unchanged,” MRK-ARQ006011, at 11.

⁶⁶ 5/30/06 Merck press release, “Merck Corrects Description of a Statistical Method Used in APPROVe Study: Study Results Unchanged,” MRK-ARQ006011, at 11.

⁶⁷ Bresalier* RS, Sandler* RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. *N Engl J Med.* 2005;352:1092-102, at 1095, MRK-ARQ0000659, at 62. Section B.3.g of this Appendix discusses the quoted language and its origin in detail.

The Data Analysis Plan for the APPROVe Trial pre-specified in detail the statistical methods for assessing efficacy of Vioxx in preventing the recurrence of colon polyps (the primary endpoint of the trial), but pre-specified only in general terms the statistical methods for analyzing other adverse event data, including cardiovascular event data. For example, it called for the calculation of the between-group difference and associated 95% confidence interval for confirmed thrombotic events,⁶⁸ but did not pre-specify any test of the proportional hazards assumption for this endpoint. It also stated that the APPROVe Trial confirmed thrombotic event data would be “combined with those of other studies in a pooled analysis [*i.e.*, Protocol 203] to better assess the effect of rofecoxib on these [confirmed thrombotic events].”⁶⁹

As discussed in detail in Section B.3.g of this Appendix, references in the APPROVe article to the “cardiovascular-analysis plan” were in fact references to the Statistical Data Analysis Plan for Protocol 203 – a combined analysis of thrombotic events from the APPROVe, VICTOR and ViP Trials (discussed in detail in Appendix M). The Statistical Data Analysis Plan for Protocol 203 pre-specified the methods for testing the proportionality of hazards for confirmed thrombotic events and the applicable level of statistical significance, stating:

Analytical and graphical methods will be employed to verify the proportional hazards assumption. The primary method for testing the proportional hazards assumption will be by including the factor $\text{treatment} * \log(\text{time})$ in the model

⁶⁸ 3/8/03 Protocol 122 Statistical Data Analysis Plan, MRK-AGO0025614, at 631.

⁶⁹ 6/11/02 Protocol 122 Statistical Data Analysis Plan, MRK-AFL0000907, at 25.

[i.e., the Logarithm of Time Test]; nonsignificance ($p > 0.050$) of this factor is not inconsistent with proportionality, i.e., constancy of treatment effect over time.

* * *

Further, the log HR [hazard ratio] will be plotted over time by stratifying time into intervals containing approximately the same number of events within each interval. The log HR [hazard ratio] within each successive 6-month time interval with confidence limits will also be calculated and plotted. Such plots will provide an indication of any time effect on the HR [hazard ratio].⁷⁰

Because Protocol 203 was a study “to assess cardiovascular safety from the combined analysis of 3 placebo-controlled studies of rofecoxib” – the APPROVe, VICTOR and ViP Trials,⁷¹ the Statistical Data Analysis Plan for Protocol 203 did not govern the analyses of the APPROVe Trial cardiovascular data standing alone. Had the Statistical Data Analysis Plan for Protocol 203 governed the analysis of the data from the APPROVe Trial alone, the $p=0.07$ obtained for the Logarithm of Time Test for the confirmed thrombotic event endpoint would not have met the pre-specified $p=0.05$ statistical significance level, meaning one could not conclude from this test that the hazard ratio of confirmed thrombotic events in the APPROVe Trial changed over time.

As discussed above, in its May 30, 2006 press release, Merck stated that the Statistical Data Analysis Plan for Protocol 203 “called for numerous statistical and graphical methods to be used to assess whether the relative risk of VIOXX compared to

⁷⁰ Protocol 203 Statistical Data Analysis Plan, MRK-AAB0083231, at 55-56.

⁷¹ Protocol 203 Statistical Data Analysis Plan, MRK-AAB0083231, at 39.

placebo was constant over time or if it changed over time.”⁷² The press release also stated that “[t]he use of the variable, logarithm of time, was an element in the primary method specified [in that Statistical Data Analysis Plan for Protocol 203].”⁷³ Finally, the press release stated that “[t]he results of diagnostic steps specified in the data analysis plan [for Protocol 203] indicate[d] that [the Linear Time Test] is an appropriate method to assess changes in relative risk over time.”⁷⁴

3. The APPROVe Article.

a. Introduction.

The cardiovascular results from the APPROVe Trial were published in the New England Journal of Medicine on February 15, 2005.⁷⁵ The article, entitled “Cardiovascular Events Associated with Rofecoxib in a Colorectal Adenoma Chemoprevention Trial,” was co-authored by members of the APPROVe Trial Administrative Committee, which included six external and three MRL scientists, along with Dr. Marvin Konstam*, a member of the APPROVe Trial External Safety Monitoring

⁷² 5/30/06 Merck press release, “Merck Corrects Description of a Statistical Method Used in APPROVe Study: Study Results Unchanged,” MRK-ARQ006011, at 11.

⁷³ 5/30/06 Merck press release, “Merck Corrects Description of a Statistical Method Used in APPROVe Study: Study Results Unchanged,” MRK-ARQ006011, at 11.

⁷⁴ 5/30/06 Merck press release, “Merck Corrects Description of a Statistical Method Used in APPROVe Study: Study Results Unchanged,” MRK-ARQ006011, at 11.

⁷⁵ The article was also published in the print edition of the New England Journal of Medicine in March 2005. Bresalier* RS, Sandler* RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med. 2005;352:1092-102, MRK-ARQ0000659.

Board, Dr. Hui Quan, the unblinded statistician for the APPROVe Trial, and

Dr. Christopher Lines, a medical writer from MRL.⁷⁶

The co-authors began drafting the manuscript in early November 2004,⁷⁷ with the intention of submitting it to the New England Journal of Medicine.⁷⁸ Dr. Baron* submitted the manuscript to Dr. Jeffrey Drazen*, Editor-in-Chief of the New England

⁷⁶ Bresalier* RS, Sandler* RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med. 2005;352:1092-102, MRK-ARQ0000659. The first and last authors were Dr. Robert S. Bresalier*, Department of Gastrointestinal Medicine and Nutrition, University of Texas M.D. Anderson Cancer Center, and Dr. John A. Baron* (Chair of the Administrative Committee), Departments of Medicine and Community and Family Medicine, Dartmouth Medical School, respectively. Other authors were: Mr. James Bolognese, MRL, the blinded statistician for the APPROVe Trial, member of the APPROVe Trial's Administrative Committee; Dr. Christopher Lines, Director, Medical Communications; Dr. Quan, MRL; Dr. Bettina Oxenius, MRL, the Clinical Monitor on the APPROVe Trial and member of the Administrative Committee; Dr. Kevin Horgan, MRL, member of the Administrative Committee and Dr. Oxenius's supervisor; Dr. Robert S. Sandler*, Department of Medicine, University of North Carolina at Chapel Hill, member of the APPROVe Trial's Administrative Committee; Dr. Marvin A. Konstam*, a cardiologist from the Department of Medicine, Tufts New England Medical Center and a member of the APPROVe Trial's External Safety Monitoring Board; Dr. Robert Riddell*, Department of Pathology, Mount Sinai Hospital, Toronto, Canada, member of the APPROVe Trial's Administrative Committee, Dr. Dion Morton*, Department of Surgery, University of Birmingham, United Kingdom, member of the APPROVe Trial's Administrative Committee; and Dr. Angel Lanas*, Department of Medicine, Clinic University Hospital, Zaragoza, Spain, member of the APPROVe Trial's Administrative Committee.

⁷⁷ 11/2/04 email from C. Lines to J. Baron* et al., MRK-AHC0027631 ("I'm attaching a draft outline of the above paper, for discussion at our teleconference on Thursday.").

⁷⁸ APPROVe article draft, MRK-AHC0027633, at 633 (referring to the word limit for articles submitted to the New England Journal of Medicine) (attached to 11/14/04 email from R. Sandler* to C. Lines, MRK-AHC0027631); see also 10/6/04 email from J. Baron* to J. Drazen*, MRK-AHC0008096 (excerpting Dr. Drazen's* email, which had stated: "It would be good to get the 'safety' article in the Journal during this calendar year.").

The Journal of the American Medical Association also had expressed to Dr. Baron* interest in considering the manuscript for publication. 11/23/04 email from J. Baron* to P. Fontanarosa*, MRK-AHC0028559 (responding to Dr. Fontanarosa's* email, which had stated: "I am writing with the hope that you might be interested in submitting the manuscript from the APPROVe study to JAMA, for consideration for rapid peer review and EXPRESS publication.").

Journal of Medicine, on February 6, 2005.⁷⁹ By letter dated February 9, 2005, Dr. Gregory Curfman*, Executive Editor of the New England Journal of Medicine, informed the authors that the manuscript “ha[d] not been accepted for publication in its current form,” but that the New England Journal of Medicine “would be very willing to consider an extensively revised version of the manuscript that successfully addressed the [attached] comments of the [peer] reviewers” (discussed in Section B.3.g of this Appendix).⁸⁰

During the following two days, the authors revised the manuscript. The article was posted on the journal’s website on February 15, 2005 – one day before the FDA’s Arthritis Advisory Committee and Drug Safety and Risk Management Advisory Committee convened on February 16 to 18, 2005 to discuss the safety and efficacy of selective and non-selective Cox-2 inhibitors, including the results of the APPROVe Trial – and was published in the March 17, 2005 print edition of the journal.⁸¹

The following subsections discuss (i) the respective roles of the various co-authors in the drafting process; (ii) the source of the erroneous description of the method used to test the proportionality of hazards; (iii) discussions among the co-authors about other notable aspects of the article prior to its initial submission; (iv) the draft

⁷⁹ 2/6/05 email from J. Baron* to R. Sandler* et al., MRK-AFV0398907, at 07.

⁸⁰ 2/9/05 letter from G. Curfman* to J. Baron*, MRK-AFV0399548.

⁸¹ Bresalier* RS, Sandler* RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med. 2005;352:1092-102, MRK-ARQ0000659.

manuscript submitted to the New England Journal of Medicine on February 6, 2005;

(v) the peer reviewers' comments and resulting revisions; and (vi) the published article.

b. Participants in the drafting process.

The initial working group drafting and reviewing the APPROVe article consisted of Dr. Baron*, Dr. Bresalier*, Mr. Bolognese, Dr. Horgan, Dr. Lines, and Dr. Sandler*.⁸² In early November 2004, Dr. Lines circulated several draft outlines of the manuscript that he had written, and Mr. Bolognese, Dr. Baron* and Dr. Sandler* commented on and proposed revisions to the draft outlines.⁸³ In addition, Dr. Quan contributed the results of analyses based on the August 2004 data used in his September 13, 2004 report, which served as a placeholder until the final data became available in December 2004.⁸⁴

After the MRL co-authors had finalized a preliminary draft of the "Introduction," "Methods," and "Results" sections of the manuscript (taking into account comments from the external co-authors), Dr. Bresalier*, with input from Drs. Baron* and Sandler*,

⁸² 10/27/04 email from K. Horgan to J. Baron*, MRK-AFO0267117 ("The working group for preparing the publication will consist of me, Chris Lines, Jim Bolognese and you, Bob Sandler and Bob Bressalier [sic] Chris and I will provide an outline of the data to be included in the paper prior to the teleconference."); see also 10/27/04 email from K. Horgan to D. Quinn, MRK-AFO0267118 (referring to a teleconference among Drs. Horgan, Lines, Baron*, Bresalier*, and Sandler* and Mr. Bolognese).

⁸³ 11/2/04 email from C. Lines to J. Baron* et al., MRK-AHC0027631 (attaching "a draft outline" of the APPROVe article); 11/2/04 email from J. Bolognese to C. Lines, MRK-AGO0029753 ("some comments are included in the attachment"); 11/14/04 email from R. Sandler* to C. Lines, K. Horgan, MRK-AHC0027631 (attaching "a marked up manuscript with comments"); 11/15/04 email from J. Baron* to K. Horgan et al., MRK-AHC0024449 (attaching comments on Dr. Lines' draft outline).

⁸⁴ 11/18/04 email from H. Quan to C. Lines and J. Bolognese, MRK-AGO0075099 ("I have added the data we have so far").

prepared a draft “Discussion” section of the manuscript.⁸⁵ By early December 2004, Dr. Konstam* also had provided comments on the draft of the manuscript.⁸⁶ In December 2004 and January 2005, Dr. Bresalier* continued to receive proposed revisions from his co-authors and to incorporate the revisions with which he agreed.⁸⁷ Prior to submission, several non-author MRL scientists (Drs. Gertz, Reicin, Braunstein, van Adelsberg, and Hirsch) also provided comments and signed off on the draft before it was submitted on February 6, 2005.⁸⁸ As with comments from the article co-authors, Dr. Bresalier* considered all revisions proposed by MRL reviewers and incorporated in the draft the revisions with which he agreed.⁸⁹

⁸⁵ E.g., 11/18/04 from C. Lines to J. Baron* et al., MRK-AHC0024585 (“we’re now handing over to you to work on the Discussion”); 11/29/04 email from K. Horgan to A. Reicin, MRK-AHC0021967 (stating that the manuscript is “awaiting draft of discussion from lead author Bresalier”); 12/1/04 email from J. Baron* to R. Bresalier* et al., MRK-AHC0024621 (“I’ve marked up [Dr. Bresalier’s*] revision of [Dr. Lines’] draft and also his discussion.”); 12/5/04 email from R. Sandler* to J. Baron* et al., MRK-AHC0030140 (attaching revised version of the draft).

⁸⁶ 12/6/04 email from R. Bresalier* to C. Lines, MRK-AHC0031383 (“Marv Konstam had several substantive comments about how he would like to see the data presented.”).

⁸⁷ See, e.g., 1/5/05 email from D. Quinn to F. Jeremiah*, MRK-AGO0063797 (referring to 1/11/05 teleconference among certain co-authors); 1/10/05 email from R. Bresalier* to K. Horgan, MRK-AGO0075320 (attaching a revised draft of the manuscript); 1/10/05 email from J. Bolognese to R. Bresalier* et al., MRK-AHD0000760 (attaching a draft of the manuscript with “a few more suggested edits”).

⁸⁸ 1/13/05 APPROVe article draft with J. van Adelsberg’s comments, MRK-AHD0000459; 1/17/05 email from L. Hirsch to C. Lines, MRK-AHD0074962 (attaching draft of the APPROVe article with his comments); APPROVe article draft with A. Reicin’s comments, MRK-AHD0000429; 12/07/04 APPROVe article draft with B. Gertz’s comments, MRK-AHD0000843 (attached to 12/13/04 email from B. Gertz to C. Inoa, MRK-AHD0000842).

⁸⁹ 12/18/04 email from R. Bresalier* to C. Lines, MRK-AHD0073546 (“I have revised [the manuscript] again taking into account many of the comments suggested by Alise [Reicin] and others. I have not incorporated all suggestions verbatim since this must be presented as an objective scientific article and frankly, a few of the comments appeared aimed at ‘damage control’.”).

c. Cardiovascular event endpoints discussed in the APPROVe article.

The APPROVe article discussed analyses of three cardiovascular thrombotic event endpoints (with emphasis on the first):

- Confirmed thrombotic event endpoint – a set of thromboembolic events pre-defined by MRL and diagnosed (or “confirmed”) by a blinded expert adjudication committee pursuant to MRL’s Cardiovascular Adjudication SOP;
- APTC composite cardiovascular events endpoint – a set of cardiovascular events comprising the endpoint developed by the Antiplatelet Trialists’ Collaboration and diagnosed (or “confirmed”) by a blinded expert adjudication committee pursuant to MRL’s Cardiovascular Adjudication SOP; and
- Investigator-reported cardiovascular events – a pre-defined set of thromboembolic events diagnosed by investigators (*i.e.*, doctors dispensing the study drug) whether or not such diagnoses were subsequently confirmed by the blinded expert adjudication committee pursuant to MRL’s Cardiovascular Adjudication SOP.

As demonstrated in Table 1, the composition of the two confirmed event endpoints overlapped, but the endpoints differed in that (i) the confirmed thrombotic event endpoint included certain peripheral vascular events, such as pulmonary embolism, that the APTC composite cardiovascular endpoint did not include; and (ii) the APTC composite cardiovascular endpoint included three hemorrhagic adverse events, such as hemorrhagic cerebrovascular stroke, that the confirmed thrombotic endpoint did not include.⁹⁰

⁹⁰ For more information regarding the two composite endpoints, see Appendix F.

d. The source of the error discovered in May 2006.

As mentioned above, the APPROVe article stated that the test of proportionality of hazards for the confirmed thrombotic event endpoint yielded a p-value of 0.01 and described the method used to perform the test as the Logarithm of Time Test.⁹¹ In May 2006, the Company discovered that the description of the test was wrong because the test that had been used to test proportionality of hazards was the Linear Time Test. This subsection discusses the circumstances that led to the error.

i. Dr. Quan's September 13, 2004 safety report to the APPROVe Trial External Safety Monitoring Board.

When the co-authors began writing the article, they used as a placeholder the data presented in Dr. Quan's September 13, 2004 report to the APPROVe Trial External Safety Monitoring Board (which included events reported and adjudicated by August 16, 2004).⁹² In preparing the September 13, 2004 report, Dr. Quan had, at the External Safety Monitoring Board's request,⁹³ performed three sets of analyses (described in Section A.2.a of this Appendix) aimed at determining whether the risk of cardiovascular events on Vioxx relative to placebo was increasing over time – such that,

⁹¹ Bresalier* RS, Sandler* RS, Quan H, et al. Cardiovascular Events Associated with Rofecoxib in a Colorectal Adenoma Chemoprevention Trial. N Engl J Med. 2005;352:1092-102, at 95, 98; MRK-ARQ0000659, at 62, 65.

⁹² 9/13/04 Pre-Meeting Report from H. Quan to APPROVe ESMB, MRK-AGO0029517, at 517.

⁹³ 8/28/04 email from J. Neaton* to H. Quan, MRK-AGO0000326.

for example, a two-fold increased risk in year 1 might grow into a four-fold increased risk in year 2 – or whether the hazard ratio remained constant over time.⁹⁴

One of these analyses was a statistical test to assess whether the hazard rates in the placebo and Vioxx groups were proportional (i.e., whether the hazard ratio stayed constant) throughout the trial. As mentioned above, biostatisticians commonly perform this test by analyzing the interaction between the assigned treatment (here, Vioxx and placebo) and either the logarithm of time or linear time in the Cox proportional-hazards model (i.e., by performing the Logarithm of Time Test or the Linear Time Test). There is no contemporaneous documentary evidence of the analyses of the proportionality of hazard rates that Dr. Quan conducted, but, after being informed of the error, Dr. Quan recalled that, as he was preparing in the summer of 2004 for the September 2004 meeting of the APPROVe Trial External Safety Monitoring Board, his analyses proceeded as follows:

⁹⁴ 9/13/04 Pre-Meeting Report from H. Quan to APPROVe ESMB, MRK-AGO0029517.

The External Safety Monitoring Board's interest in these analyses was triggered by analyses it had reviewed previously. First, at the Board's February 18, 2004 meeting, Dr. Reicin had informed the Board that, in the Alzheimer's disease trials comparing Vioxx 25 mg to placebo (Protocols 078 and 091), the constant hazard ratio assumption did not hold for the APTC and the confirmed thrombotic endpoints, suggesting that the hazard ratio for cardiovascular events on Vioxx as compared to placebo may have changed over time. Minutes of 2/18/04 APPROVe Open Session ESMB meeting, MRK-AFF0000120, at 21. Dr. Reicin's presentation to the External Safety Monitoring Board is discussed in Appendix Q. Second, data contained in Dr. Quan's February 12, 2004 safety report to the APPROVe Trial External Safety Monitoring Board suggested that the hazard ratio for cardiovascular events on Vioxx as compared to placebo might be increasing over time in the APPROVe Trial as well. 2/12/04 Pre-Meeting Report from H. Quan to APPROVe ESMB, MRK-AGO0006866, at 884-85 (indicating that the proportionality p-values were 0.0545 and 0.1312 for the confirmed thrombotic event and the APTC event endpoints, respectively). For detailed discussion of proportionality p-values, see Section B.2 of this Appendix.

- The computer program for the test of the proportionality of hazards, as originally written, used the Logarithm of Time Test (consistent with Dr. Quan's usual practice and MRL's convention),⁹⁵
- Dr. Quan ran the Logarithm of Time Test on the data available to him at the time, which yielded a p-value above 0.05 (meaning that, based on the $p \leq 0.05$ standard that Dr. Quan generally used, there was not sufficient evidence to conclude that the hazard ratio changed over time).⁹⁶ This result struck Dr. Quan as inconsistent with other results he had obtained, most notably the Kaplan-Meier plot, which seemed clearly to indicate a change in the hazard ratio;
- Dr. Quan stated that, if the hazard ratio in the APPROVe Trial changed over time, he wanted the External Safety Monitoring Board to know about it;
- Dr. Quan then ran the analysis using the Linear Time Test. The Linear Time Test generated a p-value that was lower than that obtained for the Logarithm of Time Test and therefore more consistent with Dr. Quan's prior analyses (such as the Kaplan-Meier plots). Based on his review of the two p-values, as well as "fit statistics" generated as part of the output from the two tests, Dr. Quan concluded that the Linear Time Test covariates fit the data better than the Logarithm of Time Test covariates;⁹⁷

⁹⁵ The first pre-meeting report to the APPROVe Trial External Safety Monitoring Board to cite proportionality p-values was the November 18, 2003 report; therefore, the computer program at issue was probably written before November 18, 2003. See 1/21/02 Pre-Meeting Safety Report from H. Quan to APPROVe ESMB, MRK-AGO0006753; 2/11/02 Pre-Meeting Report from H. Quan to APPROVe ESMB, MRK-AGO0006667; 5/9/02 Pre-Meeting Report from H. Quan to APPROVe ESMB, MRK-AFF0000318; 5/8/03 Pre-Meeting Report from H. Quan to APPROVe ESMB, MRK-AFK0202981; 11/18/03 Pre-Meeting Report from H. Quan to APPROVe ESMB, MRK-AGO0007185, at 202-203 (citing proportionality p-values); 2/12/04 Pre-Meeting Report from H. Quan to APPROVe ESMB, MRK-AGO0006866, at 884-85 (citing proportionality p-values); 9/13/04 Pre-Meeting Report from H. Quan to APPROVe ESMB, MRK-AGO0029517, at 536-37 (citing proportionality p-values).

⁹⁶ Dr. Quan stated that the dataset for which he obtained the non-significant result was smaller than the August 2004 dataset that was analyzed in the September 13, 2004 report to the APPROVe Trial External Safety Monitoring Board. The Logarithm of Time Test result for the August 2004 dataset is statistically significant ($p=0.048$). Draft letter to the editor, MRK-AQU0000031, at 32 (referring to the $p=0.048$ result for the August 2004 APPROVe Trial data) (attached to 6/8/06 email from J. Bolognese to T. Reiss, MRK-AQU0000030).

⁹⁷ 6/8/06 email from H. Quan to J. Bolognese et al., MRK-ARQ0007271.

- Accordingly, Dr. Quan modified the computer program to use the Linear Time Test – a test that Dr. Quan believed would give the External Safety Monitoring Board a more accurate assessment of what was happening in the APPROVe Trial because, in his view, “[i]f a model does not fit the data very well, the use of that model for checking the proportionality [of hazard rates] assumption will not serve that purpose.”⁹⁸

In preparing the September 13, 2004 report to the APPROVe Trial External Safety Monitoring Board, Dr. Quan used the same program he modified previously that summer. As a result, the report included results of the Linear Time Test for the confirmed thrombotic and APTC composite cardiovascular event endpoints. The report, did not, however, specify which covariate (linear time versus logarithm of time versus something else) was used in producing these results. The results were $p=0.0056$ and $p=0.0235$ for the confirmed thrombotic and APTC composite cardiovascular event endpoints, respectively.⁹⁹ Both p-values were low, suggesting that the cardiovascular hazard ratios for both endpoints changed over time.

On June 8, 2006, Dr. Quan wrote an email concerning the error in the APPROVe article to Mr. Bolognese and other co-authors in which he stated that he had not documented the change in the variable (other than in the program itself) and subsequently had forgotten switching the variable in the program. As a result, from September 2004 until May 2006 (including when he worked on the APPROVe article and the Clinical Study Report for the APPROVe Trial), Dr. Quan mistakenly believed that the results included in the September 13, 2004 report and the results included in the APPROVe

⁹⁸ 6/8/06 email from H. Quan to J. Bolognese et al., MRK-ARQ0007271.

⁹⁹ 9/13/04 Pre-Meeting Report from H. Quan to APPROVe ESMB, MRK-AGO0029517, at 536-37.

article – which were produced using the same computer program – were based on the Logarithm of Time Test.¹⁰⁰

ii. Final (December 2004) data.

In December 2004, MRL scientists obtained final data for the APPROVe Trial and re-ran many of the analyses that had been performed on the preliminary cardiovascular data. The APPROVe Trial abbreviated Clinical Study Report and the APPROVe article reflected these final data analyses. Among the analyses performed on the final data was a test of the proportionality of hazards using the same computer program Dr. Quan had used for his September 13, 2004 report (i.e., the program that he had modified to use the Linear Time Test). The final data yielded proportionality p-values of 0.014 and 0.119 for the confirmed thrombotic and APTC composite cardiovascular event endpoints, respectively.¹⁰¹ Table 11 below summarizes the proportionality test results for the confirmed thrombotic and APTC composite cardiovascular event endpoints based on preliminary and final data.

¹⁰⁰ 6/8/06 email from H. Quan to J. Bolognese et al., MRK- ARQ0007271 (“Unfortunately, since I could not modify the APPROVe DAP and didn’t document this in some document, I forgot all about this when I rushed through the final analysis. Thus, in my mind, I always thought we used [logarithm of time].”).

¹⁰¹ 3/15/05 APPROVe Trial abbreviated Clinical Study Report, Cardiovascular Safety Report, MRK-I8940100731, at 0986, 0994.

Table 11

Cardiovascular Thromboembolic Event Proportionality of
Hazards P-Values in the APPROVe Trial

	Preliminary (August 2004) Data	Final (December 2004) Data
Confirmed Thrombotic Event Endpoint	p=0.0056	p=0.014
APTC Composite Cardiovascular Event Endpoint	p=0.0235	p=0.119

- iii. Incorrect description of the test of proportionality
of the hazard rates in the APPROVe article.

As noted above, the published APPROVe article reported the results of the Linear Time Test performed on the final (December 2004) data (p=0.01) but described that p-value as resulting from the Logarithm of Time Test. The incorrect description that appeared in the published article was not added to the manuscript until February 13, 2005 – two days before the article was published. This subsection and Table 12 below summarize the descriptions of the proportionality of hazards test in the preceding drafts of the manuscript and the circumstances under which the incorrect description was added to the article on February 13, 2005.

The first draft to cite the results of the test for the proportionality of hazards, which was circulated by Dr. Lines on November 16, 2004, reported those results for the confirmed thrombotic event endpoint based on the preliminary data (p=0.006), but did

not describe the method that had been used to perform the test.¹⁰² On January 11, 2005, after Dr. Konstam* questioned whether the test used to obtain the $p=0.006$ result was described in the “Methods” section of the manuscript,¹⁰³ the authors supplemented the section by inserting the following sentence: “The proportional hazards assumption was tested by evaluation of the time-by-log (time) interaction in the COX model.”¹⁰⁴ This new text did not accurately describe the Logarithm of Time Test that Dr. Quan believed he had performed (there is no test that can be described as “time-by-log (time) interaction” test), nor did it describe the Linear Time Test that Dr. Quan had in fact performed. Subsequent drafts, including the one that Dr. Baron* submitted to Dr. Drazen* of the New England Journal of Medicine on February 6, 2005, continued to include this description, although the p-value was changed to $p=0.01$ on January 13, 2005 after the final data became available.¹⁰⁵

At some point between February 6, 2005 and February 12, 2005, the description was revised to state that the proportionality of hazards was tested “by evaluating the interaction between the time to an event and the total patient-years of follow-up in the

¹⁰² 11/15/04 APPROVe article draft, MRK-AHD0072653, at 663 (attached to 11/16/04 email from C. Lines to K. Horgan, MRK-AHD0072652) (“The changing pattern of treatment effect over time was confirmed by the failed test for proportionality of hazards ($p=0.006$).”).

¹⁰³ 11/18/04 APPROVe article draft with M. Konstam’s* revisions, MRK-AHC0024647, at 659 (attached to 12/1/04 email from J. Baron* to R. Sandler* et al., MRK-AHC0024646).

¹⁰⁴ 1/11/05 APPROVe article draft, MRK-AGO0075451, at 457 (attached to 1/11/05 email from R. Bresalier* to C. Lines et al., MRK-AGO0075450) (“Revised manuscript. I have taken into account all changes discussed this morning.”).

¹⁰⁵ APPROVe article draft, MRK-AGO0075953, at 962 (attached to 1/13/05 email from R. Bresalier* to C. Lines et al., MRK-AGO0075952).

Cox proportional-hazards model.”¹⁰⁶ (This description, like the one before it, did not refer to the Linear Time Test Dr. Quan had performed, or to the Logarithm of Time Test Dr. Quan believed he had performed.)

On February 12, 2005, a Journal editor asked the authors whether the description in the February 12, 2005 draft correctly defined the interaction term used in the Cox proportional hazards model.¹⁰⁷ As a result, the description was changed to that which appeared in the published article: “This [test] was accomplished by evaluating the interaction between the logarithm of time and the assigned treatment in the Cox proportional-hazards model.”¹⁰⁸ Both Mr. Bolognese and Dr. Quan believed that one of them made that change, but neither recalled which one. The following table reflects the evolution of the description of the statistical method used in various drafts of the APPROVe manuscript:

¹⁰⁶ APPROVe article draft, MRK-AHD0092254, at 60 (attached to 2/12/05 email from C. Lines to J. Wainwright et al., MRK-AHD0092252-53).

¹⁰⁷ APPROVe article draft, MRK-AHD0092254, at 60 (attached to 2/12/05 email from C. Lines to J. Wainwright et al., MRK-AHD0092252-53).

¹⁰⁸ Bresalier* RS, Sandler* RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med. 2005;352:1092-102, at 95, MRK-ARQ0000659, at 662.

Table 12

APPROVe Article Drafts – Description of the Test for Proportionality of the Hazard Rates

Date of Draft(s)	Proportionality of the Hazard Rate Result Cited	Description of the Method
11/2/04 - 11/18/04	Result not included	Method not described
12/01/04 - 1/11/05	p=0.006 (based on preliminary August 2004 data)	Method not described
1/11/05 - 1/13/05	p=0.006 (based on preliminary August 2004 data)	“The proportional hazards assumption was tested by evaluation of the time-by-log (time) interaction in the COX model.”
1/13/05 - 2/6/05 (including the draft submitted to the <u>New England Journal of Medicine</u>)	p=0.01 (based on final December 2004 data)	“The proportional hazards assumption was tested by evaluation of the time-by-log (time) interaction in the COX model.”
2/13/05	p=0.01	“A test of the proportional-hazards assumption was . . . accomplished by evaluating the interaction between the time to an event and the total patient years of follow-up in the Cox proportional hazards model.”
Published article (2/15/05)	p=0.01	“A test of the proportional-hazards assumption was . . . accomplished by evaluating the interaction between the logarithm of time and the assigned treatment in the Cox proportional-hazards model.”

e. Issues discussed during drafting process.

As discussed above, in the fall and winter of 2004, MRL scientists conducted several post-hoc analyses (i.e., analyses not pre-specified in the Data Analysis Plan for the APPROVe Trial) to explore whether and how the incidence of thrombotic events in the Vioxx and placebo groups of the APPROVe Trial varied over time. In addition, the MRL and external scientists noticed that there were very few events in the placebo group of the APPROVe Trial after the 18-month point (as depicted by the relatively flat course

of the placebo curve on the Kaplan-Meier plot), which some viewed as potentially contributing to the difference in the incidence rates on Vioxx versus placebo in the second 18 months.

In the course of drafting the APPROVe article, the authors discussed whether and to what extent to refer to – and what conclusions, if any, to draw from – these post-hoc analyses and the observed “flattening” of the placebo curve in the second 18 months. The dialogue among the APPROVe article co-authors and MRL scientists concerning whether to include these analyses in the manuscript is discussed below.

i. Discussion regarding post-18-month increase in relative risk.

In writing the APPROVe article, the authors discussed the extent to which the article should focus on the results of the post-hoc analyses (discussed above) showing that the incidence of cardiovascular events was similar in the Vioxx and placebo groups for the first 18 months of the APPROVe Trial and diverged thereafter. On November 2, 2004, Mr. Bolognese sent an email to Drs. Horgan, Lines, and Quan attaching a draft of the article with what Mr. Bolognese characterized as “some major comments.”¹⁰⁹ In his comments, Mr. Bolognese expressed the view that “the key point of the results” was that “the results for the first 18 months did not show an increased relative risk of confirmed cardiovascular events” and that “the increased relative risk was

¹⁰⁹ 11/2/04 email from J. Bolognese to C. Lines *et al.*, MRK AGO0029732 (attached to 11/1/04 APPROVe article draft with J. Bolognese’s revisions, MRK-AGO0029733).

observed beginning after 18 months of treatment.”¹¹⁰ Mr. Bolognese further stated that this “key point . . . should lead the results section [and] be followed with appropriate summary statistics broken by <18 [months] > 18 [months].”¹¹¹ In addition, Mr. Bolognese suggested that the article “maybe even NOT cite the 0-36 [month] results at all” and that “[i]f we must present the combined 0-36 [month] info, I suggest we immediately discount it because of non-proportional hazards . . . and follow it up [with relative risk analyses for the first and second 18-month intervals].”¹¹²

Later that day, Mr. Bolognese sent a follow-up email to the same group stating that a discussion with Dr. Quan had led him to modify some of his comments. In his revised comments, Mr. Bolognese stated that “[t]he thrust of the interpretation” should be that (i) the “0-36 month results represent the overall average difference across the time period,” and (ii) “[t]his average is made up of 2 distinctly different profiles, [the] <18 [months] and >18 [months].”¹¹³ Mr. Bolognese suggested that the article should report the rate and relative risk analyses for the first and second 18-month segments of

¹¹⁰ 11/1/04 APPROVe article draft with J. Bolognese’s revisions, MRK-AGO0029733, at 735 (attached to 11/2/04 email from J. Bolognese to C. Lines et al., MRK-AGO0029732).

¹¹¹ 11/1/04 APPROVe article draft, MRK-AGO0029733, at 735 (attached to 11/2/04 email from J. Bolognese to C. Lines et al., MRK-AGO0029732).

¹¹² 11/1/04 APPROVe article draft, MRK-AGO0029733, at 734-35, 744 (also stating that the table summarizing cardiovascular safety information “needs to have <18mo and >18mo results prominent – combined results are [secondary and] really not valid since non-proportional hazards was observed”) (attached to 11/2/04 email from J. Bolognese to C. Lines et al., MRK-AGO0029732).

¹¹³ 11/2/04 email from J. Bolognese to C. Lines et al., MRK-AGO0029753.

the trial, but no longer suggested that the results for the entire 0 to 36-month duration of the study be omitted.¹¹⁴

Consistent with Mr. Bolognese's suggestions, Dr. Lines revised the article to emphasize the difference in results between the two 18-month segments of the trial. For instance, a draft circulated by Dr. Lines on November 16, 2004 stated:

The difference [in cardiovascular thrombotic events] between treatments was due to events that occurred in the 2nd 18 months of the study, whereas the rates were similar for the first 18 months. This finding is illustrated in Figure 2 which shows the Kaplan-Meier plot for time to [confirmed thrombotic] events. The changing pattern of treatment effect over time was confirmed by the failed test for proportionality of hazards ($p=0.006$).¹¹⁵

In addition, pursuant to Dr. Quan's suggestion, the table reporting rate and relative risk analyses for thrombotic events was expanded to include results for the first and second 18-month segments in addition to the data for the entire duration of the APPROVe Trial.¹¹⁶

¹¹⁴ 11/2/04 email from J. Bolognese to C. Lines et al., MRK-AGO0029753, at 56-57 (attaching APPROVe article draft, MRK-AGO0029755).

¹¹⁵ APPROVe article draft, MRK-AHD0072653, at 654, 663 (also stating in the "Abstract" section: "The increased relative risk was observed beginning after 18 months of treatment; the results for the first 18 months showed similar risk for the 2 treatments.") (attached to 11/16/04 email from C. Lines to K. Horgan, MRK-AHD0072652).

¹¹⁶ APPROVe article draft, MRK-AHD0072653, at 673 (attached to 11/16/04 email from C. Lines to K. Horgan, MRK-AHD0072652).

Dr. Bresalier's* first draft of the "Results" section also referred to this observation: "The increase in thrombotic cardiovascular events associated with rofecoxib was not evident during the first 18 months of the trial. The reason for this finding is unclear." 11/18/04 APPROVe article draft, MRK-AHC0024622, at 623-624 (attached to 12/01/04 email from J. Baron* to R. Bresalier* and C. Lines, MRK-AHC0024621).

The emphasis on the difference in cardiovascular results between the first and second 18-month segments of the trial subsequently became a subject of discussion among the co-authors and non-author MRL scientists who reviewed the draft. In commenting on a draft in late November 2004, Dr. Konstam* pointed out that the analysis to evaluate the results by 18-month segments was not pre-specified in the APPROVe Trial Data Analysis Plan and the decision to choose 18 months as the breakpoint was based on the “appearance of the [Kaplan-Meier] curves.”¹¹⁷ Dr. Konstam* stated that he was “ambivalent about listing . . . early [versus] late [relative risks] at all ([as opposed to] merely point[ing] out the shape of the curve and stat[ing] the significance based on proportionality test).”¹¹⁸ Dr. Konstam* also expressed the view that, if the co-authors chose to display the 18-month data, they should explain in the article how the 18-month breakpoint was chosen and should not present the 18-month results in the same table as the results for the overall duration of the trial.¹¹⁹ (In a January 31, 2005 email to co-authors, Dr. Konstam* also expressed the view that the authors were “going out on a limb by emphasizing the 18 month issue.”¹²⁰)

¹¹⁷ 11/18/04 APPROVe article draft with M. Konstam’s* revisions, MRK-AHC0024647, at 658-59 (attached to 12/1/04 email from J. Baron* to R. Sandler* et al., MRK-AHC0024646).

¹¹⁸ 11/18/04 APPROVe article draft with M. Konstam’s* revisions, MRK-AHC0024647, at 658-59 (attached to 12/1/04 email from J. Baron* to R. Sandler* et al., MRK-AHC0024646).

¹¹⁹ 11/18/04 APPROVe article draft with M. Konstam’s* revisions, MRK-AHC0024647, at 658-59 (attached to 12/1/04 email from J. Baron* to R. Sandler* et al., MRK-AHC0024646).

¹²⁰ 1/31/05 email from M. Konstam* to C. Lines et al., MRK-AGO0077012, at 012.

Similarly, in his January 21, 2005 comments on the draft circulated to Dr. Lines (copying Dr. Gertz, Dr. Horgan, Dr. Reicin, and Ms. Lahner), Dr. Braunstein, who was not an author, wrote: “It is not clear why you are splitting the data at 18 months. Visual inspection of the curves? Are there statistical analyses to support this split[?]”¹²¹

Dr. Braunstein went on to suggest: “I would point out that there was a treatment by time interaction” – i.e., that the test for proportionality of hazards failed, indicating that the hazard ratio changed over time – “and then just discuss visual inspection of curves, unless there is some statistical basis for selecting a particular time point for analysis.”¹²²

On January 27, 2005, MRL co-authors and MRL scientists commenting on the draft held a teleconference to discuss whether, in Dr. Lines’ words, “too much [was] made in the text of the <18 month v >18 month split” and whether MRL scientists “[s]hould . . . recommend [to the external authors] reducing the amount of text on this.”¹²³ It was decided at that meeting that the 18-month analyses should be removed from the table displaying rates and relative risks for thrombotic events.¹²⁴ As discussed below, at a later point, however, Drs. Lines and Gertz decided to retain the

¹²¹ 1/13/05 APPROVe article draft, MRK-AAD0411667, at 678 (attached to 1/21/05 email from N. Braunstein to C. Lines et al., MRK-AAD0411664).

¹²² 1/13/05 APPROVe article draft, MRK-AAD0411667, at 678 (attached to 1/21/05 email from N. Braunstein to C. Lines et al., MRK-AAD0411664).

¹²³ 1/26/05 email from C. Lines to K. Horgan et al., MRK-AHD0000392, at 392 (attaching 1/26/05 APPROVe article draft, MRK-AHD0000393); see also 1/28/05 email from C. Lines to N. Braunstein et al., MRK-AGO0076263 (referring to “yesterday’s teleconf[erence]”).

¹²⁴ 1/28/05 APPROVe article draft, MRK-AGO0076264, at 297 (attached to 1/28/05 email from H. Quan to C. Lines, MRK-AGO0076263) (“Suggestion is to delete . . . month 0-18 + month 19-36 data from this table.”).

18-month-segment analysis in the table at issue because, in their view, it showed that the overall findings for the confirmed thrombotic and APTC composite cardiovascular event endpoints were similar.

The draft that Dr. Baron* submitted to the New England Journal of Medicine on February 6, 2005 set forth the 0-to-36-month results, and then stated in relevant part:

The increased relative risk was observed beginning after 18 months of treatment; in the first 18 months the risks for the 2 treatments were similar.

* * *

The difference between treatments was evident in the 2nd 18 months of the study, whereas event rates were similar for the first 18 months (Figure 2 [Kaplan-Meier plot for confirmed thrombotic events reproduced as Figure 3 above]). The changing pattern of treatment effect over time was confirmed by a failed test for proportionality of hazards ($p=0.014$). Findings for the Antiplatelet Trialists' Collaboration endpoint were similar (data not shown).¹²⁵

The table displaying rates and relative risks for thrombotic events included such results for the entire 36-month duration of the trial as well as its first and second 18-month segments.¹²⁶

¹²⁵ APPROVe article draft, MRK-AFV0398909, at 11, 19 (attached to 2/6/05 email from J. Baron* to R. Sandler* et al., MRK-AFV0398907-08).

¹²⁶ APPROVe article draft, MRK-AFV0398909, at 11, 19 (attached to 2/6/05 email from J. Baron* to R. Sandler* et al., MRK-AFV0398907-08).

ii. Discussion regarding the observed
“flattening” of the placebo curve.

As noted above, the placebo curve on the Kaplan-Meier plot of the cumulative incidence of confirmed thrombotic events in the APPROVe Trial became relatively flat after the 18-month mark, indicating that there were very few events in the placebo group after 18 months of continuous treatment (Figure 3). This, for some scientists, raised the issue of whether the seemingly anomalous flattening of the placebo curve explained or at least contributed to the separation of the Vioxx and placebo curves and the increased relative risk on Vioxx as compared to placebo.

On November 9, 2004, Dr. Lines circulated to his MRL colleagues a draft “Results” section of the article, which stated: “It can be seen that the difference between rofecoxib and placebo after 18 months partly reflects a flattening of the placebo slope after 18 months compared with the preceding [sic] 18 months.”¹²⁷ Subsequent drafts of the “Results” section expanded on this discussion:

The difference between rofecoxib and placebo after 18 months partly reflects a flattening of the placebo curve after 18 months. Within the rofecoxib group, the rate of confirmed events was similar or only slightly higher in the 2nd 18 months of the study, compared with the first 18 months, as shown by a month 19-36 versus 0-18 relative risk of 1.32 (ninety five percent confidence interval 0.74, 2.35). By contrast, the rate within the placebo group was lower in the 2nd 18 months compared with the first 18 months, as shown by a month 19-36 versus month 0-18

¹²⁷ 11/9/04 APPROVe article draft with J. Bolognese revisions, MRK-AGO0072373, at 383 (attached to 11/9/04 email from J. Bolognese to C. Lines et al., MRK-AGO0072372).

relative risk of 0.34 (ninety five percent confidence interval
0.11, 0.88).¹²⁸

Similarly, Dr. Bresalier's* first draft of the "Discussion" section of the
manuscript, circulated to co-authors on November 30, 2004, attributed the increased
relative risk on Vioxx in part to the flattening of the placebo curve:

The difference between rofecoxib and placebo after 18
months partly reflects a flattening of the placebo event
curve after 18 months, compared with a slightly higher
event rate in the rofecoxib group compared with the first 18
months of the trial. The reason for this unexpected finding
is unclear, and could represent a chance effect, or the result
of differential drop out of at-risk individuals in the placebo
group.¹²⁹

On December 10, 2004, however, Dr. Baron* sent an email to the co-author group
stating: "I deleted the sentences elaborating on the flattening of the placebo curve
because I wasn't sure where we were going with that. My guess is that we won't try to
make much of a deal out of it, so I thought maybe we should note the pattern, and leave it
at that."¹³⁰

Dr. Konstam* also expressed the view that the article should not refer to the
flattening of the placebo curve for reasons he explained in his comments on one of the
drafts:

¹²⁸ 1/13/05 APPROVe article draft, MRK-AAD0411667, at 678-679 (attached to 1/21/05 email from
N. Braunstein to C. Lines et al., MRK-AAD0411664).

¹²⁹ APPROVe article draft, MRK-AHD0072873, at 875 (attached to 11/30/04 email from R. Bresalier* to
C. Lines, MRK-AHD0072847).

¹³⁰ 12/10/04 email from J. Baron* to R. Bresalier* et al., MRK-AHC0030956, at 956.

I disagree with including this entire suggestion of a ‘flattening’ of the placebo rate. I don’t agree that anything can be made of it. Ther [sic] is no statistical test for the validity of this subjective observation. I don’t believe it; Anyway, even if it were true, it would not dispel the adverse impact of rofecoxib, which did not manifest this ‘flattening’. We will be criticized for trying to call attention to this.¹³¹

In response to Dr. Konstam’s* comment, Dr. Lines deleted the reference to the flattening of the placebo curve from the “Discussion” section.¹³² The “Results” section, however, still stated:

The difference between rofecoxib and placebo beginning after 18 months partly reflects a flattening of the placebo curve after 18 months. Within the rofecoxib group, the rate of confirmed thrombotic cardiovascular serious events was similar or only slightly higher in the 2nd 18 months of the study, compared with the first 18 months, as shown by a month 19-36 versus month 0-18 relative risk of 1.32 (ninety five percent confidence interval 0.74, 2.35). By contrast, the rate within the placebo group was lower in the 2nd 18 months compared with the first 18 months, as shown by a month 19-36 versus month 0-18 relative risk of 0.34 (ninety five percent confidence interval 0.11, 0.88).¹³³

On January 31, 2005, after seeing that the “Results” section still referred to the flattening of the placebo curve, Dr. Konstam* wrote another email to the co-authors stating:

¹³¹ 1/13/05 APPROVe article draft, MRK-AHC0025599, at 610 (attached to 1/20/05 email from M. Konstam* to K. Horgan and J. Baron*, MRK-AHC002597).

¹³² 1/31/05 email from C. Lines to R. Bresalier* et al., MRK-AAD0421787, at 87.

¹³³ 1/13/05 APPROVe article draft, MRK-AAD0421789, at 800 (attached to 1/31/05 email from C. Lines to B. Gertz et al., MRK-AAD0421787).

I am deeply concerned about mentioning the notion that the result relates to a “flattening” of the placebo curve. I believe that this observation is weak, not supportable by statistics, and of no identifiable clinical meaning. Analysis of shapes of curves are laden with hazards to begin with, and we are already going out on a limb by emphasizing the 18 month issue. I feel that mentioning this “flattening” issue, will subject us to the accusation that we are trying to minimize the key finding – an adverse effect of rofecoxib. I know we are not trying to do this, so we shouldn’t look like we are.¹³⁴

In response to Dr. Konstam’s* email, the co-authors decided to remove all references to the flattening of the placebo curve, and the draft that Dr. Baron* sent by email to the New England Journal of Medicine on February 6, 2005 did not mention the flattening of the placebo curve.¹³⁵

iii. Discussion regarding results for the APTC composite cardiovascular event endpoint.

From the outset, drafts of the APPROVe article focused on the confirmed thrombotic event endpoint, but also mentioned certain results obtained for the APTC composite cardiovascular event endpoint. For instance, the January 13, 2005 draft (which incorporated final APPROVe Trial data) cited the proportionality p-value and included a Kaplan-Meier plot only for the confirmed thrombotic event endpoint, but referred to the 18-month relative risk analyses for both endpoints. It stated:

¹³⁴ 1/31/05 email from M. Konstam* to C. Lines et al., MRK-AGO0077012, at 012; see also 2/01/05 email from B. Gertz to C. Lines et al., MRK-AHD0075697 (“I think Marv [Konstam*] makes good points on both counts.”).

¹³⁵ See APPROVe article draft, MRK-AFV0398909-40 (attached to 2/6/05 email from J. Baron* to R. Sandler* et al., MRK-AFV0398907-08).

Summary data for confirmed and investigator reported serious cardiovascular thrombotic adverse events and APTC events are shown in Table 2 [displaying data for both confirmed thrombotic and APTC composite cardiovascular event endpoints]. While the absolute rates of events in both groups was low, there was an approximately 2-fold overall increased relative risk for a confirmed event in the rofecoxib versus the placebo group [confirmed thrombotic and APTC endpoints]. The difference between treatments was evident in the 2nd 18 months of the study, whereas the rates were similar for the first 18 months (Figure 2 [Kaplan-Meier plot for the confirmed thrombotic event endpoint only]). The changing pattern of treatment effect over time was confirmed by the failed test for proportionality of hazards (p=0.014) [confirmed thrombotic endpoint only].¹³⁶

In a January 20, 2005 e-mail commenting on this paragraph of the January 13 draft, Dr. Konstam* asked: “What about APTC events, since this is mentioned in the methods?”¹³⁷ In response to Dr. Konstam’s* comment, Dr. Lines made several changes to the paragraph quoted above, including a new sentence at the end stating: “Similar results were seen for the APTC endpoint.”¹³⁸ The paragraph as revised by Dr. Lines read:

For confirmed thrombotic cardiovascular serious events, the absolute rates of events in both groups were low, but there was a 1.92 overall increased relative risk for a confirmed event in the rofecoxib versus the placebo group.

¹³⁶ 1/13/05 APPROVe article draft, MRK-AHC0025599, at 610 (attached to 1/20/05 email from M. Konstam* to K. Horgan and J. Baron*, MRK-AHC0025597). The reference to the “failed test for proportionality of hazards” was to the fact that the p-value of 0.014 was less than p=0.05.

¹³⁷ 1/13/05 APPROVe article draft, MRK-AHC0025599, at 610 (attached to 1/20/05 email from M. Konstam* to K. Horgan and J. Baron*, MRK-AHC0025597).

¹³⁸ 1/26/05 APPROVe article draft, MRK-AHD0000393, at 404 (attached to 1/26/05 email from C. Lines to K. Horgan and et al., MRK-AHD0000392).

The difference between treatments was evident in the 2nd 18 months of the study, whereas the rates were similar for the first 18 months (Figure 2 [Kaplan-Meier plot for confirmed thrombotic events reproduced as Figure 3 above]). The changing pattern of treatment effect over time was confirmed by the failed test for proportionality of hazards ($p=0.014$). Similar results were seen for the APTC endpoint.¹³⁹

In response to this revision, on January 28, 2005, Dr. Quan noted: “[A]ctually, for APTC, $p=0.119$ ” – a p-value too high to conclude based on this test alone that the hazard ratio for the APTC composite cardiovascular event endpoint changed over time – “even though the first 18 months and the last 18 months results were similar to those of [the confirmed thrombotic event endpoint].”¹⁴⁰ As discussed in Section B.3.e.i of this Appendix, at the January 27, 2005 teleconference among MRL authors and reviewers of the draft, in response to the concern voiced by Dr. Konstam* and Dr. Braunstein that the 18-month analysis was being overemphasized, it was suggested that the draft manuscript omit rate and relative risk data for the two 18-month intervals from Table 2 and retain such data only for the overall 36-month duration of the trial. When Dr. Quan noted that the proportionality p-value for the APTC composite cardiovascular event endpoint, unlike that for the confirmed thrombotic event endpoint, was not statistically significant, Dr. Lines responded:

¹³⁹ 1/26/05 APPROVe article draft, MRK-AHD0000393, at 404 (attached to 1/26/05 email from C. Lines to K. Horgan et al., MRK-AHD0000392).

¹⁴⁰ 1/28/05 APPROVe article draft, MRK-AGO0076264, at 275 (attached to 1/28/05 email from H. Quan to C. Lines, MRK-AGO0076263).

Thanks Hui

The fact that the APTC endpoint did not fail the stat test of proportionality, unlike the confirmed endpoint [sic], might be an argument for keeping in the month 0-18, month 19-36 split in Table 2 (so people can see that the pattern [for the APTC composite cardiovascular event endpoint] is similar to [the pattern for the confirmed thrombotic event endpoint]).¹⁴¹

Dr. Lines suggested as an alternative including a Kaplan-Meier plot for the APTC composite cardiovascular event endpoint in addition to the Kaplan-Meier plot for the confirmed thrombotic event endpoint already included in the draft so that readers could see the similarity between the two plots (see Figures 3 and 4).¹⁴² The following day, Dr. Gertz responded to Dr. Lines: “I would tend to agree with that . . . provides some utility for the table. Would not see need for a second APTC [Kaplan-Meier] plot.”¹⁴³

Consistent with Dr. Gertz’s email, the draft of the APPROVe article submitted to Dr. Drazen* of the New England Journal of Medicine and the published version of the article included a table displaying confirmed thrombotic and APTC composite cardiovascular event endpoint results for the first and second 18-month intervals of the APPROVe Trial as well as the entire 36-month period, included a Kaplan-Meier plot for the confirmed thrombotic event endpoint, and did not include a Kaplan-Meier plot for the

¹⁴¹ 1/28/05 email from C. Lines to H. Quan (cc: N. Braunstein, B. Gertz, A. Reicin, J. Lahner, K. Horgan, B. Oxenius, J. Bolognese), MRK-AHC0025721, at 21.

¹⁴² 1/28/05 email from C. Lines to H. Quan (cc: N. Braunstein, B. Gertz, A. Reicin, J. Lahner, K. Horgan, B. Oxenius, J. Bolognese), MRK-AHC0025721, at 21.

¹⁴³ 1/29/05 email from B. Gertz to C. Lines, MRK-AHC0025721, at 21.

APTC event endpoint. The proportionality p-value for the APTC composite cardiovascular endpoint was not stated.¹⁴⁴

- f. The manuscript submitted to the New England Journal of Medicine on February 6, 2005.

On or shortly before February 6, 2005, Dr. Drazen* emailed Dr. Baron* to request that the authors submit the manuscript to the New England Journal of Medicine for review by February 7, 2005.¹⁴⁵ Around the same time, Dr. Drazen* informed Dr. Baron* that the New England Journal of Medicine had received manuscripts concerning cardiovascular results from trials of other selective Cox-2 inhibitors and wanted to publish the APPROVe article in the same issue with the other manuscripts prior to the FDA Advisory Committee meeting regarding the safety of selective and non-selective Cox-2 inhibitors, which was scheduled for February 16 to 18, 2005.¹⁴⁶ At that point, the co-authors of the article had not yet finalized the text of the article or proofread the parts

¹⁴⁴ Bresalier* RS, Sandler* RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med. 2005;352:1092-102, at 097, MRK-ARQ0000659, at 64.

¹⁴⁵ 2/6/05 email from J. Baron* to R. Sandler* et al., MRK-AFV0398907.

¹⁴⁶ 2/7/05 email from J. Baron* to R. Sandler* et al., MRK-AFV0398907; 1/10/05 email from J. Drazen* to J. Baron*, MRK-AAD0421174 (“If we can get the paper soon we could potentially have it up on the web for the [February 16 to 18, 2005] advisory panel. I anticipate this will increase its impact.”); 2/12/05 email from K. Horgan to J. Wainwright et al., MRK-AFV0403673, at 74 (“the journal is eager to place [the APPROVe article] on its web site prior to the FDA [Advisory Committee Meeting].”); see also 2/14/04 email from K. Sokolic to C. Smith et al., MRK-AJJ0018179 (attaching articles “being published in NEJM this week,” including articles entitled “Cardiovascular Risk Associated with Celecoxib in a Clinical Trial for Colorectal Adenoma Prevention,” MRK-AJJ0018181, and “Complications of the Cox-2 Inhibitors Parecoxib and Valdecoxib After Cardiac Surgery,” MRK-AJJ0018198).

that had been completed.¹⁴⁷ Given Dr. Drazen's* request, however, Dr. Baron* incorporated some of the comments he had received previously from his co-authors and, on February 6, 2005, sent the manuscript to Dr. Drazen* by email.¹⁴⁸

In the draft of the manuscript submitted to Dr. Drazen*, the paragraph discussing the hazard ratio over time stated:

There were a total of 121 patients with investigator reported thrombotic cardiovascular adverse events (77 in the rofecoxib group and 44 in the placebo group) The absolute rates of confirmed thrombotic cardiovascular events were low in both groups, but there was increased relative risk in the rofecoxib group (relative risk 1.92; ninety five percent confidence interval 1.19 – 3.11). The difference between treatments was evident in the 2nd 18 months of the study, whereas event rates were similar for the first 18 months (Figure 2 [Kaplan-Meier plot for confirmed thrombotic events reproduced as Figure 3 above]). The changing pattern of treatment effect over time was confirmed by a failed test for proportionality of hazards (p=0.014). Findings for the Antiplatelet Trialists' Collaboration endpoint were similar (data not shown).¹⁴⁹

The proportionality p-value of p=0.014 resulted from the Linear Time Test although the submitted draft erroneously described the test for proportionality of hazards

¹⁴⁷ See 1/28/05 email from J. Baron* to C. Lines, MRK-AHD0075359 (“I thinks people need to see the new [i.e., updated December 2004] and also have more time with a draft.”).

¹⁴⁸ 2/6/05 email from J. Baron* to R. Sandler* et al., MRK-AFV0398907 (“[The New England Journal of Medicine] needed the paper sooner than our strategy would easily allow us Bob [Bresalier*] was still in the air, so I patched together the paper as best as I could in a few minutes – taking into account Marv Konstam's comments and the numbers that Jim [Bolognese] had sent. . . . Then I ‘submitted’ it to Drazin [sic] personally (not to the NEJM....). . . . We can change the paper in the next few days if we want, however.”).

¹⁴⁹ APPROVe article draft, MRK-AFV0398909, at 19 (attached to 2/6/05 email from J. Baron* to R. Sandler* et al., MRK-AFV0398907-08).

performed to obtain the reported $p=0.014$ as one evaluating “the time-by-log (time) interaction in the COX model” (i.e., the Logarithm of Time Test).¹⁵⁰ The draft also referred in several other places to the 18-month issue, stating that “[t]he increased relative risk was observed beginning after 18 months of treatment,” and that, “in the first 18 months[,] the risks for the 2 treatments were similar,”¹⁵¹ and that there was an “absence of an adverse effect treatment [sic] on cardiovascular outcomes during the first 18 months of APPROVe”¹⁵² and a “delayed emergence” of an increased risk.¹⁵³

The table displaying the rate and relative risk analyses for thrombotic events included results for the confirmed thrombotic and APTC composite cardiovascular event endpoints for the entire 36-month duration of the trial as well as for the first and second 18-month segments of the APPROVe Trial.¹⁵⁴

Finally, the draft reported analyses of certain adverse events that might have been related to thrombotic cardiovascular events but were not included in either composite endpoint – namely, hypertension, edema, and a composite endpoint of congestive heart

¹⁵⁰ APPROVe article draft, MRK-AFV0398909, at 16 (attached to 2/6/05 email from J. Baron* to R. Sandler* et al., MRK-AFV0398907-08).

¹⁵¹ APPROVe article draft, MRK-AFV0398909, at 11 (attached to 2/6/05 email from J. Baron* to R. Sandler* et al., MRK-AFV0398907-08).

¹⁵² APPROVe article draft, MRK-AFV0398909, at 24 (attached to 2/6/05 email from J. Baron* to R. Sandler* et al., MRK-AFV0398907-08).

¹⁵³ APPROVe article draft, MRK-AFV0398909, at 25 (attached to 2/6/05 email from J. Baron* to R. Sandler* et al., MRK-AFV0398907-08).

¹⁵⁴ APPROVe article draft, MRK-AFV0398909, Table 2 at 39 (attached to 2/6/05 email from J. Baron* to R. Sandler* et al., MRK-AFV0398907-08).

failure, pulmonary edema and cardiac failure. Specifically, the draft stated:

(i) “[c]ompared to placebo, the rofecoxib group had a higher percentage of patients with hypertension adverse experiences (rofecoxib = 29.3%, placebo = 16.9%, $p < 0.001$), edema adverse experiences (rofecoxib = 8.6%, placebo = 5.9%, $p = 0.008$), and the grouping of congestive heart failure, pulmonary edema or cardiac failure (rofecoxib = 1.3%, placebo = 0.3%, $p = 0.004$)”; and (ii) “[t]he Kaplan-Meier plots for hypertension, edema, and congestive heart failure, pulmonary edema, or cardiac failure all showed early separation between treatment groups with no statistically significant departures from [proportional] hazards over time (data not shown).”¹⁵⁵

g. Peer reviewers’ comments on the submitted draft and subsequent revisions.

i. Summary.

By letter dated February 9, 2005, Dr. Curfman*, Executive Editor of the New England Journal of Medicine, informed Dr. Baron* that the manuscript had “not been accepted for publication in its current form,” but that the New England Journal of Medicine “would be very willing to consider an extensively revised version of the manuscript that successfully addressed” the attached comments by five anonymous peer reviewers.¹⁵⁶ In the letter, Dr. Curfman* stated that the New England Journal of

¹⁵⁵ APPROVe article draft, MRK-AFV0398909, at 20 (attached to 2/6/05 email from J. Baron* to R. Sandler* et al., MRK-AFV0398907-08).

¹⁵⁶ 2/9/05 letter from G. Curfman* to J. Baron*, MRK-AFV0399548, at 48.

Medicine hoped the authors would be able to “complete the revisions very soon, given the timely nature and the public health importance of the topic.”¹⁵⁷

Over the next few days, Drs. Baron* and Bresalier*, with input from other co-authors, worked on revising the article.¹⁵⁸ In revising the article, Dr. Baron* held several telephone conferences with Dr. Curfman*.¹⁵⁹ In addition, the authors also prepared a document responding to each comment made by the peer reviewers, although it is not clear whether that document was sent to the New England Journal of Medicine.

This section provides an overview of the reviewers’ comments and discusses in detail comments (and authors’ subsequent revisions of the article) concerning two specific issues: (i) the manuscript’s discussion of the post-hoc analyses showing that, in the APPROVe Trial, the increased relative risk on Vioxx versus placebo did not become apparent until 18 months of continuous treatment; and (ii) the authors’ pre-submission decision not to include in the article analyses of investigator-reported events (and instead include only analyses of confirmed events, i.e., the confirmed thrombotic and APTC cardiovascular events).

¹⁵⁷ 2/9/05 letter from G. Curfman* to J. Baron*, MRK-AFV0399548, at 49 (“The editors recognize that we are asking for extensive changes, but we hope you will be able to complete the revisions very soon, given the timely nature and the public health importance of the topic. Please understand that we will carefully check your revised manuscript to be certain that all of the changes requested by the reviewers and editors have been made.”).

¹⁵⁸ 2/11/05 email from C. Lines to B. Gertz et al., MRK-AJK0016400 (“Here’s the revised paper from the external [authors].”).

¹⁵⁹ 11/21/05 deposition of G. Curfman* at 132 (In re Vioxx Litig., MDL No. 1657, E.D. La.).

ii. The 18-month issue.

(a) Peer reviewers' comments.

In their comments, four of the five peer reviewers asked that the article place less emphasis on analyses suggesting that the increased relative risk on Vioxx versus placebo emerged in the second 18 months of the trial because (i) these analyses were post-hoc, that is, were not conducted pursuant to a pre-specified data analysis plan; and (ii) Kaplan-Meier curves for events involving congestive heart failure, pulmonary edema, and cardiac failure – cardiovascular-related adverse events that were not included in either endpoint – began to separate much earlier (at around the 5-month point). These Kaplan-Meier curves had been discussed but not included in the submitted draft. For example, one of the reviewers wrote:

The [18-month] analysis . . . seems to be presented as if it were a pre-specified hypothesis. Was the stratification of the analysis of 0-18 months and 19-36 months pre-specified in the original protocol? If not, then this analysis and the post-hoc choice of the cutpoint at 18 months need to be presented as an exploratory analysis, and they do not belong in Table 2 [presenting relative risks with 95% confidence intervals for the two composite endpoints].¹⁶⁰

A different reviewer made a similar comment:

The '18 month hypothesis'. The [manuscript] aggressively promotes the safety of up to 18 months of use of rofecoxib. This goes beyond the data of the study. The analysis by duration of use is an example of a post hoc subgroup analysis, because the variation of the effect by duration was not an a priori hypothesis and the 18 month cutpoint was

¹⁶⁰ Reviewer B's comments on APPROVe article draft, MRK-AFV0399552, at 53 (attached to 2/9/05 letter from G. Curfinan* to J. Baron*, MRK-AFV0399548).

chosen by inspection of the data.

* * *

The data do provide evidence, based on a post hoc analysis, that risk increases with duration of use. The remarks in the [manuscript] should be congruent with these data and appropriately cautious about the finding of a subgroup analysis.¹⁶¹

Another reviewer stated:

Much has been made of the lack of a safety signal on thrombotic events until 18 months. However it is clear that the heart failure and pulmonary edema plots do show such a separation from the beginning. It is hard to see why these data were omitted. They must be included.¹⁶²

In light of the peer reviewers' comments, in his February 9, 2005 letter (attaching the comments), Dr. Curfman* asked that the co-authors, among other things:

Remove assertion that increased risk was observed only after 18 months, because the event curves for other adverse events [i.e., (i) hypertension; (ii) edema; and (iii) congestive heart failure, pulmonary edema, and cardiac failure] separate early. Also, the focus on the first 18 months was not pre-specified and is a post hoc analysis.¹⁶³

¹⁶¹ Reviewer C's comments on APPROVe article draft, MRK-AFV0399556, at 56-57 (attached to 2/9/05 letter from G. Curfman* to J. Baron*, MRK-AFV0399548).

¹⁶² Reviewer D's comments on APPROVe article draft, MRK-AFV0399560, at 60 (attached to 2/9/05 letter from G. Curfman* to J. Baron*, MRK-AFV0399548).

¹⁶³ 2/9/05 letter from G. Curfman* to J. Baron*, MRK-AFV0399548, at 49; see also 11/21/05 deposition of G. Curfman* at 144 (In re Vioxx Litig., MDL No. 1657, E.D. La.) (Dr. Curfman* testifying that he "thought that if you . . . repeated the study in exactly the same way, it would be very unlikely you would see that kind of sharp separation at 18 months" and that "a significant amount" of his February 9-15, 2005 telephone communications with Dr. Baron* focused on the way in which the manuscript discussed the analysis showing that the increased relative risk on Vioxx did not emerge until 18 months of continuous treatment).

(b) Authors' responses.

(i) Question of pre-specification.

Upon reviewing the reviewers' comments regarding the 18-month analyses, Mr. Bolognese sent to other co-authors, including Dr. Baron*, what Mr. Bolognese characterized as, "comments that might help resolve some of the reviewers' [sic] comments."¹⁶⁴ With regard to the reviewers' comments that the analysis comparing cardiovascular results from the first 18 months of the trial to the results from the second 18 months was not pre-specified, Mr. Bolognese expressed the view that the co-authors "probably should mention"¹⁶⁵ that:

- the APPROVe Trial was part of the pre-planned Protocol 203 – a combined analysis of cardiovascular event data from the APPROVe, VICTOR and ViP Trials;
- Protocol 203 had a Statistical Data Analysis Plan;¹⁶⁶
- "endpoints & analyses reported for APPROVe were intended for [Protocol] 203"; and
- "the unexpected definitive signal [in] APPROVe precluded completion of [Protocol] 203."¹⁶⁷

Further, Mr. Bolognese suggested mentioning that:

- the "test of proportional hazards was pre-specified for [Protocol] 203";

¹⁶⁴ 2/10/05 email from J. Bolognese to J. Baron* et al., MRK-AGO0107791.

¹⁶⁵ It is unclear whether Mr. Bolognese suggested that these comments be mentioned in the article or in responses to reviewers' comments.

¹⁶⁶ The Statistical Data Analysis Plan for Protocol 203 had been finalized in January 2004. Protocol 203 Statistical Data Analysis Plan, MRK-AAB0083231.

¹⁶⁷ 2/10/05 email from J. Bolognese to J. Baron* et al., MRK-AGO0107791.

- the Data Analysis Plan for Protocol 203 stated that in the event of a statistically significant departure from the proportional hazards assumption, additional exploratory analyses would be carried out to explain the departure; and
- “[t]he 18[-month] breakdown was strongly suggested by [the Kaplan-Meier plot pre-specified in the Statistical Data Analysis Plan for Protocol 203], [and additional exploratory analyses] were employed for APPROVe.”¹⁶⁸

On February 11, 2005, Dr. Baron* circulated the first draft of a proposed written response to the peer reviewers’ comments. The draft, among other things, stated that “[t]he time-dependent analyses were part of a pre-specified analysis,” referred to “a pre-planned assessment of the cardiovascular safety of rofecoxib, with specified endpoints and analysis,” and a few pages later explained:

The stratification of analysis on 0-18 months and 19-36 months was not prespecified. However, the test for proportional hazards WAS part of the planned safety analysis referred to above, and was to be followed by an exploration of time-period patterns. This exploration resulted in the 0-18 months and 19-36 months groupings.¹⁶⁹

Later that day, Mr. Bolognese circulated “suggested edits/comments” regarding Dr. Baron’s* draft response to the peer reviewers’ comments.¹⁷⁰ Next to Dr. Baron’s* proposed statement that “time-dependent analyses were part of a pre-specified analysis,” Mr. Bolognese wrote: “time-dependent analyses were not pre-specified; [the APPROVe

¹⁶⁸ 2/10/05 email from J. Bolognese to J. Baron* et al., MRK-AGO0107791.

¹⁶⁹ Draft responses to reviewers’ comments on APPROVe article draft, MRK-AGO0103978, at 79, 81, 83 (attached to 2/11/05 email from J. Baron* to K. Horgan et al., MRK-AGO0103976).

¹⁷⁰ 2/11/05 email from J. Bolognese to K. Horgan et al., MRK-AGO0103998 (attaching draft responses to peer reviewers’ comments with J. Bolognese’s comments, MRK-AGO0104000).

Trial External Safety Monitoring Board] carried some of them out to assist their evaluation of patient safety; we added to them to pursue their findings.”¹⁷¹

Mr. Bolognese did not comment on Dr. Baron’s* proposed statement that “the test for proportional hazards WAS part of the planned safety analysis . . . and was to be followed by an exploration of time-period patterns.”¹⁷²

While it is not clear whether the document responding to the reviewers’ comments was sent to the New England Journal of Medicine, what appears to be the latest version of the document stated with regard to the pre-specification question:

¹⁷¹ Draft responses to reviewers’ comments on APPROVe article draft with J. Bolognese’s comments, MRK-AGO0104000, at 01 (attached to 2/11/05 email from J. Bolognese to K. Horgan et al., MRK-AGO0103998).

¹⁷² Draft responses to reviewers’ comments on APPROVe article draft with J. Bolognese’s comments, MRK-AGO0104000, at 05 (attached to 2/11/05 email from J. Bolognese to K. Horgan et al., MRK-AGO0103998).

Peer Reviewers' / Dr. Curfman's* Comments	Authors' Proposed Responses
<p>“Remove assertion that increased risk was observed only after 18 months, because the event curves for other adverse events separate early. Also, the focus on the first 18 months was not pre-specified.”¹⁷³</p>	<p>“These matters are discussed in detail below. The time-dependent analyses were part of a pre-specified analysis.”¹⁷⁴</p>
<p>“Cardiovascular events do not appear to have been an a priori outcome of primary interest and defined as such in the original protocol. . . . At what point in the trial was the decision made to review carefully and adjudicate blindly cardiovascular events?”¹⁷⁵</p>	<p>“We have clarified in the manuscript that the cardiovascular analysis was part of a pre-planned assessment of the cardiovascular safety of rofecoxib, with specified endpoints and analysis.”¹⁷⁶</p>
<p>“The analysis stratified on follow-up time seems to be presented as if it were a pre-specified hypothesis. Was the stratification of the analysis of 0-18 months and 19-36 months pre-specified in the original protocol? If not, then this analysis and the post-hoc choice of the cutpoint at 18 months need to be presented as an exploratory analysis, and they do not belong in Table 2 [presenting relative risks with 95% confidence intervals for the two composite endpoints].”¹⁷⁷</p>	<p>“The stratification of analysis on 0-18 months and 19-36 months was not prespecified. However, the test for proportional hazards WAS part of the planned safety analysis referred to above, and in [sic] the specified analysis was to be followed by consideration of time-period patterns. This resulted in the 0-18 months and 19-36 months groupings.”¹⁷⁸</p>

¹⁷³ 2/9/05 letter from G. Curfman* to J. Baron*, MRK-AFV0399548, at 49.

¹⁷⁴ Authors' responses to reviewers' comments on APPROVe article draft, MRK-AHD0000061, at 62.

¹⁷⁵ Reviewer B's comments on APPROVe article draft, MRK-AFV0399552, at 52 (attached to 2/9/05 letter from G. Curfman* to J. Baron*, MRK-AFV0399548).

¹⁷⁶ Authors' responses to reviewers' comments on APPROVe article draft, MRK-AHD0000061, at 65.

¹⁷⁷ Reviewer B's comments on APPROVe article draft, MRK-AFV0399552, at 53 (attached to 2/9/05 letter from G. Curfman* to J. Baron*, MRK-AFV0399548).

¹⁷⁸ Authors' responses to reviewers' comments on APPROVe article draft, MRK-AHD0000061, at 67.

Peer Reviewers’/Dr. Curfman’s* Comments	Authors’ Proposed Responses
<p>“If the test for proportional hazards revealed this issue, it is reasonable to include the finding within a table that explores other potential interactions. Stratification on aspirin should be fully presented. . . .”¹⁷⁹</p>	<p>“As noted above the proportional hazards analysis was part of the prespecified analysis plan. We are hesitant to emphasize further the subgroup/interaction analyses . . . because of the well-known variability of subgroup findings.”¹⁸⁰</p>
<p>“. . . The ‘18 month hypothesis’. The [manuscript] aggressively promotes the safety of up to 18 months of use of rofecoxib. This goes beyond the data of the study. The analysis by duration of use is an example of a post hoc subgroup analysis, because the variation of the effect by duration was not an a priori hypothesis and the 18 month cutpoint was chosen by inspection of the data.”¹⁸¹</p>	<p>“As noted above, the test for proportional hazards was part of the intended safety analysis. We now emphasize this in the methods section”¹⁸²</p>
<p>“Given that the duration finding is from a post-hoc subgroup analysis, is it appropriate to give it the prominence that it has in Table 2?”¹⁸³</p>	<p>“See response above.”¹⁸⁴</p>

¹⁷⁹ Reviewer B’s comments on APPROVe article draft, MRK-AFV0399552, at 53 (attached to 2/9/05 letter from G. Curfman* to J. Baron*, MRK-AFV0399548).

¹⁸⁰ Authors’ responses to reviewers’ comments on APPROVe article draft, MRK-AHD0000061, at 67.

¹⁸¹ Reviewer C’s comments on APPROVe article draft, MRK-AFV0399556, at 56 (attached to 2/9/05 letter from G. Curfman* to J. Baron*, MRK-AFV0399548).

¹⁸² Authors’ responses to reviewers’ comments on APPROVe article draft, MRK-AHD0000061, at 70.

¹⁸³ Reviewer C’s comments on APPROVe article draft, MRK-AFV0399556, at 57 (attached to 2/9/05 letter from G. Curfman* to J. Baron*, MRK-AFV0399548).

¹⁸⁴ Authors’ responses to reviewers’ comments on APPROVe article draft, MRK-AHD0000061, at 71.

Peer Reviewers’/Dr. Curfman’s* Comments	Authors’ Proposed Responses
<p>“I am not impressed that we understand enough about the effect of exposure time to provide the reassurance offered by the authors here.”¹⁸⁵ (The peer reviewer was referring to the following statement in the submitted draft: “Differences between the present results and these earlier clinical trial data may be related to differences in exposure time, a possibility supported by the absence of an adverse effect treatment [sic] on cardiovascular outcomes during the first 18 months of APPROVe.”¹⁸⁶)</p>	<p>“We have softened the inference somewhat – and strengthened the scientific basis for the exposure time considerations.”¹⁸⁷ (As stated below, the revised article read: “Differences between our results and these earlier clinical-trial data may be related to differences in defined end points or the duration of treatment, a possibility supported by the apparent absence of a difference in adjudicated thrombotic events during the first 18 months of our study.”¹⁸⁸)</p>

In response to the peer reviewers’ comments regarding the post-hoc nature of the analyses at issue and the manuscript’s emphasis on the results of such analyses, the authors made several revisions to the article. First, they identified the analysis assessing the difference between results for the first and second 18 months of continuous treatment as a “post-hoc” analysis. Second, they added sentences stating that “[m]onitoring and analysis of the cardiovascular events in the [APPROVe] trial were part of a planned assessment of the cardiovascular safety of rofecoxib” and that “[a] test of the

¹⁸⁵ Reviewer D’s comments on APPROVe article draft, MRK-AFV0399560, at 61 (attached to 2/9/05 letter from G. Curfman* to J. Baron* MRK-AFV0399548).

¹⁸⁶ APPROVe article draft, MRK-AFV0398909, at 24 (attached to 2/6/05 email from J. Baron* to R. Sandler* et al., MRK-AFV0398907-08).

¹⁸⁷ Authors’ responses to reviewers’ comments on APPROVe article draft, MRK-AHD0000061, at 76.

¹⁸⁸ Bresalier* RS, Sandler* RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med. 2005;352:1092-102, at 099, MRK-ARQ0000659, at 666.

proportional-hazards assumption was specified in the cardiovascular-analysis plan.”¹⁸⁹

Third, they changed as follows the wording and order of certain paragraphs in order to place less emphasis on the results of the post-hoc time-dependent analyses:

Text of the draft submitted to the <u>New England Journal of Medicine</u> for peer review	Text of the published article (incorporating authors’ post-peer review revisions)
“Abstract” section	
“The increased relative risk was observed beginning after 18 months of treatment; in the first 18 months the risks for the 2 treatments were similar.”	“The increased relative risk became apparent after 18 months of treatment; during the first 18 months, the event rates were similar in the two groups.” ¹⁹⁰

¹⁸⁹ Bresalier* RS, Sandler* RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med. 2005;352:1092-102, at 094, 095, MRK-ARQ0000659, at 661, 662.

¹⁹⁰ Bresalier* RS, Sandler* RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med. 2005;352:1092-102, at 092, MRK-ARQ0000659, at 659.

Text of the draft submitted to the <u>New England Journal of Medicine</u> for peer review	Text of the published article (incorporating authors' post-peer review revisions)
"Results" section	
<p>"There were a total of 121 patients with investigator reported thrombotic cardiovascular adverse events (77 in the rofecoxib group and 44 in the placebo group) (Table 2). The absolute rates of confirmed thrombotic cardiovascular events were low in both groups, but there was increased relative risk in the rofecoxib group (relative risk 1.92; ninety five percent confidence interval 1.19 – 3.11). The difference between treatments was evident in the 2nd 18 months of the study, whereas event rates were similar for the first 18 months (Figure 2). The changing pattern of treatment effect over time was confirmed by a failed test for proportionality of hazards (p=0.014). Findings for the Antiplatelet Trialists' Collaboration endpoint were similar (data not shown).</p> <p>The types of confirmed thrombotic cardiovascular serious adverse events are shown in Table 3. The difference between treatments was mainly due to an increase in myocardial infarctions and strokes in the rofecoxib group. There were 10 deaths in the rofecoxib group and 10 deaths in the placebo group. Cardiovascular causes of death were myocardial infarction (2 rofecoxib, 3 placebo), sudden cardiac death (3 rofecoxib, 1 placebo), ischemic cerebrovascular stroke (1 rofecoxib), and hemorrhagic cerebrovascular stroke (1 placebo)."¹⁹¹</p>	<p>"A total of 121 patients had investigator-reported serious thrombotic events (77 in the rofecoxib group and 44 in the placebo group). A total of 46 patients in the rofecoxib group had confirmed (i.e., adjudicated) thrombotic events during 3059 patient-years of follow-up (1.50 events per 100 patient-years), and 26 patients in the placebo group had such events during 3327 patient-years of follow-up (0.78 event per 100 patient-years). As compared with the placebo group, the rofecoxib group had an increased risk of confirmed thrombotic events (relative risk, 1.92; 95 percent confidence interval, 1.19 to 3.11). The types of confirmed serious thrombotic events are shown in Table 2. The difference between the two groups was mainly due to an increased number of myocardial infarctions and strokes in the rofecoxib group. There were 10 deaths in each group. Myocardial infarction was the cause of death in two patients in the rofecoxib group and three in the placebo group, sudden death from cardiac causes occurred in three patients in the rofecoxib group and one in the placebo group, ischemic stroke was the cause of death in one patient in the rofecoxib group, and hemorrhagic stroke was the cause of death in one patient in the placebo group.</p> <p>In a post hoc analysis, the difference between the two groups in the incidence of thrombotic events was evident in the second 18 months of the study, whereas the event rates were similar for the first 18 months (Fig. 2 and Table 3). The changing pattern of the treatment effect over time was confirmed by a failed test for proportionality of hazards (P=0.01). Findings for the APTC endpoint were similar (Table 3)."¹⁹²</p>

¹⁹¹ APPROVe article draft, MRK-AFV0398909, at 19 (attached to 2/6/05 email from J. Baron* to R. Sandler* et al., MRK-AFV0398907-08).

Text of the draft submitted to the <u>New England Journal of Medicine</u> for peer review	Text of the published article (incorporating authors' post-peer review revisions)
"Discussion" section	
"The increased relative risk was observed beginning after approximately 18 months of treatment." ¹⁹³	"In post hoc analyses, the increased relative risk of adjudicated thrombotic events was first observed after approximately 18 months of treatment." ¹⁹⁴
"Differences between the present results and these earlier clinical trial data may be related to differences in exposure time, a possibility supported by the absence of an adverse effect treatment [sic] on cardiovascular outcomes during the first 18 months of APPROVe." ¹⁹⁵	"Differences between our results and these earlier clinical-trial data may be related to differences in defined end points or the duration of treatment, a possibility supported by the apparent absence of a difference in adjudicated thrombotic events during the first 18 months of our study." ¹⁹⁶

- (ii) Kaplan-Meier curves for congestive heart failure, pulmonary edema and cardiac failure.

As noted above, one of the reviewers had asked that the Kaplan-Meier plots for heart failure and pulmonary edema be included in the article, noting that, while "[m]uch

¹⁹² Bresalier* RS, Sandler* RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med. 2005;352:1092-102, at 096-97, MRK-ARQ0000659, at 663-64.

¹⁹³ APPROVe article draft, MRK-AFV0398909, at 22 (attached to 2/6/05 email from J. Baron* to R. Sandler* et al., MRK-AFV0398907-08).

¹⁹⁴ Bresalier* RS, Sandler* RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med. 2005;352:1092-102, at 098, MRK-ARQ0000659, at 665.

¹⁹⁵ APPROVe article draft, MRK-AFV0398909, at 24 (attached to 2/6/05 email from J. Baron* to R. Sandler* et al., MRK-AFV0398907-08).

¹⁹⁶ Bresalier* RS, Sandler* RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med. 2005;352:1092-102, at 099, MRK-ARQ0000659, at 666.

[had] been made of the lack of a safety signal on thrombotic events until 18 months,”
“the heart failure and pulmonary edema plots [showed] such a separation from the
beginning.”¹⁹⁷ In addition, Dr. Curfman* asked in his February 9, 2005 letter
summarizing peer reviewers’ comments that the authors “[s]how the [Kaplan-Meier]
curves for all serious CV events to demonstrate the early separation.”¹⁹⁸

After discussing with Dr. Drazen* what Dr. Curfman* meant by “all serious CV
events,” Dr. Baron* asked that the MRL co-authors prepare a Kaplan-Meier plot for a
wide composite endpoint including (i) cardiovascular events comprising the confirmed
thrombotic and APTC event endpoints; (ii) congestive heart failure; (iii) pulmonary
edema; and (iv) cardiac failure.¹⁹⁹ Dr. Quan had included such a Kaplan-Meier plot in
his September 13, 2004 report to the APPROVe Trial External Safety Monitoring Board,
and, in light of Dr. Baron’s* request, Mr. Bolognese asked Dr. Quan to prepare an
updated Kaplan-Meier plot based on the final December 2004 data.²⁰⁰

After learning of Dr. Baron’s* request, Dr. Reicin stated in an email to the MRL
co-authors and other Merck personnel that the Kaplan-Meier plot in Dr. Quan’s

¹⁹⁷ Reviewer D’s comments on APPROVe article draft, MRK-AFV0399550, at 60 (attached to 2/9/05
letter from G. Curfman* to J. Baron*, MRK-AFV0399548).

¹⁹⁸ 2/9/05 letter from G. Curfman* to J. Baron*, MRK-AFV0399548, at 48.

¹⁹⁹ 2/8/05 email from J. Baron* to J. Bolognese and H. Quan, MRK-ALE0045390; 2/8/05 email from
J. Bolognese to K. Horgan *et al.*, MRK-ALE0045390 (“[Dr. Baron*] mentioned the endpoint which
was in the ESMB report last Sept[ember] w/ CHF+edema and other things. I copied this from
[Dr. Quan’s] report. Adjudicated/confirmed thromboembolic or APTC serious AEs or (congestive
heart failure, pulmonary edema, or cardiac failure from CTS data base).”).

²⁰⁰ 2/8/05 email from J. Bolognese to K. Horgan *et al.*, MRK-ALE0045390; 9/13/04 report from H. Quan
to APPROVe ESMB, MRK-AGO0029517, at 60.

September 13, 2004 report was “a post hoc analysis” that “mix[ed] apples and oranges,” and that such mixing of events was “not appropriate for the paper.”²⁰¹ In an email sent the next day to the same group, Dr. Gertz agreed with Dr. Reicin: “I’m assuming we will not pool such data in a [Kaplan-Meier] plot for the NEJM in this completely post hoc analysis with events of very likely different pharmacologic origin.”²⁰²

On February 9, 2005, the authors sent to the editors Kaplan-Meier plots for edema-related and hypertension adverse events as well as a Kaplan-Meier plot for the composite endpoint of congestive heart failure, pulmonary edema and cardiac failure.²⁰³ After reviewing the Kaplan-Meier plots, Dr. Curfman* requested that the article include one Kaplan-Meier plot for the composite endpoint of congestive heart failure, pulmonary edema and cardiac failure, but did not ask for a Kaplan-Meier plot for the endpoint including the composite of (i) cardiovascular events comprising the confirmed thrombotic and APTC event endpoints; (ii) congestive heart failure; (iii) pulmonary edema; and (iv) cardiac failure.²⁰⁴

²⁰¹ 2/9/05 email from A. Reicin to N. Braunstein et al., MRK-ALE0045389 (also stating: “[Congestive heart failure] and [hypertension] are known side effects of the drug and we already acknowledge these in the paper (as well as our label). If [the journal] want[s] a [Kaplan-Meier plot] of [hypertension or congestive heart failure] that is fine but mixing all together is inappropriate for this paper which focuses on CV thrombotic safety.”).

²⁰² 2/9/05 email from B. Gertz to C. Lines et al., MRK-ALE0045288; see also 2/9/05 email from B. Gertz to A. Reicin et al., MRK-ALE0045289 (“We need to speak first to [Dr. Baron*] and not just provide inappropriate analyses.”).

²⁰³ 2/10/05 email from C. Lines to A. Reicin, MRK-AFV0399928.

²⁰⁴ 2/10/05 email from C. Lines to A. Reicin, MRK-AFV0399928 (“The editor’s letter is confusing, but [Dr. Baron*] determined that what they want incorporated is the 3 K-M plots that we sent yesterday for 1) edema, 2) hypertension, 3) congestive heart failure + pulmonary edema + cardiac failure.”); Bresalier* RS, Sandler* RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a

As a result of these discussions, the authors made three revisions in response to the peer reviewers' comments regarding congestive heart failure, pulmonary edema and cardiac failure. First, they included in the article a Kaplan-Meier plot for a composite endpoint of congestive heart failure, pulmonary edema and cardiac failure (Figure 3 of the published article, reproduced in full in Section B.3.h of this Appendix), not the broader composite endpoint originally mentioned by Dr. Baron*.²⁰⁵ Second, they added to the article a table reporting the rate and hazard ratio results for hypertension, edema, and congestive heart failure, pulmonary edema, and cardiac failure.²⁰⁶ Third, they revised the "Abstract" section of the article, which previously did not mention non-adjudicated investigator-reported congestive heart failure, pulmonary edema, and cardiac failure, to state that "[t]here was earlier separation (at approximately five months) between groups in the incidence of [such events]."²⁰⁷ Finally, the authors revised a paragraph in the "Discussion" section of the article as follows"

colorectal adenoma chemoprevention trial. N Engl J Med. 2005;352:1092-102, at 097, MRK-ARQ0000659, at 664.

²⁰⁵ Bresalier* RS, Sandler* RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med. 2005;352:1092-102, at 097, MRK-ARQ0000659, at 664; see also Authors' responses to reviewers' comments on APPROVe article draft, MRK-AHD0000061, at 61 ("[W]e have supplied an event curve for the most serious and homogeneous of the endpoints, congestive failure [sic]/pulmonary edema/cardiac failure.").

²⁰⁶ 2/11/05 email from J. Baron* to K. Horgan et al., MRK-AGO0103976 ("I . . . added a table for the three . . . events we included."); see also Bresalier* RS, Sandler* RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med. 2005;352:1092-102, at 099, MRK-ARQ0000659, at 666 (Table 4).

²⁰⁷ Bresalier* RS, Sandler* RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med. 2005;352:1092-102, at 092, MRK-ARQ0000659, at 659.

Text of the draft submitted to the <u>New England Journal of Medicine</u> for peer review	Text of the published article (incorporating authors' post-peer review revisions)
<p>“The increase in thrombotic cardiovascular events associated with rofecoxib was not evident during the first 18 months of the trial. This delayed emergence does not suggest a pro-thrombotic effect, which might be expected to be manifest early upon treatment initiation.”²⁰⁸</p>	<p>“The increase in adjudicated thrombotic events associated with rofecoxib therapy was not evident during the first 18 months of the trial. Other investigator-reported cardiovascular events known to be associated with NSAID use, such as congestive heart failure and pulmonary edema, although less well defined, occurred earlier (at approximately five months) and at a higher rate among patients taking rofecoxib than among those taking placebo.”²⁰⁹</p>

iii. Investigator-reported cardiovascular thromboembolic events.

As discussed above, the manuscript submitted on February 6, 2005 reported results for three composite cardiovascular event endpoints: (i) investigator-reported cardiovascular thrombotic events; (ii) confirmed thrombotic events; and (iii) APTC composite cardiovascular event endpoint.²¹⁰ The manuscript reported the total number of

²⁰⁸ APPROVe article draft, MRK-AFV0398909, at 25 (attached to 2/6/05 email from J. Baron* to R. Sandler* et al., MRK-AFV0398907-08).

²⁰⁹ Bresalier* RS, Sandler* RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med. 2005;352:1092-102, at 100, MRK-ARQ0000659, at 667.

²¹⁰ As with all trials of Vioxx of four weeks of longer, investigator-reported cardiovascular thrombotic events from the APPROVe Trial were adjudicated pursuant to Merck’s Cardiovascular Adjudication SOP. As discussed in Appendix A, an independent committee, which consisted of external cardiologists, neurologists, and vascular medicine internists, adjudicated investigator-reported cardiovascular events – that is, decided, based on its evaluation of the underlying medical documentation, whether the event met the pre-specified criteria for a thromboembolic event, such as a myocardial infarction or a stroke. If an event was adjudicated to be a thromboembolic event, then the event was classified as a “confirmed” thromboembolic event and included in the relevant composite endpoint (e.g., (i) confirmed thrombotic event endpoint as defined by MRL scientists in the Cardiovascular Adjudication SOP, and/or (ii) the widely used APTC endpoint). The difference between the two endpoints is discussed in detail in Appendix F and in Section B.3.c of this Appendix.

events for each of the three endpoints, but reported the rate and relative risk analyses only for the two confirmed event endpoints (*i.e.*, the confirmed thrombotic and APTC composite cardiovascular event endpoints).²¹¹ Pre-submission drafts of the manuscript had included the rate and relative risk analyses for the investigator-reported event endpoint as well,²¹² but, after some discussion, the authors decided not to include such analyses on the theory that the confirmed event data constituted more reliable and therefore more relevant evidence of the cardiovascular profile of Vioxx.²¹³ Table 13

²¹¹ APPROVe article draft, MRK-AFV0398909, at 19, Table 2 at 39 (attached to 2/6/05 email from J. Baron* to R. Sandler* *et al.*, MRK-AFV0398907-08).

²¹² APPROVe article draft, MRK-AGO0072373, at 392 (attached to 11/9/04 email from J. Bolognese to C. Lines *et al.*, MRK-AGO0072372); *see also* APPROVe article draft, MRK-AHC0027633, at 653 (attached to 11/14/04 email from R. Sandler* to C. Lines and K. Horgan, MRK-AHC0027631); APPROVe article draft, MRK-AGO0075101, at 122 (attached to 11/18/04 email from H. Quan to C. Lines *et al.*, MRK-AGO0075099).

²¹³ APPROVe article draft, MRK-AHC0024450, at 457 (attached to 11/15/04 email from J. Baron* to K. Horgan *et al.*, MRK-AHC0024449) (Dr. Baron* suggesting that the article not include any analyses of investigator-reported events, reasoning: “Let’s not do this. It takes up space and is dominated by the adjudicated analysis anyway.”); APPROVe article draft with M. Konstam’s* comments, MRK-AHC0024647, at 655 (“I think that the principal focus of the results should be on the event rates based on the pre-specified adjudication process. We can consider including the investigator-reported events as a secondary analysis, but I’m not sure that’s necessary.”) (attached to 12/1/04 email from J. Baron* to R. Sandler* *et al.*, MRK-AHC0024646); 1/17/05 email from C. Lines to K. Horgan, MRK-AHD0074962 (“I’m beginning to think we should get rid of the investigator events.”); APPROVe article draft with N. Braunstein’s comments, MRK-AAD0411667, at 675, 678 (attached to 1/21/05 email from N. Braunstein to C. Lines *et al.*, MRK-AAD0411664) (Dr. Braunstein noted that he “wouldn’t do or report any analyses based on investigator-reported [thromboembolic] events” “other than perhaps to report their number” because “the entire rationale for the [Cardiovascular Adjudication SOP] was that analyses would be based on confirmed events.”); 1/26/05 email from C. Lines to K. Horgan *et al.*, MRK-AHD0000392 (“Issues raised by a number of people that we need to agree upon before discussing with externals are: 1) Should we recommend deletion of investigator reported events, other than just brief mention of overall numbers at beginning of Results section?”); 1/31/05 email from C. Lines to R. Bresalier* *et al.*, MRK-AAD0421787 (attaching APPROVe article draft, MRK-AAD0421789) (emailing to external authors a revised draft with “suggestions,” including “[d]eletion of investigator-reported events (except for 1 sentence describing the totals) – improves clarity.”).

below is the version of Table 2, including investigator-reported data, that appeared in the January 28, 2005 draft. The shaded rows subsequently were deleted pre-submission.

Table 13

“Table 2. Summary of rates and relative risks for confirmed thrombotic cardiovascular serious adverse events and APTC events” from 1/28/05 draft of the APPROVe article.²¹⁴

	Rofecoxib 25 mg				Placebo				Comparison	
	N*	Events	Patient-years at risk	Rate per 100 patient-years	N*	Events	Patient-years at risk	Rate per 100 patient-years	Rate Difference [95% CI]	Relative risk [95% CI]
<i>Confirmed</i>										
Confirmed	1287	46	3059	1.50	1299	26	3327	0.78	0.72 [0.19, 1.25]	1.92 [1.19, 3.11]
Month 0-18	1287	22	1656	1.33	1299	20	1765	1.13	0.20 [-0.55, 0.94]	1.18 [0.64, 2.15]
Month 19-36	989	24	1403	1.71	1079	6	1561	0.38	1.33 [0.58, 2.08]	4.45 [1.77, 13.32]
<i>APTC</i>										
APTC	1287	34	3070	1.11	1299	18	3334	0.54	0.57 [0.12, 1.02]	2.06 [1.16, 3.64]
Month 0-18	1287	14	1658	0.84	1299	12	1769	0.68	0.17 [-0.42, 0.75]	1.25 [0.58, 2.69]
Month 19-36	994	20	1412	1.42	1083	6	1565	0.38	1.03 [0.34, 1.73]	3.69 [1.43, 11.24]
<i>Investigator-reported</i>										
Overall	1287	77	3035	2.54	1299	44	3312	1.33	1.21 [0.52, 1.90]	1.90 [1.31, 2.76]
Month 0-18	1287	39	1651	2.36	1299	25	1763	1.42	0.94 [0.02, 1.87]	1.67 [1.01, 2.76]
Month 19-36	982	38	1384	2.74	1076	19	1549	1.23	1.52 [0.49, 2.55]	2.22 [1.28, 3.85]
* Number of patients at risk										

²¹⁴ APPROVe article draft, MRK-AGO0076264, at 297-98 (attached to 1/28/05 email from H. Quan to C. Lines, MRK-AGO0076263).

The person who deleted investigator-reported events from the manuscript prior to submission, however, inadvertently left behind a note stating: “SUGGESTION IS TO DELETE INVESTIGATOR-REPORTED EVENTS [from Table 2].”²¹⁵ In reviewing the draft manuscript, the New England Journal of Medicine peer reviewers noticed that note and asked that the analyses of investigator-reported events be restored.²¹⁶ One of the reviewers also commented: “There is a worrying note to delete some reports from table 2 that the authors ought to explain – we were obviously not supposed to see that!”²¹⁷

In the document responding to each of the peer reviewers’ comments, the co-authors explained:

We will be happy to [include analyses of investigator-reported events if] necessary, but believe it would detract from the paper to present numbers of candidate events, including those deemed by a blinded adjudication committee to NOT be “real.” The investigator reported thrombotic events ARE included in the text, however.²¹⁸

* * *

After extensive discussion the authors chose to omit non-adjudicated investigator reported events from the table (they are included in the text). This was a choice based on

²¹⁵ APPROVe article draft, MRK-AFV0398909, at 39 (attached to 2/6/05 email from J. Baron* to R. Sandler* et al., MRK-AFV0398907-08). In addition, although Table 2 no longer included analyses of investigator-reported events, the text of the submitted draft still referred the reader to Table 2 for data on such events. Id. at 19.

²¹⁶ 2/9/05 letter from G. Curfman* to J. Baron*, MRK-AFV0399548, at 48 (highlighting several reviewers’ comments, including request that authors “[r]estore the investigator-reported events category to Table 2”).

²¹⁷ Reviewer D’s comments on APPROVe article draft, MRK-AFV0399560, at 61 (attached to 2/9/05 letter from G. Curfman* to J. Baron*, MRK-AFV0399548).

²¹⁸ Authors’ responses to reviewers’ comments on APPROVe article draft, MRK-AHD0000061, at 61.

the authors' decision about weight of evidence, and it is
inappropriate to infer anything else.²¹⁹

The co-authors did not reinstate and the published article did not include the analyses of
investigator-reported events that had been removed from the draft manuscript.

iv. Peer reviewers' other comments on
the submitted manuscript.

The reviewers also commented on the overall tone and quality of the article. One
reviewer noted that the manuscript was "excellent"²²⁰ and another that it was "clearly
written."²²¹ Several reviewers, however, asked that the authors change the overall tone of
the article. For example, one reviewer wrote:

The "hand" of the study sponsor seems too evident
throughout the manuscript, which is written consistently in
a fashion designed to support the company's public
positions. In the process, the scientific integrity of the
manuscript is compromised.

* * *

The manuscript really needs a complete re-working to
avoid the current "spin."²²²

²¹⁹ Authors' responses to reviewers' comments on APPROVe article draft, MRK-AHD0000061, at 76.

²²⁰ Reviewer A's comments on APPROVe article draft, MRK-AFV0399550, at 550 (attached to 2/9/05
letter from G. Curfman* to J. Baron*, MRK-AFV0399548).

²²¹ Reviewer D's comments on APPROVe article draft, MRK-AFV0399560, at 560 (attached to 2/9/05
letter from G. Curfman* to J. Baron*, MRK-AFV0399548).

²²² Reviewer E's comments on APPROVe article draft, MRK-AFV0399562, at 62 (attached to 2/9/05
letter from G. Curfman* to J. Baron*, MRK-AFV0399548); see also Reviewer C's comments on
APPROVe article draft, MRK-AFV0399556, at 56 (attached to 2/9/05 letter from G. Curfman* to
J. Baron*, MRK-AFV0399548) ("[M]uch of the Introduction and Discussion of the present
[manuscript] is devoted to a selective review of previous studies that seems to have as a goal
defending the manufacturer's conduct in the entire rofecoxib affair.").

As discussed above, after receiving peer reviewers' comments, Dr. Bresalier* and Dr. Baron* (with input from other co-authors) made extensive revisions to the manuscript. In comparison to the draft submitted on February 6, 2005, the revised manuscript, among other things, placed less emphasis on the results of the post-hoc analysis suggesting that the increased risk on Vioxx did not emerge until 18 months of continuous use and focused more on the early separation of the Kaplan-Meier curves for Vioxx and placebo for congestive heart failure, pulmonary edema and cardiac failure. In their response to the peer reviewers' comments, the authors also stated: "The manuscript has been extensively edited. . . . We are confident the manuscript fairly describes the literature and the findings of APPROVe."²²³

h. The published article.

The article was posted on the website of the New England Journal of Medicine on February 15, 2005 and published in its March 17, 2005 edition. The following is a summary of the article (which is reproduced in full below) with regard to the issues discussed above:

- The article in several places, including the Abstract section, noted that the event rates in the Vioxx and placebo groups were similar during the first 18 months of the study and that the increased relative risk became apparent after 18 months of treatment.²²⁴

²²³ Authors' responses to reviewers' comments on APPROVe article draft, MRK-AHD0000061, at 70, 78.

²²⁴ Bresalier* RS, Sandler* RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med. 2005;352:1092-102, at 1092, 1097, 1100.

- The article stated that the 18-month analysis (*i.e.*, the analysis comparing the relative risk results between the first and second 18-month intervals of the APPROVe Trial) was a post-hoc analysis and that “[a] test of the proportional-hazards assumption was specified in the cardiovascular-analysis plan.”²²⁵
- The table reporting cardiovascular results of the study reported data for the entire duration of the trial as well as for its first and second 18-month intervals for the confirmed thrombotic and APTC composite cardiovascular endpoints.²²⁶
- The article stated that the proportional hazards assumption was tested using the Logarithm of Time Test and that the test yielded a statistically significant result ($p=0.01$) for the confirmed thrombotic event endpoint.²²⁷
- The article also stated that “[f]indings” for the confirmed thrombotic and APTC composite cardiovascular endpoints were “similar” and did not state the proportionality p -value ($p=0.119$) obtained for the APTC endpoint:

“In a post hoc analysis, the difference between the two groups in the incidence of thrombotic events was evident in the second 18 months of the study, whereas the event rates were similar for the first 18 months (Fig. 2 and Table 3). The changing pattern of the treatment effect over time was confirmed by a failed test for proportionality of hazards ($P=0.01$). Findings for the APTC end point were similar (Table 3).”²²⁸

²²⁵ Bresalier* RS, Sandler* RS, Quan H, *et al.* Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. *N Engl J Med.* 2005;352:1092-102, at 1097, 1095.

²²⁶ Bresalier* RS, Sandler* RS, Quan H, *et al.* Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. *N Engl J Med.* 2005;352:1092-102, at 1098.

²²⁷ Bresalier* RS, Sandler* RS, Quan H, *et al.* Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. *N Engl J Med.* 2005;352:1092-102, at 1095, 1097.

²²⁸ Bresalier* RS, Sandler* RS, Quan H, *et al.* Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. *N Engl J Med.* 2005;352:1092-102, at 1097.

- The article included Kaplan-Meier plots for confirmed thrombotic events as well as for congestive heart failure, pulmonary edema and cardiac failure. Furthermore, the Abstract and Discussion sections mentioned the early separation of the Kaplan-Meier curves for congestive heart failure, pulmonary edema and cardiac failure.²²⁹
- The article reported total event numbers for investigator-reported events but did not include analyses of investigator-reported events.²³⁰
- The article did not mention the observation that the Kaplan-Meier curve for placebo for confirmed thrombotic events turned relatively flat after the 18-month mark.

²²⁹ Bresalier* RS, Sandler* RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med. 2005;352:1092-102, at 1092, 1097.

²³⁰ Bresalier* RS, Sandler* RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med. 2005;352:1092-102, at 1096.

ORIGINAL ARTICLE

Cardiovascular Events Associated with Rofecoxib in a Colorectal Adenoma Chemoprevention Trial

Robert S. Bresalier, M.D., Robert S. Sandler, M.D., Hui Quan, Ph.D.,
James A. Bolognese, M.Stat., Bettina Oxenius, M.D., Kevin Horgan, M.D.,
Christopher Lines, Ph.D., Robert Riddell, M.D., Dion Morton, M.D.,
Angel Lanas, M.D., Marvin A. Konstam, M.D., and John A. Baron, M.D.,
for the Adenomatous Polyp Prevention on Vioxx (APPROVe) Trial Investigators*

ABSTRACT

BACKGROUND

From the Department of Gastrointestinal Medicine and Nutrition, University of Texas M.D. Anderson Cancer Center, Houston (R.S.B.); the Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill (R.S.S.); Merck Research Laboratories, West Point, Pa. (H.Q., J.A. Bolognese, B.O., K.H., C.L.); the Department of Pathology, Mount Sinai Hospital, Toronto (R.R.); the Department of Surgery, University of Birmingham, Birmingham, United Kingdom (D.M.); the Department of Medicine, Clinic University Hospital, Zaragoza, Spain (A.L.); the Department of Medicine, Tufts–New England Medical Center, Boston (M.A.K.); and the Departments of Medicine and Community and Family Medicine, Dartmouth Medical School, Hanover, N.H. (J.A. Baron). Address reprint requests to Dr. Bresalier at the Department of Gastrointestinal Medicine and Nutrition, University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd., Houston, TX 77030-4009, or at rbresali@mdanderson.org.

*The members of the APPROVe Trial are listed in the Appendix.

This article was published at www.nejm.org on February 15, 2005.

N Engl J Med 2005;352:1092-102.
Copyright © 2005 Massachusetts Medical Society.

Selective inhibition of cyclooxygenase-2 (COX-2) may be associated with an increased risk of thrombotic events, but only limited long-term data have been available for analysis. We report on the cardiovascular outcomes associated with the use of the selective COX-2 inhibitor rofecoxib in a long-term, multicenter, randomized, placebo-controlled, double-blind trial designed to determine the effect of three years of treatment with rofecoxib on the risk of recurrent neoplastic polyps of the large bowel in patients with a history of colorectal adenomas.

METHODS

A total of 2586 patients with a history of colorectal adenomas underwent randomization: 1287 were assigned to receive 25 mg of rofecoxib daily, and 1299 to receive placebo. All investigator-reported serious adverse events that represented potential thrombotic cardiovascular events were adjudicated in a blinded fashion by an external committee.

RESULTS

A total of 46 patients in the rofecoxib group had a confirmed thrombotic event during 3059 patient-years of follow-up (1.50 events per 100 patient-years), as compared with 26 patients in the placebo group during 3327 patient-years of follow-up (0.78 event per 100 patient-years); the corresponding relative risk was 1.92 (95 percent confidence interval, 1.19 to 3.11; $P=0.008$). The increased relative risk became apparent after 18 months of treatment; during the first 18 months, the event rates were similar in the two groups. The results primarily reflect a greater number of myocardial infarctions and ischemic cerebrovascular events in the rofecoxib group. There was earlier separation (at approximately five months) between groups in the incidence of nonadjudicated investigator-reported congestive heart failure, pulmonary edema, or cardiac failure (hazard ratio for the comparison of the rofecoxib group with the placebo group, 4.61; 95 percent confidence interval, 1.50 to 18.83). Overall and cardiovascular mortality was similar in the two groups.

CONCLUSIONS

Among patients with a history of colorectal adenomas, the use of rofecoxib was associated with an increased cardiovascular risk.

CARDIOVASCULAR EVENTS ASSOCIATED WITH ROFECOXIB

NONSTEROIDAL ANTIINFLAMMATORY drugs (NSAIDs) alleviate pain and inflammation but may cause gastrointestinal ulceration and bleeding, presumably by inhibiting cyclooxygenase (COX)-mediated production of prostaglandins. The discovery that there were two forms of cyclooxygenase, 1 (COX-1) and 2 (COX-2), provided the impetus for the development of selective inhibitors of COX-2 with a reduced risk of gastrointestinal complications whose analgesic and antiinflammatory efficacy was likely to be similar to that of nonselective COX inhibitors.

COX-2 is expressed at sites of inflammation, such as in atheromatous plaques, and in neoplasms, raising the possibility that COX-2 inhibition might also be useful in the treatment or prevention of atherosclerosis and various cancers.^{1,2} However, predicting the consequences of COX-2 inhibition on cardiovascular disease is not a straightforward proposition. COX-2 inhibition has several effects that could increase the risk of cardiovascular disease, including reducing prostacyclin levels, increasing blood pressure, decreasing angiogenesis,³⁻¹¹ and destabilizing plaque.¹²

Rofecoxib is a selective COX-2 inhibitor that has been shown to be associated with significantly fewer gastrointestinal adverse events than NSAIDs.¹³ In one trial,¹³ there were more cardiovascular events among patients given a high dose of rofecoxib than among those given naproxen, an NSAID with platelet-inhibiting properties of unclear clinical relevance.^{4,5,14-16} Pooled data from other randomized trials have not shown a significant difference in cardiovascular risk between rofecoxib and placebo or other nonselective NSAIDs.^{4,5,17} Observational studies have provided conflicting data on the association of rofecoxib with cardiovascular risk: some studies suggested that there was no effect, some suggested that the risk was increased only at high doses, and others indicated a possible increase in the risk of cardiovascular events at standard or unspecified doses.^{6,9,11,18-21}

The Adenomatous Polyp Prevention on Vioxx (APPROVE) Trial was designed to evaluate the hypothesis that three years of treatment with rofecoxib would reduce the risk of recurrent adenomatous polyps among patients with a history of colorectal adenomas. Potential thrombotic events were adjudicated by an independent committee, and all safety data were monitored by an external safety-monitoring committee. We report the cardiovascular findings from the study.

METHODS

DESIGN OF THE TRIAL

Enrollment occurred from February 2000 to November 2001 at 108 centers in 29 countries. Participating investigators are listed in the Appendix. Men and women who were at least 40 years old were eligible if they had had at least one histologically confirmed large-bowel adenoma removed within 12 weeks before study entry and were not anticipated to need long-term NSAID therapy (including high-dose aspirin) during the study. Initially, patients who were taking low-dose aspirin (no more than 100 mg daily) were excluded from the study. However, in May 2000, after the results of the Vioxx Gastrointestinal Research (VIGOR) trial¹³ had become available, the protocol was amended to allow randomized subjects to take low-dose aspirin (no more than 100 mg daily) for cardiovascular protection. The proportion of subjects taking low-dose aspirin at enrollment was capped at 20 percent because of the possible chemopreventive effects of the drug.²² Exclusion criteria were evidence of uncontrolled hypertension (defined by a blood pressure of more than 165/95 mm Hg); angina or congestive heart failure, with symptoms evoked by minimal activity; myocardial infarction, coronary angioplasty, or coronary-artery bypass grafting within the preceding year; or stroke or transient ischemic attack within two years before screening.

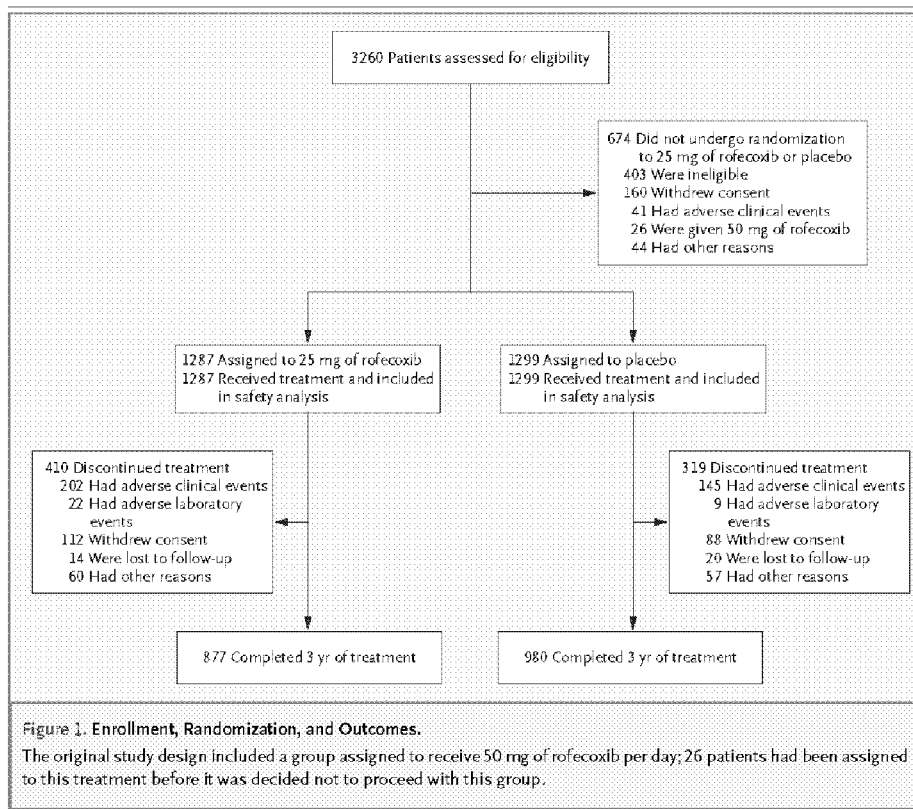
Written informed consent was obtained from all patients. The institutional review board at each center approved the study.

TREATMENT

The randomized treatment period was preceded by a six-week, single-blind, placebo run-in period to assess patients' compliance. Patients who took at least 80 percent of their tablets during the placebo run-in period were randomly assigned to receive either one 25-mg tablet of rofecoxib per day (the maximal recommended long-term daily dose) or one identical-appearing placebo tablet per day for three years. The computer-derived randomization was stratified according to the clinical center and the use or nonuse of low-dose aspirin, with blocks of 2. Patients, investigators, and sponsor personnel who monitored the study, other than the unblinded study statistician, were unaware of the treatment assignments.

Patients were evaluated clinically at randomization and at weeks 4, 17, 35, 52, 69, 86, 104, 121,

The NEW ENGLAND JOURNAL of MEDICINE



138, 156, and 158 or after the discontinuation of treatment. Vital signs, including blood pressure obtained while the patient was seated, were measured at each clinic visit during the study according to the usual clinical practice. Adverse events occurring during the study were recorded and evaluated in a blinded fashion by the investigators. Follow-up of the patients for one year after the discontinuation of treatment is ongoing.

CARDIOVASCULAR EVENTS

Monitoring and analysis of the cardiovascular events in the trial were part of a planned assessment of the cardiovascular safety of rofecoxib. Data presented include events occurring during treatment and up to 14 days after the last dose of the study drug. Serious vascular events were reviewed in a blinded fashion by adjudication committees, which confirmed events that met prespecified case definitions for two sets of events. Thrombotic events included fatal and

nonfatal myocardial infarction, unstable angina, sudden death from cardiac causes, fatal and nonfatal ischemic stroke, transient ischemic attack, peripheral arterial thrombosis, peripheral venous thrombosis, and pulmonary embolism. The end point used in the Antiplatelet Trialists' Collaboration (APTC) study²³ was also analyzed: the combined incidence of death from cardiovascular, hemorrhagic, and unknown causes; nonfatal myocardial infarction; and nonfatal ischemic and hemorrhagic stroke. Other relevant but not independently adjudicated events were also analyzed, including hypertension-related events, edema-related events, and the combined end point of congestive heart failure, pulmonary edema, or cardiac failure.

The procedure for confirming cardiovascular events was prespecified in the protocol. All serious adverse events were identified and recorded by the clinical investigators. Potential thromboembolic events, components of the APTC end point, and all

CARDIOVASCULAR EVENTS ASSOCIATED WITH ROFECOXIB

deaths (regardless of cause) were prespecified as eligible for adjudication according to standard procedures for rofecoxib studies initiated by the sponsor in 1998. For each eligible event, source documents were collected and sent to the cardiac, cerebrovascular, or peripheral vascular adjudication committee. Decisions were made on the basis of majority rule with the use of prespecified criteria.

STATISTICAL ANALYSIS

An independent, external safety-monitoring board met periodically to review safety data provided by a statistician who was aware of patients' study-group assignments. No formal stopping rule was specified for terminating the study.

Data were collected and held by the sponsor. The investigators had full and unfettered access to the data. A statistician who was aware of patients' study-group assignments but who was not otherwise involved in the study analyzed the data using SAS software (version 8.2). All patients who underwent randomization and took at least one dose of study medication were included in the analyses. For confirmed serious thrombotic events and the APTC end point, event rates were determined and relative risks (with 95 percent confidence intervals) were calculated with the use of Cox proportional-hazards models. However, if there were fewer than 11 events in either group, the rate ratio was computed with the use of the binomial distribution.²⁴ A test of the proportional-hazards assumption was specified in the cardiovascular-analysis plan. This was accomplished by evaluating the interaction between the logarithm of time and the assigned treatment in the Cox proportional-hazards model. Kaplan-Meier estimates of the cumulative event rates over time were also made.

Several exploratory analyses were performed to delineate the relation between mean arterial pressure and the study findings. One analysis summarized the relative risk of confirmed serious thrombotic adverse events according to the quartiles of change in mean arterial pressure at week 4. This time was chosen because treatment-based differences in mean arterial pressure occurred early and remained constant throughout the treatment period and because only two confirmed serious thrombotic events had occurred by week 4 (one in each group). The second analysis included changes from baseline in mean arterial pressure as a time-varying covariate in a Cox proportional-hazards model in which treatment was the main effect. This model

was used to investigate the association of the change in blood pressure over time with the occurrence of confirmed serious thrombotic events.

The data reported here are those available to the authors as of February 14, 2005.

RESULTS

PARTICIPANTS

A total of 3260 patients were screened for the study, of whom 2586 were deemed to be eligible; 1287 of the eligible patients were randomly assigned to receive rofecoxib, and 1299 to receive placebo (Fig. 1). The two groups were generally similar with regard to baseline characteristics, including age, sex, use or nonuse of low-dose aspirin, and cardiovascular-risk status (Table 1). Concomitant medications used at some time during the study included low-dose aspirin (in 20 percent of the rofecoxib group and

Table 1. Baseline Characteristics of the Patients.

Characteristic	Rofecoxib (N=1287)	Placebo (N=1299)
Age (yr)		
Mean	59	59
Range	40-96	40-86
Height (cm)		
Mean	170	170
Range	137-198	133-199
Weight (kg)		
Mean	81	81
Range	38-160	34-159
Male sex (%)	62	62
White race (%)*	84	84
Use of low-dose aspirin (%) †	17	16
Use of antihypertensive medication (%)	30	29
High cardiovascular risk (%) ‡	30	26
History of symptomatic atherosclerotic cardiovascular disease (%)	9	8
History of hypertension (%)	36	34
History of hypercholesterolemia (%)	29	26
History of diabetes (%)	9	9
Current cigarette use (%)	22	22

* Race was self-reported.

† Low-dose aspirin was defined as 100 mg per day or less.

‡ A high cardiovascular risk was defined by a history of symptomatic atherosclerotic cardiovascular disease or the presence of at least two of the following risk factors for cardiovascular disease: history of hypertension, history of hypercholesterolemia, history of diabetes, or current cigarette use.

The NEW ENGLAND JOURNAL of MEDICINE

Table 2. Incidence of Adjudicated Thrombotic Adverse Events.*

Adverse Event	Rofecoxib Group (N=1287)		Placebo Group (N=1299)		Hazard Ratio (95% CI)
	No. of Patients (%)	Rate/100 Patient-yr	No. of Patients (%)	Rate/100 Patient-yr	
Total	46 (3.6)	1.50	26 (2.0)	0.78	1.92 (1.19–3.11)
Cardiac events	31 (2.4)	1.01	12 (0.9)	0.36	2.80 (1.44–5.45)
Myocardial infarction	21		9		
Fatal myocardial infarction	2		3		
Sudden death from cardiac causes	3		1		
Unstable angina pectoris	7		4		
Cerebrovascular events	15 (1.2)	0.49	7 (0.5)	0.21	2.32 (0.89–6.74)
Fatal ischemic stroke	1		0		
Ischemic stroke	11		6		
Transient ischemic attack	5		2		
Peripheral vascular events	3 (0.2)	0.10	7 (0.5)	0.21	0.46 (0.08–2.03)
Peripheral arterial thrombosis	1		1		
Peripheral venous thrombosis	2		4		
Pulmonary embolism	0		2		

* The total duration of follow-up was 3059 patient-years in the rofecoxib group and 3327 patient-years in the placebo group. Although a patient may have had two or more clinical adverse events, the patient was counted once within a category. The same patient may appear in different categories. CI denotes confidence interval.

19 percent of the placebo group, $P=0.52$), antihypertensive drugs (44 percent and 36 percent, respectively; $P<0.001$), lipid-lowering agents (31 percent and 28 percent, respectively; $P=0.09$), antiplatelet agents such as clopidogrel (4 percent and 2 percent, respectively; $P=0.003$), insulin (3 percent and 2 percent, respectively; $P=0.10$), and oral hypoglycemic agents (13 percent and 11 percent, respectively; $P=0.12$).

The study was terminated on September 30, 2004, approximately two months ahead of the planned date of completion, at the recommendation of the external safety-monitoring board and the steering committee. At the time of termination, a total of 877 patients in the rofecoxib group and 980 patients in the placebo group had completed the scheduled three years of treatment. The mean duration of treatment was 2.4 years in the rofecoxib group and 2.6 years in the placebo group.

Before September 30, 2004, more patients discontinued rofecoxib treatment than placebo (32 percent vs. 25 percent) (Fig. 1). The main reason for discontinuation was an adverse clinical event. The three most common adverse events resulting in the discontinuation of treatment were hypertension (25 patients in the rofecoxib group and 7 patients in

the placebo group), increased blood pressure (6 in the rofecoxib group and 1 in the placebo group), and peripheral edema (7 in the rofecoxib group and 1 in the placebo group).

INCIDENCE OF THROMBOTIC EVENTS AND THE APTC END POINT

A total of 121 patients had investigator-reported serious thrombotic events (77 in the rofecoxib group and 44 in the placebo group). A total of 46 patients in the rofecoxib group had confirmed (i.e., adjudicated) thrombotic events during 3059 patient-years of follow-up (1.50 events per 100 patient-years), and 26 patients in the placebo group had such events during 3327 patient-years of follow-up (0.78 event per 100 patient-years). As compared with the placebo group, the rofecoxib group had an increased risk of confirmed thrombotic events (relative risk, 1.92; 95 percent confidence interval, 1.19 to 3.11). The types of confirmed serious thrombotic events are shown in Table 2. The difference between the two groups was mainly due to an increased number of myocardial infarctions and strokes in the rofecoxib group. There were 10 deaths in each group. Myocardial infarction was the cause of death in two patients in the rofecoxib group and three in the placebo

CARDIOVASCULAR EVENTS ASSOCIATED WITH ROFECOXIB

group, sudden death from cardiac causes occurred in three patients in the rofecoxib group and one in the placebo group, ischemic stroke was the cause of death in one patient in the rofecoxib group, and hemorrhagic stroke was the cause of death in one patient in the placebo group.

In a post hoc analysis, the difference between the two groups in the incidence of thrombotic events was evident in the second 18 months of the study, whereas the event rates were similar for the first 18 months (Fig. 2 and Table 3). The changing pattern of the treatment effect over time was confirmed by a failed test for proportionality of hazards ($P=0.01$). Findings for the APTC end point were similar (Table 3).

There were no significant interactions between treatment group and subgroups ($P>0.10$ for all comparisons) for confirmed serious thrombotic events in subgroup analyses based on country (United States vs. other); age; sex; use or nonuse of antihypertensive drugs at baseline, low-dose aspirin at baseline, or low-dose aspirin for more than 50 percent of follow-up; presence or absence of a history of hypertension, hypercholesterolemia, or ischemic heart disease; presence or absence of current cigarette use; or presence or absence of a high cardiovascular risk. A high cardiovascular risk was defined by a history of symptomatic atherosclerotic cardiovascular disease or the presence of at least two of the following risk factors for coronary artery disease: a history of hypertension, a history of hypercholesterolemia, a history of diabetes, or current cigarette use. However, point estimates for the relative risk in the rofecoxib group as compared with the placebo group were particularly high among patients with a history of symptomatic atherosclerotic cardiovascular disease (9.59; 95 percent confidence interval, 1.36 to 416) relative to those without such a clinical history (1.58; 95 percent confidence interval, 0.95 to 2.64; P for interaction=0.096). Also, the relative risk in the rofecoxib group as compared with the placebo group was 6.10 among patients with a history of diabetes (95 percent confidence interval, 1.36 to 56.1), in contrast to a relative risk of 1.55 among patients with no history of diabetes (95 percent confidence interval, 0.92 to 2.61; P for interaction=0.091).

NONADJUDICATED CARDIOVASCULAR EVENTS

As compared with the placebo group, the rofecoxib group had higher percentages of patients with hypertension-related events and edema-related events.

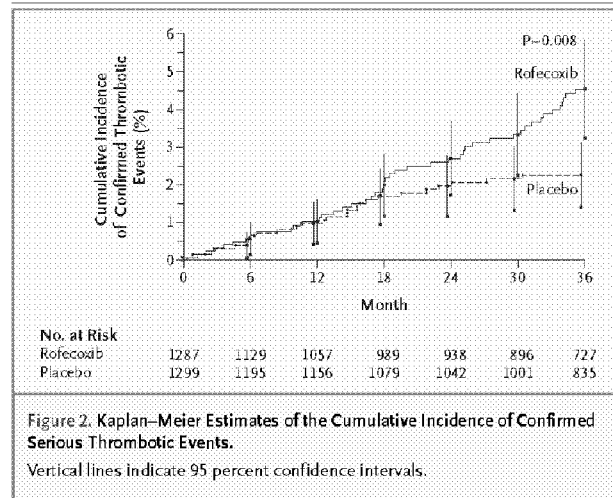


Figure 2. Kaplan-Meier Estimates of the Cumulative Incidence of Confirmed Serious Thrombotic Events. Vertical lines indicate 95 percent confidence intervals.

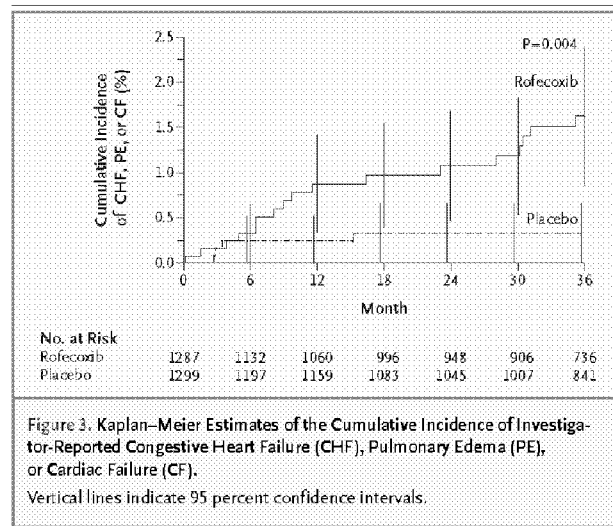


Figure 3. Kaplan-Meier Estimates of the Cumulative Incidence of Investigator-Reported Congestive Heart Failure (CHF), Pulmonary Edema (PE), or Cardiac Failure (CF). Vertical lines indicate 95 percent confidence intervals.

The Kaplan-Meier curves for the cumulative incidence of congestive heart failure, pulmonary edema, and cardiac failure (Fig. 3) showed early separation of the two groups (at approximately five months), with no significant departures from proportional hazards over time and a hazard ratio of 4.61 for the comparison of the rofecoxib group with the placebo group (95 percent confidence interval, 1.50 to 18.83). The hazard ratios for edema and hyperten-

The NEW ENGLAND JOURNAL of MEDICINE

Table 3. Summary of Rates and Relative Risks of Confirmed Serious Thrombotic Events and the APTC End Point.*

Adverse Event	Rofecoxib Group				Placebo Group				Difference in Rate (95% CI)	Relative Risk (95% CI)
	No. at Risk	No. of Events	No. of Patient-yr at Risk	Rate/100 Patient-yr	No. at Risk	No. of Events	No. of Patient-yr at Risk	Rate/100 Patient-yr		
Confirmed event										
Overall	1287	46	3059	1.50	1299	26	3327	0.78	0.72 (0.19 to 1.25)	1.92 (1.19 to 3.11)
Month 0–18	1287	22	1656	1.33	1299	20	1765	1.13	0.20 (-0.55 to 0.94)	1.18 (0.64 to 2.15)
Month 19–36	989	24	1403	1.71	1079	6	1561	0.38	1.33 (0.58 to 2.08)	4.45 (1.77 to 13.32)
APTC end point										
Overall	1287	34	3070	1.11	1299	18	3334	0.54	0.57 (0.12 to 1.02)	2.06 (1.16 to 3.64)
Month 0–18	1287	14	1658	0.84	1299	12	1769	0.68	0.17 (-0.42 to 0.75)	1.25 (0.58 to 2.69)
Month 19–36	994	20	1412	1.42	1083	6	1565	0.38	1.03 (0.34 to 1.73)	3.69 (1.43 to 11.24)

* CI denotes confidence interval, and APTC Antiplatelet Trialists' Collaboration.

sion were lower than those for the combined end point of congestive heart failure, pulmonary edema, or cardiac failure (Table 4), but the event curves showed an early separation similar to that for the combined end point (data not shown).

During the trial, the rofecoxib group had mean (\pm SE) increases of 3.4 ± 0.4 mm Hg in systolic blood pressure and 0.9 ± 0.2 mm Hg in diastolic blood pressure, as compared with respective changes of -0.5 ± 0.3 mm Hg and -0.8 ± 0.2 mm Hg in the placebo group ($P<0.01$ for the comparison between the two groups). Blood-pressure effects were seen by four weeks and remained relatively constant throughout the study. To investigate the relation between changes in blood pressure and confirmed thrombotic events, we categorized patients according to the change from baseline in mean arterial pressure at four weeks. The relative risks of a confirmed thrombotic event in the rofecoxib group, as compared with the placebo group, were broadly similar across quartile categories of the change in blood pressure (data not shown). The mean arterial pressure throughout the study, included as a time-varying covariate, did not materially modify the treatment effect (relative risk for the comparison of the rofecoxib group with the placebo group, 1.87; 95 percent confidence interval, 1.14 to 3.06).

DISCUSSION

COX-2 inhibitors have been widely used as anti-inflammatory and pain-relief agents and may hold promise as chemopreventive agents for a variety of epithelial cancers. In this randomized, placebo-controlled, double-blind trial, we found that long-term use of the COX-2 inhibitor rofecoxib was associated with an increased risk of cardiovascular events. In post hoc analyses, the increased relative risk of adjudicated thrombotic events was first observed after approximately 18 months of treatment. The overall risk did not appear to be significantly influenced by baseline or subsequent use of low-dose aspirin. In addition, there was an increased frequency of investigator-reported events, such as hypertension, edema, and congestive heart failure, which occurred much earlier in the study.

Thromboxane A_2 , a major COX-1-mediated product of arachidonic acid metabolism, causes irreversible platelet aggregation, vasoconstriction, and smooth-muscle proliferation, whereas prostacyclin is an inhibitor of platelet aggregation, a vasodilator, and an inhibitor of smooth-muscle proliferation. COX-2 is the chief source of systemic prostacyclin synthesis,²⁵ and COX-2 inhibitors may increase the cardiovascular risk by shifting the functional bal-

CARDIOVASCULAR EVENTS ASSOCIATED WITH ROFECOXIB

Table 4. Incidence of Nonadjudicated Cardiovascular Adverse Events.*

Adverse Event†	Rofecoxib Group (N=1287)		Placebo Group (N=1299)		Hazard Ratio (95% CI)
	No. of Patients (%)	Rate/100 Patient-yr	No. of Patients (%)	Rate/100 Patient-yr	
Hypertension	377 (29.3)	14.9	219 (16.9)	7.3	2.02 (1.71–2.38)
Serious event	11		1		
Edema	111 (8.6)	3.8	76 (5.9)	2.4	1.57 (1.17–2.10)
Serious event	3		0		
Congestive heart failure, pulmonary edema, or cardiac failure	17 (1.3)	0.6	4 (0.3)	0.1	4.61 (1.50–18.83)
Serious event	12		2		

* The total duration of follow-up was 3059 patient-years in the rofecoxib group and 3327 patient-years in the placebo group. Although a patient may have had two or more clinical adverse events, the patient was counted once within a category. The same patient may appear in different categories. CI denotes confidence interval.

† A serious event was defined as one that was life-threatening, resulted in (or prolonged) hospitalization, or caused permanent disability.

ance of these vasoactive eicosanoids toward the promotion of thrombosis or atherogenesis. COX-2 inhibition combined with thromboxane-receptor antagonism may also lead to the destabilization of atheromatous plaque.¹² In addition, COX-2 plays a role in angiogenesis.¹ How these pharmacologic observations relate to the clinical cardiovascular findings with COX-2 inhibition is unknown. It is also not clear whether the partial inhibition of COX-1 by various nonselective NSAIDs offsets any adverse cardiovascular effects of COX-2 inhibition, since this possibility has not been evaluated explicitly in trials.

The VIGOR study²³ compared 50 mg of rofecoxib daily with 500 mg of naproxen twice daily in patients with rheumatoid arthritis and found rofecoxib to be associated with a higher incidence of myocardial infarction. It was unclear how much of the increase in risk was due to a deleterious effect of high-dose rofecoxib, a protective effect of naproxen, chance, or a combination of these factors.²⁶ A recent meta-analysis²¹ suggested that the magnitude of any cardioprotective effect of naproxen is unlikely to account entirely for these findings.

In aggregate, previous randomized, controlled trials comparing rofecoxib with placebo or conventional NSAIDs other than naproxen have not demonstrated an increased cardiovascular risk associated with rofecoxib use. Analysis of a database including 5435 patients with osteoarthritis in eight double-blind, placebo-controlled, phase 2B or

phase 3 trials reported similar rates of thrombotic cardiovascular adverse events with rofecoxib, placebo, and various nonselective NSAIDs.⁵ A pooled analysis of data from more than 28,000 patients with various diseases (representing more than 14,000 patient-years at risk) from 23 previous trials of rofecoxib (phase 2B through phase 5), including patients from the VIGOR trial, also did not demonstrate a significant increase in cardiovascular risk for rofecoxib as compared with placebo or NSAIDs other than naproxen.⁴ This analysis used the APTC end point we evaluated. An updated analysis that included data from various placebo-controlled studies investigating rofecoxib for the treatment or prevention of Alzheimer's disease did not demonstrate an excess of cardiovascular events associated with rofecoxib therapy.⁵ A recent meta-analysis comparing cardiovascular risk in trials that included various doses of rofecoxib suggested an increased relative risk among patients taking rofecoxib, as compared with those taking naproxen, but not placebo.²² Differences between our results and these earlier clinical-trial data may be related to differences in defined end points or the duration of treatment, a possibility supported by the apparent absence of a difference in adjudicated thrombotic events during the first 18 months of our study.

Observational studies have provided conflicting data on the cardiovascular safety of rofecoxib. A Canadian retrospective cohort study did not demonstrate an increased risk of myocardial infarction

THE NEW ENGLAND JOURNAL OF MEDICINE

among new users of rofecoxib as compared with control subjects,¹⁸ but a case-control study of patients 65 years of age or older suggested a dose-dependent elevation in the relative risk of acute myocardial infarction with rofecoxib therapy.⁷ Unlike the findings in the current study, this risk was elevated during the first 90 days of use, but not thereafter. A retrospective cohort study that assessed the occurrence of serious coronary heart disease among NSAID users¹⁹ showed an elevated cardiovascular risk associated with the use of high-dose rofecoxib, but no increased risk with the use of doses of 25 mg or less.

In our randomized, placebo-controlled trial, we found an increased risk of confirmed thrombotic events associated with the long-term use of rofecoxib. The increase in adjudicated thrombotic events associated with rofecoxib therapy was not evident during the first 18 months of the trial. Other investigator-reported cardiovascular events known to be associated with NSAID use, such as congestive heart failure and pulmonary edema, although less well defined, occurred earlier (at approximately five months) and at a higher rate among patients taking rofecoxib than among those taking placebo.

Patients in the rofecoxib group had increases in systemic arterial pressure during the trial, a finding that is consistent with the previously reported renovascular effects of NSAIDs. These changes in blood pressure were observed early in the study, along with investigator-reported edema and congestive heart failure. Mean arterial pressure did not appear to have a significant association with confirmed thrombotic events, however, according to an assessment of changes from baseline to four weeks and an analysis that included mean arterial pressure as a time-varying covariate in a model of treatment effects. On the basis of these findings, it is unlikely that changes in blood pressure were the explana-

tion for the excess cardiovascular risk in our study. However, hemodynamic changes could have contributed to a degree that is difficult to determine from the available data.

It is unclear whether the results seen with rofecoxib represent a general effect of COX-2 inhibitors or a specific effect of rofecoxib. A recent case-control study²⁷ suggested that the odds of nonfatal myocardial infarction differ between patients who take rofecoxib and those who take celecoxib, and a nested case-control study²⁰ also suggested that there are differences in the risk of serious coronary heart disease between the two agents. Elsewhere in this issue of the *Journal*, Nussmeier et al. report that patients who received parecoxib and valdecoxib for pain in the first 10 days after coronary-artery bypass grafting had an increased risk of cardiovascular events during 30 days of follow-up.²⁸ Also in this issue, Solomon et al. report that an ongoing safety review of the Adenoma Prevention with Celecoxib Trial revealed that the risk of fatal or nonfatal cardiovascular events was increased by a factor of 2.3 among patients who were randomly assigned to receive celecoxib, as compared with those who were assigned to receive placebo,²⁹ leading the National Cancer Institute to suspend the trial. The possibility that conventional NSAIDs may have similar effects also has to be considered. Possible cardiovascular effects will need to be taken into account in an assessment of the potential ability of any of these drugs to prevent neoplasia in the large bowel and other organs.

Funded by Merck Research Laboratories.

Drs. Bresalier, Sandler, Riddell, Morton, Lanas, and Baron report having received consulting fees from Merck Research Laboratories. Dr. Baron also reports having served as an unpaid consultant to Bayer. Dr. Konstam reports having received consulting fees from Merck. Mr. Bolognese and Drs. Quan, Oxenius, Horgan, and Lines are employees of Merck, and Drs. Quan, Oxenius, Horgan, and Lines and Mr. Bolognese own equity in the company.

APPENDIX

The following persons and institutions participated in the APPROVe Trial: Steering Committee — J.A. Baron (chair), R.S. Bresalier, R.S. Sandler, R. Riddell, D. Morton, A. Lanas, B. Oxenius (nonvoting member), J.A. Bolognese (nonvoting member), K. Horgan (nonvoting member); External Safety Monitoring Board — J. Neaton (chair), M.A. Konstam, D. Bjorkman, R. Logan, H. Quan (nonvoting member); Adjudication Committees — *Cardiology*: L.S. Dreifus, G. Vetrovec, B. Chaitman; *Neurology*: H. Adams, J.P. Mohr, J. Zivin; *Peripheral Vascular*: J. Ginsberg, C. Kearon, T. Rooke; *Gastrointestinal*: M. Griffin, M. Langman, D. Jensen; Investigators — M. Aguilar, Clinica Aguilar Bonilla, San Jose, Costa Rica; P. Angus, Austin & Repatriation Medical Centre, Heidelberg, Australia; N. Arber, Tel Aviv Sourasky Medical Center, Tel Aviv; J.M.P. Badia, Hospital Clinic I Provincial, Barcelona, Spain; R.D. Baerg, Tacoma Digestive Disease Center, Tacoma, Wash.; H. Baistrocchi, Unidad de Aparato Digestivo Julio Dante Baistrocchi, Cordoba, Argentina; M.L. Barclay, Christchurch Hospital, Christchurch, New Zealand; C. Beglinger, University of Basel, Basel, Switzerland; G. Bianchi-Portro, Ospedale Luigi Sacco, Milan; T. Bolin, Prince of Wales Hospital, Randwick, Australia; R.M. Bostick, Palmetto Health South Carolina Cancer Center, Columbia; R.S. Bresalier, A.A. Dekovich, T. Ben-Menachem, S.K. Batra, Henry Ford Hospital, Detroit; E. Bruun, J. Christiansen, Amtssygehuset i Herlev, Herlev, Denmark; C. Burke, Cleveland Clinic Foundation, Cleveland; E. Butruk, Akademia Medyczna w Warszawie, Warsaw; L. Capurso, Azienda Ospedaliera San Filippo, Rome; J.P. Cello, San Francisco General Hospital, San Francisco; S. Chaussade, Hospital Cochin Saint-Jacques, Paris; D.P. Cleland, Montreal General Hospital, Montreal; G. Costamagna, Università Cattolica del Sacro Cuore, Rome; P. Crone, Kobenhavns Amtssygehus i Glos-

CARDIOVASCULAR EVENTS ASSOCIATED WITH ROFECOXIB

trup, Glostrup, Denmark; E.V. Cutsem, Universitaire Ziekenhuizen, Leuven, Belgium; G.R. D'Haens, Imeldaziekenhuis, Bon Heiden, Belgium; W. Dekker, J. Ferwerda, Kennemer Gasthuis, Haarlem, the Netherlands; E. Domínguez-Munoz, Hospital de Conxo, La Coruña, Spain; D.S. Eskreis, R.E. Tepper, Long Island Clinical Research Associates, Great Neck, N.Y.; R. Estela, Hospital Clínico San Borja-Arriaran, Santiago, Chile; M. Färkkilä, University Central Hospital, Helsinki; G.M. Fugarolas, J.F. deDios, Hospital Universitario Reina Sofía, Córdoba, Spain; A. Giacosa, Istituto Nazionale Per La Ricerca Sul Cancro, Genoa, Italy; M.J. Goldstein, Long Island Gastro Intestinal Research Group, Great Neck, N.Y.; E. Gomollon-García, Hospital Universitario Miguel Servet, Zaragoza, Spain; P. Gandrup, Aalborg Sygehus, Aalborg, Denmark; A. Habr-Gama, Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, São Paulo; M. Haque, S. Parry, Middlemore Hospital, Auckland, New Zealand; R. Hardi, Metropolitan Gastroenterology Group, Chevy Chase, Md.; W. Hartford, Veterans Affairs Medical Center, Dallas; S.M. Harris, N.B. Vakil, Aurora Sinai Medical Center, Milwaukee; P.R. Holt, D.P. Kotler, Saint Luke's-Roosevelt Hospital, New York; P.A. Holt, Endoscopic Microsurgery Associates, Towson, Md.; R.W. Hulterantz, Karolinska Universitetssjukhuset-Solna, Stockholm; S.H. Itzkowitz, Mount Sinai Medical Center, New York; R.F. Jacoby, University of Wisconsin Medical School, Madison; K.E.J. Jensen, Centralsygehuset Esbjerg Varde, Esbjerg, Denmark; J.E. Johanson, Rockford Gastroenterology Associates, Rockford, Ill.; P.W. Jørgensen, Bispebjerg Hospital, Copenhagen; K.E. Kim, University of Chicago Medical Center, Chicago; P. Knoflach, AKH Barmherzige, Schwestern vom heiligen Kreuz, Wels, Austria; M. Koch, Capital Gastroenterology Consultants, Silver Spring, Md.; B. Koch, St. Vincenz Krankenhaus, Datteln, Germany; A. Lanas, University Clinical Hospital, Zaragoza, Spain; M.R. Lane, Auckland City Hospital, Auckland, New Zealand; T.R. Liebermann, Radiant Research, Austin, Tex.; M. Lukas, Univerzita Karlova, Prague; C.M. Schmitt, Southern Clinical Research, Chattanooga, Tenn.; R. Macrae, Cabrini Hospital, Malvern, Australia; E.E. Maiza, Hospital Jose Joaquin Aguirre, Santiago, Chile; N.E. Marcon, Toronto; R.D. Marks, Alabama Digestive Research Center, Alabaster; C.E. Martínez, Hospital Militar Central, Bogotá, Colombia; R. McLeod, Mount Sinai Hospital, Toronto; K.R. McQuaid, Veterans Affairs Medical Center, San Francisco; G. Minoli, Ospedale Valduce Reparto, Como, Italy; M. Montoro, Hospital San Jorge, Huesca, Spain; A. Montoya, Clínica Shaio, Bogotá, Colombia; G. Morelli, Optimum Clinical Research, Montreal; D.G. Morton, Queen Elizabeth Hospital, Edgbaston, United Kingdom; T.J. Myrhol, J.R. Andersen, Hvidovre Hospital, Hvidovre, Denmark; A. Nakad, Hospital Notre Dame, Tournai, Belgium; V. Narayan, Gastrointestinal Diagnostic Center, Baltimore; Y. Niv, Rabin Medical Center, Petah Tikva, Israel; P.M. Pardoll, S. Scheinert, Center for Digestive Diseases, St. Petersburg, Fla.; C. Phino, J.M. Soares, Celestial Ordem Terceira da Santissima, Porto, Portugal; I. Pokorny, Prerov, Czech Republic; J. Ponce-García, Hospital Universitari La Fe Valencia, Valencia, Spain; T. Ponchou, Hospital Edouard Herriot, Lyon, France; J.H. Pressman, San Diego Digestive Disease Consultants, San Diego, Calif.; V. Prochazka, Fakultni Nemocnice Olomouc, Olomouc, Czech Republic; J.M. Provenza, Louisiana Research Center, Shreveport; W.S. Putnam, Seattle Gastroenterology Associates, Seattle; E. Quintero-Carrion, Hospital Universitario de Canarias Tenerife, Santa Cruz de Tenerife, Spain; J.P. Raufman, V.Raj, University of Arkansas for Medical Sciences, Little Rock; D.K. Rex, Indiana University Hospital, Indianapolis; F.P. Rossini, M. Spandre, Azienda Sanitaria Ospedaliera, Turin, Italy; R.I. Rothstein, Dartmouth Hitchcock Medical Center, Lebanon, N.H.; A.K. Rustgi, University of Pennsylvania, Philadelphia; R. Sandler, University of North Carolina at Chapel Hill, Chapel Hill; B. Schweizer, Sunninghill Clinic, Sandton, South Africa; R.E. Schoen, University of Pittsburgh Medical Center, Pittsburgh; T.T. Schubert, Allenmore Medical Center, Tacoma, Wash.; H.L. Schwartz, Miami Research Associates, Miami; E. Segal, Hospital General de Agudos Carlos G. Durand, Buenos Aires; E. Seow-Choen, K.W. Eu, Singapore General Hospital, Singapore; N.R. Shah, Philip J. Bean Medical Center, Hollywood, Md.; N. Skandalis, Peripheral General Hospital, Athens; E.L. Szego, Royal Victoria Hospital, Montreal; N. Toribara, University of Colorado Health Sciences Center, Denver; J. Torosis, Gastrointestinal Research, Redwood City, Calif.; D.K. Turgeon, University of Michigan, Ann Arbor; S.W. Van der Merwe, Pretoria, South Africa; R. VanStolk, C.W. Howden, Northwestern University, Chicago; P. Vergauwe, Algemeen Ziekenhuis Groeninge, Kortrijk, Belgium; B. Vergeau, C. Nizou, Hôpital d'Instruction des Armees, Paris; G. Winde, Klinikum Kreis Herford, Herford, Germany; J. Wolosin, M.W. Swaim, Regional Research Institute, Jackson, Tenn.; J.C. Wolper, E.L. Yudelman, Digestive Health Physicians, Fort Myers, Fla.; B.C.Y. Wong, University of Hong Kong Queen Mary Hospital, Hong Kong; J.P. Wright, Kinsbury Hospital, Claremont, South Africa; A. Zambelli, Ospedale Maggiore Azienda Ospedaliera, Crema, Italy.

REFERENCES

1. Dannenberg AJ, Subbaramaiah K. Targeting cyclooxygenase-2 in human neoplasia: rationale and promise. *Cancer Cell* 2003;4:431-6.
2. Crofford LJ, Lipsky PE, Brooks P, Abramson SB, Simon LS, van de Putte LB. Basic biology and clinical application of specific cyclooxygenase-2 inhibitors. *Arthritis Rheum* 2000;43:4-13.
3. Howard PA, Delafontaine P. Nonsteroidal anti-inflammatory drugs and cardiovascular risk. *J Am Coll Cardiol* 2004;43:519-25.
4. Konstam MA, Weir MR, Reicin A, et al. Cardiovascular thrombotic events in controlled, clinical trials of rofecoxib. *Circulation* 2001;104:2280-8.
5. Weir MR, Sperling RS, Reicin A, Gertz BJ. Selective COX-2 inhibition and cardiovascular effects: a review of the rofecoxib development program. *Am Heart J* 2003;146:591-604.
6. Solomon DH, Glynn RJ, Levin R, Avorn J. Nonsteroidal anti-inflammatory drug use and acute myocardial infarction. *Arch Intern Med* 2002;162:1099-104.
7. Solomon DH, Schneeweiss S, Glynn RJ, et al. Relationship between selective cyclooxygenase-2 inhibitors and acute myocardial infarction in older adults. *Circulation* 2004;109:2068-73.
8. Strand V, Hochberg MC. The risk of cardiovascular thrombotic events with selective cyclooxygenase-2 inhibitors. *Arthritis Rheum* 2002;47:349-55.
9. Mamdani M, Juurlink DN, Lee DS, et al. Cyclo-oxygenase-2 inhibitors versus non-selective non-steroidal anti-inflammatory drugs and congestive heart failure outcomes in elderly patients: a population-based cohort study. *Lancet* 2004;363:1751-6.
10. Whelton A, Fort JG, Puma JA, Normandin D, Bello AB, Verburg KM. Cyclooxygenase-2-specific inhibitors and cardiorenal function: a randomized, controlled trial of celecoxib and rofecoxib in older hypertensive osteoarthritis patients. *Am J Ther* 2001;8:85-95. [Erratum, *Am J Ther* 2001;8:220.]
11. Whelton A, White WB, Bello AB, Puma JA, Fort JG. Effects of celecoxib and rofecoxib on blood pressure and edema in patients ≥65 years of age with systemic hypertension and osteoarthritis. *Am J Cardiol* 2002;90:959-63.
12. Egan KM, Wang M, Lucitt MB, et al. Cyclooxygenases, thromboxane, and atherosclerosis: plaque destabilization by cyclooxygenase-2 inhibition combined with thromboxane receptor antagonism. *Circulation* 2005;111:334-42.
13. Bombardier C, Laine L, Reicin A, et al. Comparison of upper gastrointestinal toxicity of rofecoxib and naproxen in patients with rheumatoid arthritis. *N Engl J Med* 2000;343:1520-8.
14. Rahme E, Pilote L, LeLorier J. Association between naproxen use and protection against acute myocardial infarction. *Arch Intern Med* 2002;162:1111-5. [Erratum, *Arch Intern Med* 2002;162:1858.]
15. Watson DJ, Rhodes T, Cai B, Guess HA. Lower risk of thromboembolic cardiovascular events with naproxen among patients with rheumatoid arthritis. *Arch Intern Med* 2002;162:1105-10. [Erratum, *Arch Intern Med* 2002;162:1779.]

CARDIOVASCULAR EVENTS ASSOCIATED WITH ROFECOXIB

16. Farkonh ME, Kirshner H, Harrington RA, et al. Comparison of lornoxicoxib with naproxen and ibuprofen in the Therapeutic Arthritis Research and Gastrointestinal Event Trial (TARGET), cardiovascular outcomes: randomised controlled trial. *Lancet* 2004;364:675-84.
17. Reicin AS, Shapiro D, Sperling RS, Barr E, Yu Q. Comparison of cardiovascular thrombotic events in patients with osteoarthritis treated with rofecoxib versus nonselective nonsteroidal anti-inflammatory drugs (ibuprofen, diclofenac, and nabumetone). *Am J Cardiol* 2002;89:204-9.
18. Mamdani M, Rochon P, Juurlink DN, et al. Effect of selective cyclooxygenase 2 inhibitors and naproxen on short-term risk of acute myocardial infarction in the elderly. *Arch Intern Med* 2003;163:481-6.
19. Ray WA, Stein CM, Daugherty JR, Hall K, Arbogast PG, Griffin MR. COX-2 selective non-steroidal anti-inflammatory drugs and risk of serious coronary heart disease. *Lancet* 2002;360:1071-3.
20. Graham DJ, Campen D, Hui R, et al. Risk of acute myocardial infarction and sudden cardiac death in patients treated with cyclo-oxygenase 2 selective and non-selective non-steroidal anti-inflammatory drugs: nested case-control study. *Lancet* 2005;365:475-81.
21. Juni P, Nartey L, Reichenbach S, Sterchi R, Dieppe PA, Egger M. Risk of cardiovascular events and rofecoxib: cumulative meta-analysis. *Lancet* 2004;364:2021-9.
22. Baron JA, Cole BF, Sandler RS, et al. A randomized trial of aspirin to prevent colorectal adenomas. *N Engl J Med* 2003;348:891-9.
23. Antiplatelet Trialists' Collaboration. Collaborative overview of randomised trials of antiplatelet therapy. I. Prevention of death, myocardial infarction, and stroke by prolonged antiplatelet therapy in various categories of patients. *BMJ* 1994;308:81-106. [Erratum, *BMJ* 1994;308:1540.]
24. Guess HA, Lydick EG, Small RD, Miller LP. Epidemiologic programs for computers and calculators: exact binomial confidence intervals for the relative risk in follow-up studies with sparsely stratified incidence density data. *Am J Epidemiol* 1987;125:340-7.
25. McAdam BF, Catella-Lawson F, Mardini IA, Kapoor S, Lawson JA, FitzGerald GA. Systemic biosynthesis of prostacyclin by cyclooxygenase (COX)-2: the human pharmacology of a selective inhibitor of COX-2. *Proc Natl Acad Sci U S A* 1999;96:272-7.
26. Fitzgerald GA, Patrono C. The coxibs, selective inhibitors of cyclooxygenase-2. *N Engl J Med* 2001;345:433-42.
27. Kimmel SE, Berlin JA, Reilly M, et al. Patients exposed to rofecoxib and celecoxib have different odds of nonfatal myocardial infarction. *Ann Intern Med* 2005;142:157-64.
28. Nussmeier NA, Whelton AA, Brown MT, et al. Complications of the COX-2 inhibitors parecoxib and valdecoxib after cardiac surgery. *N Engl J Med* 2005;352:1081-91.
29. Solomon SD, McMurray JJV, Pfeffer MA, et al. Cardiovascular risk associated with celecoxib in a clinical trial for colorectal adenoma prevention. *N Engl J Med* 2005;352:1071-80.

Copyright © 2005 Massachusetts Medical Society.

POWERPOINT SLIDES OF JOURNAL FIGURES AND TABLES

At the *Journal's* Web site, subscribers can automatically create PowerPoint slides of *Journal* figures and tables. Click on a figure or table in the full-text version of any article at www.nejm.org, and then click on PowerPoint Slide for Teaching. A PowerPoint slide containing the image, with its title and reference citation, can then be downloaded and saved.

4. The Abbreviated Clinical Study Report for the APPROVe Trial.

a. Introduction.

Shortly after unblinding of the APPROVe Trial data on September 23, 2004, MRL scientists began to prepare a submission to the FDA reporting the APPROVe Trial results. As discussed in Appendix I, MRL generally reports safety data from a trial to the FDA in the form of a Clinical Study Report. Procedure No. 21 of Merck's Medical Affairs Procedures and Policies provides that an abbreviated Clinical Study Report, as opposed to the full Clinical Study Report normally filed with the FDA, may be used "when a program is canceled or a study is discontinued prematurely and yields no usable efficacy or immunogenicity data."²³¹

In this case, because Merck had withdrawn Vioxx from the market and discontinued the Vioxx development program, MRL scientists prepared an abbreviated Clinical Study Report to report the APPROVe Trial data to domestic and worldwide regulatory agencies. Given the fact that the APPROVe Trial resulted in the withdrawal of Vioxx and the obvious public interest in the data from this Trial, however, the abbreviated Clinical Study Report was much more extensive than required by Merck's Medical Affairs Procedures and Policies.²³²

²³¹ Merck Medical Affairs Procedures and Policies – Procedure 21, MRK-AGE0001636, at 37. This procedure further provided that this format may be used "when a study is in progress and, for full disclosure in a regulatory submission, the safety information must be briefly summarized." Id.

²³² See 10/25/04 email from J. Bolognese to T. Capizzi, MRK-AFO0264630 (noting that for "obvious reasons," the abbreviated Clinical Study Report for APPROVe would be "much more extensive" than "the usual aCSR."); 10/08/04 email from K. Horgan to B. Oxenius and T. Reiss, MRK-AHC0014922.

Shortly after the Company withdrew Vioxx from the market, MRL representatives met with the FDA a number of times regarding timing for the submission of the APPROVe Trial data and the analyses to be included.²³³ On January 26, 2005, in accordance with the timeline agreed upon with the FDA, Merck submitted to the FDA the cardiovascular safety portion of the yet-to-be completed abbreviated Clinical Study Report.²³⁴ This version was, to the extent relevant to this Report, similar to the final abbreviated Clinical Study Report. The final abbreviated Clinical Study Report was submitted to the FDA on June 6, 2005 and included a short synopsis and three appendices: the Efficacy Report (concerning colon polyps), the General Safety Report, and the Cardiovascular Safety Report.²³⁵

The following sections describe the drafting process of the abbreviated Clinical Study Report, with focus on the drafting and content of the cardiovascular safety portion of the Report. In particular, this section discusses the abbreviated Cardiovascular Safety Report's description of the test for proportionality of the hazard rate, the characterization of the results from these analyses, and the Report's conclusion that the APPROVe Trial showed an increased cardiovascular risk on Vioxx beginning after 18 months.

²³³ 9/29/04 email from P. Davish to J. Wainwright *et al.*, MRK-ACM0005819; 6/6/05 letter from P. Huang to B. Harvey* (FDA), MRK-S0420050996; 10/07/04 email from D. Louie to N. Braunstein *et al.*, MRK-ABY0179883-85; Minutes of 11/10/04 FDA teleconference, MRK-AFK0192456, at 57; see also 6/14/05 email from Regulatory Release Mailbox to N. Braunstein *et al.*, MRK-AOY0032263 (attaching FDA eDossier reflecting timeline for APPROVe regulatory submissions).

²³⁴ 1/26/05 letter from P. Huang to B. Harvey* (FDA) attaching 1/17/05 APPROVe Trial Cardiovascular Safety Report draft, MRK-AIU0057340-406.

²³⁵ 6/6/05 letter from P. Huang to B. Harvey* (FDA), MRK-S0420050996; 12/21/04 memorandum from S. Loftus to A. Reicin *at al.*, MRK-AHC0032613, at 13-14.

b. Summary of contents.

The final Cardiovascular Safety Report contained in the abbreviated Clinical Study Report presented detailed analyses of the APPROVe Trial cardiovascular data, including Kaplan-Meier plots and hazard plots with proportionality p-values for the confirmed thrombotic and APTC composite cardiovascular event endpoints, investigator-reported cardiovascular events, and subgroup analyses of confirmed thrombotic cardiovascular events.²³⁶ The final Cardiovascular Safety Report also included analyses of the relationship between blood pressure and cardiovascular events in the APPROVe Trial.²³⁷

Like the APPROVe article published in the New England Journal of Medicine, the Cardiovascular Safety Report stated that (i) the Kaplan-Meier plot for confirmed thrombotic events revealed an increased relative risk on Vioxx “after approximately 18 months,” (ii) that the apparent change in the hazard ratio was confirmed by a failed test for proportionality of hazards,²³⁸ and (iii) that the proportionality of hazards was analyzed

²³⁶ 03/15/05 APPROVe Trial abbreviated Clinical Study Report, Cardiovascular Safety Report, MRK-I8940100962, passim; see also 1/17/05 APPROVe Trial Cardiovascular Safety Report draft, MRK-AIU0057345, passim (attached to 1/26/05 letter from P. Huang to B. Harvey* (FDA), MRK-AIU0057340).

²³⁷ 03/15/05 APPROVe Trial abbreviated Clinical Study Report, Cardiovascular Safety Report, MRK-I8940100962, at 1010-18; see also 1/17/05 APPROVe Trial Cardiovascular Safety Report draft, MRK-AIU0057345, at 386-92 (attached to 1/26/05 letter from P. Huang to B. Harvey* (FDA), MRK-AIU0057340).

²³⁸ 3/15/05 APPROVe Trial abbreviated Clinical Study Report, Cardiovascular Safety Report, MRK-I8940100962, at 0969.

using the Logarithm of Time Test. Unlike the APPROVe article, it did not state that the test of proportionality of hazards was “specified.”²³⁹

c. Procedural aspects of the drafting process.

The principal author and point person for the abbreviated Clinical Study Report was Dr. Raymond Joseph,²⁴⁰ who had joined Dr. Bettina Oxenius as co-clinical monitor of the APPROVe Trial in mid-October 2002. Dr. Joseph’s co-authors were Dr. Hui Quan, the unblinded statistician for the APPROVe Trial, Dr. Bettina Oxenius, clinical monitor on the APPROVe Trial beginning in fall 2001, Ms. Susan Loftus, Medical Program Clinical Specialist and Dr. Eric Maller, Director of Clinical Research.²⁴¹

²³⁹ 3/15/05 APPROVe Trial abbreviated Clinical Study Report, Cardiovascular Safety Report, MRK-I8940100962, at 0969, 0973-977. As discussed above, in May 2006, MRL Scientists discovered that the APPROVe article and the abbreviated Clinical Study Report mistakenly reported that the Logarithm of Time Test resulted in a p-value of 0.014 for confirmed thrombotic events and 0.119 for APTC events. These results were in fact obtained by running a linear time method to assess the proportionality assumption. Although the abbreviated Clinical Study Report did not indicate that the test used to assess the proportionality assumption was pre-specified and that there was no pre-specified cardiovascular data analysis plan for the APPROVe Trial alone, recent internal correspondence in connection with discovery of the error suggests that the Statistical Data Analysis Plan for Protocol 203 may have been looked to when preparing the cardiovascular analyses for APPROVe. In a May 25, 2006 memorandum to Drs. Theodore Reiss, Barry Gertz, Raymond Bain, Bettina Oxenius, Ned Braunstein, Philip Huang and Ms. Joanne Lahner, Messrs. Jim Bolognese and Tom Cook explained that “[t]here was no Statistical Analyses Plan (SAP) for the confirmed thrombotic CV data from APPROVe alone. The CV data from APPROVe, along with that from two other trials (VICTOR and VIP), form the basis for VIOXX Protocol 203. Although the SAP for Protocol 203 specified that the primary assessment of proportional hazards would be the test of treatment-by-natural-log(time) interaction [logarithm of time], a detailed assessment of relationship of hazard ratio with time [analyses performed by Drs. Jennifer Ng and Hongwei Wang] revealed a better fit of the [confirmed thrombotic event hazard ratio] over time relationship by a model including linear time in comparison to one including log(time).” 5/25/06 memorandum from J. Bolognese and T. Cook to R. Bain et al., MRK-S0420112192.

²⁴⁰ 11/09/04 email from B. Oxenius to T. Biftu, MRK-AFN0073160.

²⁴¹ 03/15/05 APPROVe Trial abbreviated Clinical Study Report, Cardiovascular Safety Report, MRK-I8940100962.

Although the drafting was collaborative,²⁴² Dr. Quan had primary responsibility for preparing the Cardiovascular Safety Report and the Efficacy Report, while Ms. Loftus provided overall support and had primary responsibility for the General Safety Report.²⁴³

Preparation and review of the abbreviated Clinical Study Report was guided by Merck's Medical Affairs Procedures and Policies.²⁴⁴ Procedure No. 26 detailed formatting guidelines for full or modified Clinical Study Reports, and Procedure No. 21 provided guidelines for format and content of abbreviated Clinical Study Report.²⁴⁵ While the guidelines for content and format of an abbreviated Clinical Study Report differed from those applying to a full report, the procedures for planning, review and authorship delineated in Policy No. 16 were the same for both formats.²⁴⁶

Procedure No. 16 required that the statistical author (in this case Dr. Quan) "review the CSR for accuracy in reporting statistical results and data, and agreement with the [Data Analysis Plan]."²⁴⁷ It also provided the following:

²⁴² Dr. Quan distributed the Cardiovascular Report to the other authors for comment. See 11/24/2004 email from H. Quan to B. Oxenius *et al.*, attaching draft Cardiovascular Safety Report for review, MRK-AGO0074753.

²⁴³ 11/12/04 email from H. Quan to S. Loftus, MRK-AFN0073765.

²⁴⁴ Merck Medical Affairs Procedures and Policies – Guidance Document, MRK-AFK0180481, at 82.

²⁴⁵ Merck Medical Affairs Procedures and Policies – Guidance Document, MRK-AFK0180481, at 82; Merck Medical Affairs Procedures and Policies – Procedure 21, MRK-AGE0001636.

²⁴⁶ Merck Medical Affairs Procedures and Policies – Guidance Document, MRK-AFK0180481, at 83; Merck Medical Affairs Procedures and Policies – Procedure 21, MRK-AGE0001636, at 38 ("The review and approval of ACSRs is described in MAPP 16."). Procedure No. 16 provided guidance on the planning, writing, reviewing and approving of regulatory documents for use in regulatory submissions. Merck Medical Affairs Procedures and Policies – Procedure 16, MRK-AGE0000001.

²⁴⁷ Merck Medical Affairs Procedures and Policies – Procedure 16, MRK-AGE0000001, at 20.

A mandatory QC [quality check] for internal consistency with standards and a data check should be completed by all authors, according to the QC Guidance for Authors Preparing Clinical & Regulatory Documents.

* * *

Prior to releasing the CSR into the Management Review cycle, the [Medical Program Clinical Specialists – in this case, Ms. Loftus] should confirm that all authors have completed a thorough QC of the CSR and approved it for entry into Management Review Cycle. The [Medical Program Clinical Specialists] should also complete a final QC of the CSR before releasing it into Management Review cycle.²⁴⁸

Mr. Bolognese described “QC” as a process intended to ensure that all data generated by statistical programs were reflected in the proper tables in the final product. According to Mr. Bolognese, the APPROVe Trial abbreviated Clinical Study Report was subject to the “QC process” provided in Policy No. 16.²⁴⁹

²⁴⁸ Merck Medical Affairs Procedures and Policies – Procedure 16, MRK-AGE0000001, at 21.

²⁴⁹ Separate from the “QC process” described in Policy No. 16, an internal departmental (Biostatistics and Research Decision Sciences) Standard Operating Procedure – SOP No. SP-S002 – mandated peer review and validation of analyses of results from “pivotal studies that [would] be submitted to support a claim for a [Product Licensing Application, Worldwide Marketing Application, New Drug Application] or supplementary [New Drug Application], as well as any program that create[d] an analysis data set used in these analyses.” Merck Standard Operating Procedure No. SP-S002: Validating Internally Written Programs, MRK-AGE0003145, at 51. The validation and peer review process required an independent statistical programmer or statistician who was not the program developer to test the statistical programs used to ensure replication/validation of results. For certain complex programs, the SOP also required statistical peer review testing by an independent statistician. *Id.* at 53-56. Although the procedures delineated in this SOP were typically required for Clinical Study Reports, they were not mandatory for the APPROVe Trial, because after withdrawal of Vioxx, the APPROVe Trial was no longer considered a “pivotal study.” See 10/29/04 email from N. Braunstein to I. Pescek and C.Q. Song, MRK-AFO0310044 (“Since APPROVe will have an aCSR (not a full CSR) and since it is not being used as a pivotal trial, or to support any regulatory claims, we do not see any need for our usual formal validation/peer review processes applying. Rather, this seems more atuned [sic] to how we handle regulatory queries – i.e., we ‘validate’ the results by inspection of the output & its consistency with the entire set of results available.”).

In addition, Policy No. 16 provided for formal review and approval by the first line of the authors' management within each therapeutic area as well as a broader review by management in Clinical Sciences, Biostatistics and Research Decision Sciences, and Global Strategic Regulatory Development.²⁵⁰

The APPROVe Trial abbreviated Clinical Study Report was reviewed internally by Dr. Reicin, Mr. Bolognese, Dr. Reiss, and Dr. Horgan.²⁵¹ In addition, the Cardiovascular Safety Report, which was primarily prepared by Dr. Quan, was reviewed by Mr. Bolognese, Dr. Quan's supervisor, and by the other authors.²⁵² On January 11, 2005, it was distributed among a larger group for review before it was submitted to the FDA on January 26, 2005.²⁵³

d. Substantive review: drafts of the Cardiovascular Safety Report.

As noted above, the abbreviated Clinical Study Report, like the APPROVe article, erroneously stated that the proportionality of hazards had been tested using the Logarithm of Time Test, when in fact the reported p-value resulted from the Linear Time Test. In addition, like the APPROVe article, the abbreviated Clinical Study Report characterized the results for the APTC composite cardiovascular event endpoint as "similar" to those

²⁵⁰ Merck Medical Affairs Procedures and Policies – Procedure 16, MRK-AGE0000001, at 7-8.

²⁵¹ 11/03/04 email from S. Loftus to A. Reicin et al., MRK-AHC0012140; see also 12/21/04 memorandum from S. Loftus to A. Reicin et al., MRK-AHC0032613-14.

²⁵² 11/22/2004 email from H. Quan to J. Bolognese attaching APPROVe Trial Cardiovascular Safety Report draft, MRK-AGO0073798; 11/24/2004 email from H. Quan to B. Oxenius et al., attaching APPROVe Trial Cardiovascular Safety Report draft, MRK-AGO0074753.

²⁵³ 1/11/05 email from P. Huang to M. Stimson et al., attaching APPROVe Trial Cardiovascular Safety Report draft, MRK-AFK0200644.

for the confirmed thrombotic event endpoint and stated that the increased relative risk on Vioxx was revealed after 18 months of continuous therapy. The following section discusses the drafting and review process relevant to those statements.

i. Description of the Cox proportional hazards model and outputs from the analyses.

The first draft of the Cardiovascular Safety Report was distributed on November 12, 2004 and provided the following description of the proportionality of hazards test in the “Statistical Methodology” section:

Cox proportional hazards model with covariate for treatment was used to calculated [sic] the relative risk (i.e., hazard ratio) estimates The proportionality assumption for the Cox model was assessed by testing a time-dependant covariate Treatment*log (Time) in the Cox model [i.e., the Logarithm of Time Test].²⁵⁴

The above description was drafted by Dr. Quan and remained essentially the same, with minor typographical edits, throughout the drafting process. Drafts do not reflect any edits or commentary by reviewers questioning the correctness of the description.²⁵⁵ The January 26, 2005 submission to the FDA, as well as the

²⁵⁴ 11/12/04 APPROVe Trial Cardiovascular Safety Report draft, MRK-AGO0018092, at 94.

²⁵⁵ 11/22/04 APPROVe Trial Cardiovascular Safety Report draft, MRK-AGO0074646, at 49; 11/24/04 APPROVe Trial Cardiovascular Safety Report draft, MRK-AGO0074754, at 57; 12/9/04 APPROVe Trial Cardiovascular Safety Report draft, MRK-AGO0060794, at 97; APPROVe Trial Cardiovascular Safety Report draft, MRK-AGO0063695, at 701 (attached to 1/5/05 email from J. Bolognese to H. Quan, MRK-AGO0063694); 1/5/05 APPROVe Trial Cardiovascular Safety Report draft reflecting comments by A. Reicin, MRK-AHC0002489, at 496; 1/9/05 APPROVe Trial Cardiovascular Safety Report draft, MRK-AGO0035207, at 210; 1/10/05 APPROVe Trial Cardiovascular Safety Report draft, MRK-AFK0200837, at 845; 1/10/05 APPROVe Trial Cardiovascular Safety Report draft reflecting edits, MRK-AFK0200645, at 653; 1/10/05 APPROVe Trial Cardiovascular Safety Report draft reflecting comments by J. van Adelsberg, MRK-AFK0185270, at 278 (attached to 1/11/05 email from J. van Adelsberg to S. Curtis et al., MRK-AFK0185269); 1/10/05 APPROVe Trial Cardiovascular Safety Report draft reflecting comments by K. Horgan, MRK-AFN0093232, at 239

March 15, 2005 final abbreviated Clinical Study Report synopsis and Cardiovascular

Safety Report, described the test in similar language.²⁵⁶

- ii. Description of results of proportionality testing for confirmed thrombotic and APTC composite cardiovascular event endpoints.

- (a) November 12 through December 9, 2004 drafts.

The November 12, 2004 draft, which was based on dummy data, did not report the p-value outputs obtained from tests of the proportionality of hazards. However, in the sections discussing results for both confirmed thrombotic and APTC composite cardiovascular events, the draft reported the following:

Bothe [sic] figure xx and Table mm show that the hazard ratio or the relative risk for thrombotic cardiovascular [or APTC] events appeared not to be a constant over time. Thus, the proportional hazard assumption for the Cox proportional hazards model did not hold.²⁵⁷

(attached to 1/12/05 email from K. Horgan to H. Quan, MRK-AFN0093231); 1/14/05 APPROVe Trial Cardiovascular Safety Report draft, MRK-AHC0002981, at 993; 1/14/05 APPROVe Trial Cardiovascular Safety Report draft reflecting comments by A. Ko, MRK-AJJ0024204, at 211 (attached to 1/24/05 email from A. Ko to H. Quan, MRK-AJJ0024203) (description of language stating that a Cox proportional hazard model was used to calculate the relative risk is highlighted; however, description of the logarithm of time as the test used to assess the proportionality assumption is not highlighted. There is note stating “Need to explain . . . better” with regard to following statement: “If the proportionality assumption was not satisfied, the estimate of the hazard ratio from the Cox model should be interpreted as an average of the treatment effect over the time range of the study.”); 1/17/05 APPROVe Trial Cardiovascular Safety Report draft MRK-AIU0057345, at 353 (attached to 1/26/05 letter from P. Huang to B. Harvey* (FDA), MRK-AIU0057340); 1/30/05 APPROVe Trial Cardiovascular Safety Report draft with handwritten comments by H. Quan, MRK-AFN0054274, at 286; 3/15/05 APPROVe Trial abbreviated Clinical Study Report, Cardiovascular Safety Report, MRK-I8940100962, at 976.

²⁵⁶ 1/17/05 APPROVe Trial Cardiovascular Safety Report draft, MRK-AIU0057345, at 353 (attached to 1/26/05 letter from P. Huang to B. Harvey* (FDA), MRK-AIU0057340); 3/15/05 APPROVe Trial abbreviated Clinical Study Report, MRK-I8940100731, at 737 (synopsis), 976 (Cardiovascular Safety Report).

²⁵⁷ 11/12/04 APPROVe Trial Cardiovascular Safety Report draft, MRK-AGO0018092, at 99, 102.

The November 22, November 24 and December 9 drafts were based on an interim frozen data set that was created on November 10, 2004 to enable work on the abbreviated Clinical Study Report to begin pending availability of the final December 2004 data. These drafts reported statistically significant p-values of 0.0109 for confirmed thrombotic events and 0.0449 for APTC composite cardiovascular events.²⁵⁸

(b) January 5, 2005 through final
March 15, 2005 drafts.

The January 5, 2005 draft, prepared by Dr. Quan and reflecting revisions by Mr. Bolognese, appears to have been the first draft based on the final frozen dataset. This draft, and all future drafts – including the January 26, 2005 submission to the FDA and the final approved March 15, 2005 Report – reported a statistically significant p-value of 0.014 for confirmed thrombotic events and a non-statistically significant p-value of 0.119 for APTC composite cardiovascular events.²⁵⁹ Although the p-value for the APTC

²⁵⁸ 11/22/04 APPROVe Trial Cardiovascular Safety Report draft, MRK-AGO0074646, Figure 3 at 57, Figure 5 at 63; 11/24/04 APPROVe Trial Cardiovascular Safety Report draft, MRK-AGO0074754, Figure 3 at 66, Figure 5 at 71; 12/9/04 APPROVe Trial Cardiovascular Safety Report draft, MRK-AGO0060794, Figure 3 at 807, Figure 5 at 811.

²⁵⁹ APPROVe Trial Cardiovascular Safety Report draft, MRK-AGO0063695, Figure 3 at 710, Figure 5 at 717 (attached to 1/5/05 email from J. Bolognese to H. Quan, MRK-AGO0063694); 1/10/05 APPROVe Trial Cardiovascular Safety Report draft reflecting edits, MRK-AFK0200645, Figure 3 at 665, Figure 5 at 672; 1/10/05 APPROVe Trial Cardiovascular Safety Report draft reflecting comments by J. van Adelsberg, MRK-AFK0185270, Figure 3 at 290, Figure 5 at 297 (attached to 1/11/05 email from J. van Adelsberg, MRK-AFK0185269); 1/10/05 APPROVe Trial Cardiovascular Safety Report draft reflecting comments by K. Horgan, MRK-AFN0093232, Figure 3 at 249, Figure 5 at 256 (attached to 1/12/05 email from K. Horgan to H. Quan, MRK-AFN0093231); 1/14/05 APPROVe Trial Cardiovascular Safety Report draft reflecting comments by A. Ko, MRK-AJJ0024204, Figure 3 at 222, Figure 5 at 229 (attached to 1/24/05 email from A. Ko to H. Quan, MRK-AJJ0024203); 1/17/05 APPROVe Trial Cardiovascular Safety Report draft, MRK-AIU0057412, Figure 3 at 431, Figure 5 at 438 (attached to 1/26/05 letter from P. Huang to B. Harvey* (FDA), MRK-AIU0057340); APPROVe Trial Cardiovascular Safety Report draft, MRK-AFN0096304, Figure 3 at 327, Figure 5 at 334 (attached to 1/26/05 email from J. McGovern to

endpoint was not statistically significant, the executive summary of this draft and all

future drafts reported the following, with only slight modification:

For confirmed thrombotic events, there were 46 patients with the events in the rofecoxib group and 26 patients with the events in the placebo group. Over the 3-year treatment period, the patient-year adjusted event rates were 1.50 and 0.78 events per 100 patient-years for the rofecoxib and placebo groups, respectively. The relative risk of rofecoxib versus placebo was 1.92 with a 95% confidence interval of (1.19, 3.11). The relative risk was significantly greater than 1 ($p=0.008$). Examination of the Kaplan-Meier plot for cumulative proportion of patients with these events revealed similarity of risk between treatments for the first approximately 18 months of the study and increased risk with rofecoxib after approximately 18 months. The test for proportional hazards was statistically significant, indicating different relative risks over time. . . .

Results for confirmed [Antiplatelet Trialists' Collaboration] APTC events and for investigator reported events were similar to those for confirmed thrombotic events.²⁶⁰

B. Oxenius, MRK-AFN0096302); APPROVe Trial Cardiovascular Safety Report draft, MRK-AFN0096985, Figure 3 at 008, Figure 5 at 015 (attached to 1/28/05 email from H. Quan to B. Oxenius et al., MRK-AFN0096981); 1/30/05 APPROVe Trial Cardiovascular Safety Report draft with handwritten comments by H. Quan, MRK-AFN0054274, Figure 3 at 297, Figure 5 at 306; 3/15/05 APPROVe Trial abbreviated Clinical Study Report, Cardiovascular Safety Report, MRK-I8940100962, Figure 3 at 987, Figure 5 at 994.

²⁶⁰ APPROVe Trial Cardiovascular Safety Report draft, MRK-AGO0063695, at 695 (attached to 1/5/05 email from J. Bolognese to H. Quan, MRK-AGO0063694); 1/10/05 APPROVe Trial Cardiovascular Safety Report draft, MRK-AFK0200837, at 838; 1/10/05 APPROVe Trial Cardiovascular Safety Report draft reflecting edits, MRK-AFK0200645, at 646; 1/10/05 APPROVe Trial Cardiovascular Safety Report draft reflecting comments by J. van Adelsberg, MRK-AFK0185270, at 271 (attached to 1/11/05 email from J. van Adelsberg to S. Curtis et al., MRK-AFK0185269); 1/10/05 APPROVe Trial Cardiovascular Safety Report draft reflecting comments by K. Horgan, MRK-AFN0093232, at 233 (attached to 1/12/05 email from K. Horgan to H. Quan, MRK-AFN0093231); 1/14/05 APPROVe Trial Cardiovascular Safety Report draft with comments by A. Ko, MRK-AJJ0024204, at 205 (attached to 1/24/05 email from A. Ko to H. Quan, MRK-AJJ0024203); 1/17/05 APPROVe Trial Cardiovascular Safety Report draft, MRK-AIU0057345, at 346 (attached to 1/26/05 letter from P. Huang to B. Harvey* (FDA), MRK-AIU0057340); APPROVe Trial Cardiovascular Safety Report draft, MRK-AFN0096304, at 309 (attached to 1/26/05 email from J. McGovern to B. Oxenius, MRK-AFN0096302); APPROVe Trial Cardiovascular Safety Report draft, MRK-AFN0096985,

The synopsis to the final March 15, 2005 abbreviated Clinical Study Report had slightly stronger language:

Notably, results for confirmed APTC events and for investigator reported events were similar to those for confirmed thrombotic events.²⁶¹

Unlike the APPROVe article, however, the body of the Cardiovascular Safety Report provided the proportionality p-value for both confirmed thrombotic and APTC composite cardiovascular endpoints. The January 5, 2005 draft reflected the following revisions with regard to the p-value for the APTC composite cardiovascular endpoint.²⁶²

Original (Authored by Dr. Quan)	Revised (Revised by Mr. Bolognese)
“The testing of the proportional hazard assumption for the Cox proportional hazards model did not show violation of the assumption (p = 0.119).”	“The testing of the proportional hazard assumption for the Cox proportional hazards model did not show <u>significant</u> violation of the assumption (p = 0.119), <u>but it approached significance.</u> ”

The January 10, 2005 and future drafts, including the final Report, reflect the revised language.²⁶³

at 990 (attached to 1/28/05 email from H. Quan to B. Oxenius et al., MRK-AFN0096981); 1/30/05 APPROVe Trial Cardiovascular Safety Report draft with handwritten comments by H. Quan, MRK-AFN0054274, at 279; 3/15/05 APPROVe Trial abbreviated Clinical Study Report, Cardiovascular Safety Report, MRK-I8940100962, at 969.

²⁶¹ 3/15/05 APPROVe Trial abbreviated Clinical Study Report, Synopsis, MRK-I8940100731, at 740.

²⁶² APPROVe Trial Cardiovascular Safety Report draft, MRK-AGO0063695, at 717 (attached to 1/5/05 email from J. Bolognese to H. Quan, MRK-AGO0063694).

²⁶³ 1/10/05 APPROVe Trial Cardiovascular Safety Report draft, MRK-AFK0200837, at 864; 1/10/05 APPROVe Trial Cardiovascular Safety Report draft reflecting edits, MRK-AFK0200645, at 672;

iii. Conclusion that the increased cardiovascular risk
of Vioxx was revealed after 18 months.

As with the APPROVe article published in the New England Journal of Medicine, the Cardiovascular Safety Report stated that, in the APPROVe Trial, the use of Vioxx was associated with an increased risk of thrombotic events as compared to placebo beginning “after approximately 18 months” of therapy.²⁶⁴ This statement does not appear to have been the subject of much editing or discussion during the drafting and review of the Cardiovascular Safety Report.

The Executive Summary and Conclusion sections of the January 26, 2005 draft submission to the FDA read as follows:

Executive Summary	Conclusion
Examination of the Kaplan-Meier plot for the cumulative proportion of patients with [confirmed thrombotic] events revealed similarity of risk between treatments for the first 18 months of the	-- Increased risk of thrombotic cardiovascular events beginning approximately 18 months associated with rofecoxib therapy.

1/10/05 APPROVe Trial Cardiovascular Safety Report draft reflecting comments by J. van Adelsberg, MRK-AFK0185270, at 297 (attached to 1/11/05 email from J. van Adelsberg to S. Curtis, MRK-AFK0185269); 1/10/05 APPROVe Trial Cardiovascular Safety Report draft reflecting comments by K. Horgan, MRK-AFN0093232, at 256 (attached to 1/12/05 email from K. Horgan to H. Quan, MRK-AFN0093231); 1/14/05 APPROVe Trial Cardiovascular Safety Report draft reflecting comments by A. Ko, MRK-AJJ0024204, at 229 (attached to 1/24/05 email from A. Ko to H. Quan, MRK-AJJ0024203); 1/17/05 APPROVe Trial Cardiovascular Safety Report draft, MRK-AIU0057345, at 371 (attached to 1/26/05 letter from P. Huang to B. Harvey* (FDA), MRK-AIU0057340); APPROVe Trial Cardiovascular Safety Report draft, MRK-AFN0096304, at 334 (attached to 1/26/05 email from J. McGovern to B. Oxenius); APPROVe Trial Cardiovascular Safety Report draft, MRK-AFN0096985, at 015 (attached to 1/28/05 email from H. Quan to B. Oxenius et al., MRK-AFN0096981); 1/30/05 APPROVe Trial Cardiovascular Safety Report draft with handwritten comments by H. Quan, MRK-AFN0054274, at 306; 3/15/05 APPROVe Trial abbreviated Clinical Study Report, Cardiovascular Safety Report, MRK-I8940100962, at 994.

²⁶⁴ 3/15/05 APPROVe Trial abbreviated Clinical Study Report, Cardiovascular Safety Report, MRK-I8940100962, at 969.

Executive Summary	Conclusion
study and increased risk with rofecoxib therapy after approximately 18 months. The test for the proportional hazards was statistically significant, indicating different relative risks over time. ²⁶⁵	-- First 18 months consistent with prior placebo-controlled and non-naproxen controlled data. ²⁶⁶

The final March 15, 2005 Cardiovascular Safety Report had identical language in the Executive Summary and similar language in the Conclusion section.²⁶⁷ The synopsis to the March 15, 2005 abbreviated Clinical Study Report stated:

Examination of the Kaplan-Meier plot for the cumulative proportion of patients with [confirmed thrombotic] events revealed similarity of risk between treatments for the first 18 months of the study and significantly increased risk with rofecoxib versus placebo beginning after 18 months.

Notably, results for confirmed APTC events and for investigator reported events were similar to those for confirmed thrombotic events.²⁶⁸

5. MRL's Modeling of Hazard Ratios of Cardiovascular Thrombotic Events Over Time in the APPROVe Trial.

As discussed above, MRL scientists' analyses of the APPROVe Trial cardiovascular hazard ratio had:

²⁶⁵ 1/17/05 APPROVe Trial Cardiovascular Safety Report draft, MRK-AIU0057345, at 346 (attached to 1/26/05 letter from P. Huang to B. Harvey* (FDA), MRK-AIU0057340).

²⁶⁶ 1/17/05 APPROVe Trial Cardiovascular Safety Report draft, MRK-AIU0057345, at 394 (attached to 1/26/05 letter from P. Huang to B. Harvey* (FDA), MRK-AIU0057340)

²⁶⁷ 3/15/05 APPROVe Trial abbreviated Clinical Study Report, Cardiovascular Safety Report, MRK-I8940100962, at 0969, 0020.

²⁶⁸ 3/15/05 APPROVe Trial abbreviated Clinical Study Report, Synopsis, MRK-I8940100731, at 740.

- established through testing the proportionality of hazards that the hazard ratio was not constant throughout the course of the APPROVe Trial;
- provided visual evidence through Kaplan-Meier plots suggesting that the hazard rates for Vioxx and placebo were similar for the first 18 months of the trial and then diverged thereafter; and
- provided some statistical evidence, through 6-month and 18-month segment analyses, to support the visual evidence suggesting that the hazard ratio increased beginning at approximately 18 months.

The limited data available, however, made it difficult to draw more detailed conclusions about the precise way in which the hazard ratio changed over time. As a result, contemporaneously with the preparation of the APPROVe manuscript and the abbreviated Clinical Study Report, Merck statisticians undertook an effort to “model” the cardiovascular hazard ratio in the APPROVe Trial over time by identifying various mathematical models that fit the known data and explained its progression so that inferences about the data might be drawn. This section describes the purpose of modeling analyses in general, the analyses that the Company undertook in order to further understand how the hazard ratio changed over time, and the results of those analyses.

a. Overview of statistical modeling.

Statistical modeling is a means of describing and understanding data mathematically. It may be used, among other things, to help statisticians extrapolate or draw inferences from a limited pool of data. In the context of the APPROVe Trial, MRL statisticians sought to identify appropriate models – in essence, mathematical equations – that “fit” the data with the goal of furthering their understanding of the relation of Vioxx

and placebo to the development of cardiovascular events and the constancy (or lack thereof) of the hazard ratio.

Statisticians rely on an array of standard pre-existing models that, over time, have come to be accepted as providing reasonably accurate representations of certain types of phenomena. An example of such a model is the bell shaped curve (the normal distribution) which is used often to model medical variables such as total cholesterol and systolic blood pressure. For investigating and modeling the time it takes to develop specified events (such as heart attacks), the Cox proportional hazard model and an alternative model called the Weibull model are often used.

The first step in the modeling process is often to examine how well a model fits the data. This can be accomplished with the aid of “model fit” measures (or “fit statistics”) and graphical techniques. Fit statistics indicate the degree to which the model matches or fits the data. They can be either associated with formal statistical tests or simple numerical indexes.

By analyzing the value of the fit statistics and graphical plots a statistician can conclude which of the several models being investigated is the most appropriate to represent the phenomenon under investigation. In selecting a model or models for the data, the joint use of graphical and fit statistics is the usual modality. In addition substantive concerns such as biological plausibility are essential.

Once a model is selected, the parameters of the model can be estimated. For example, with systolic blood pressure the mean blood pressure can be estimated from the data. For the APPROVe study, once a model such as the Cox model is deemed

appropriate, parameters such as the hazard ratio can be estimated. In addition, deviation of data from proportional hazard ratio (or constant hazard) can be evaluated. For the Cox model the latter can be achieved by examining an interaction covariate consisting of the product of time and treatment.

The final step in a modeling exercise is often to identify the most appropriate model and the inclusion of appropriate covariates (such as treatment and the interaction of treatment and time). Fit statistics and graphical methods, in addition to statistical significance tests, are usually employed.

When interpreting the plots produced by the chosen models, it is important to bear in mind two considerations: First, the plots represent *estimated* values, not actual observed values. The accuracy of these estimates may be influenced by the amount of data available for analysis. (In the case of the APPROVe Trial, for example, a mere 70 confirmed thrombotic events over the course of three years – 45 on Vioxx and 25 on placebo – were insufficient to produce wholly reliable estimates of the hazard ratio over time.)²⁶⁹ Second, the model chosen influences to a large extent certain characteristics of the curve produced. For example, in a model assessing changes in a hazard ratio over time, a model using a single time-varying covariate such as the interaction of time and treatment almost always shows change beginning immediately because time starts

²⁶⁹ Dr. Ng performed her modeling analyses on preliminary data from August 2004, which included 70 confirmed thrombotic events. 1/27/05 memorandum from J. Ng and H. Wang to R. Bain et al., MRK-AOY023523, at 28. The final data from December 2004 on which the APPROVe manuscript was based included two additional events, one on Vioxx and one on placebo. See Bresalier* RS, Sandler* RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med. 2005;352:1092-102, Table 3 at 1098.

increasing immediately, even though the actual phenomenon may not involve an immediate change.

b. Conduct of the modeling.

In October or November of 2004, Mr. Jim Bolognese or Dr. Murray Selwyn asked Dr. Jennifer Ng, a statistician in the Biostatistics and Research Decisions Department, to undertake, based on preliminary APPROVe Trial data from August 2004, exploratory analyses specified in the Statistical Data Analysis Plan for Protocol 203.²⁷⁰ Specifically, Dr. Ng was assigned to model how the hazard ratio changed over time.²⁷¹ At that time, MRL planned to use the results of Dr. Ng's work to create back-up slides to be presented in response to potential questions at the February 16 to 18, 2005 joint meeting of the FDA's Arthritis Advisory Committee and Drug Safety and Risk Management Advisory Committee convened to discuss the safety and efficacy of selective and non-selective Cox-2 inhibitors, including the results of the APPROVe Trial.²⁷² Witnesses stated that Dr. Ng's analyses were not intended to confirm the findings resulting from the proportionality of hazards analyses conducted earlier by Dr. Quan.

²⁷⁰ The Statistical Data Analysis Plan for Protocol 203 (analysis of combined cardiovascular event data from the APPROVe, ViP and VICTOR Trials) stated that "[a]dditional diagnostic steps based on residuals and HR [hazard ratio] will be performed to examine model fitting if the previously described analyses [e.g., the Logarithm of Time Test] indicate issues that warrant further assessment." Protocol 203 Statistical Data Analysis Plan, MRK-AAB0083231, at 56.

²⁷¹ Another Merck statistician, Dr. Hongwei Wang, assisted Ng with one of the calculations she used. See, e.g., 1/6/05 email from H. Wang to J. Bolognese and J. Ng, MRK-AFO0283723. Although Dr. Quan was the unblinded statistician for the APPROVe Trial, Dr. Ng received this assignment because Dr. Quan was fully occupied with other duties.

²⁷² 1/3/05 email from M. Selwyn to J. van Adelsberg et al., MRK-AFK0198025.

According to Dr. Ng, she began her assignment by holding discussions with other Merck statisticians, including Mr. Bolognese, Dr. Selwyn, and Dr. Quan, in order to narrow down the potential models and variables to those most likely to be appropriate for her task.²⁷³ Thereafter, Dr. Ng and Dr. Beilei Xu, an MRL computer programmer, modeled the APPROVe Trial data using SAS statistical data analysis computer software.

Among others, Dr. Ng tested a variety of iterations of the Cox proportional hazards model discussed above, including the model used to perform the Linear Time Test²⁷⁴ that Dr. Quan had used in preparing the September 2004 report to the APPROVe Trial External Safety Monitoring Board and the abbreviated Clinical Study Report for the APPROVe Trial. Dr. Ng also tested the model used to perform the Logarithm of Time Test. Although the Logarithm of Time Test was specified in the Statistical Data Analysis Plan for Protocol 203 as “[t]he primary method for testing the proportional hazards assumption,”²⁷⁵ Dr. Ng did not refer to any data analysis plan in connection with her work and did not understand her work to have any relationship to the drafting of the APPROVe article.

Dr. Ng analyzed the data generated by the SAS software for each model to determine which models “fit” the data well and which did not. To do so, she focused on

²⁷³ In addition to these discussions, the Company solicited input from Dr. Scott Zeger*, a member of Merck’s Board of Scientific Advisors. 11/10/04 email from J. Bolognese to S. Zeger*, MRK-AGO0072511. According to Dr. Zeger*, he requested that the plots show the 95% confidence intervals surrounding the estimated values.

²⁷⁴ 1/27/05 memorandum from J. Ng and H. Wang to R. Bain et al., MRK-AOY0023523, at 524.

²⁷⁵ Protocol 203 Statistical Data Analysis Plan, MRK-AAB0083231, at 55.

the “fit statistics” described above. Although each model also generated a proportionality p-value, as well as approximately 10 other pages of output, Dr. Ng did not recall reviewing any p-values reported by the software and did not regard them as relevant to her work.

c. Results of the modeling.

Dr. Ng summarized the results of her analyses in a memorandum dated January 27, 2005 to Dr. Bain, Dr. Scott Zeger*, Dr. James Neaton*, Mr. Bolognese, Dr. Quan, and Dr. Selwyn (the “Ng Memorandum”). The Ng Memorandum included numerous analyses, including diagnostic tests and an assessment of the fit of various parametric models, semi-parametric models (such as the Cox proportional hazards model) and non-parametric models.²⁷⁶ In addition, the Ng Memorandum attached several diagnostic plots and several plots that showed application of the models to the APPROVE Trial data.²⁷⁷

Dr. Ng concluded that a version of the Cox proportional hazards model using a single covariate of treatment*time fit the data better than the Linear Time Test, the

²⁷⁶ 1/27/05 memorandum from J. Ng and H. Wang to R. Bain et al., MRK-AOY0023523, at 23-26. Semi-parametric models, like the Cox proportional hazards model, do not assume any particular form of probability distribution for survival times. Collett* D. Modelling Survival Data in Medical Research. London, UK: Chapman & Hall; 1994, at 54.

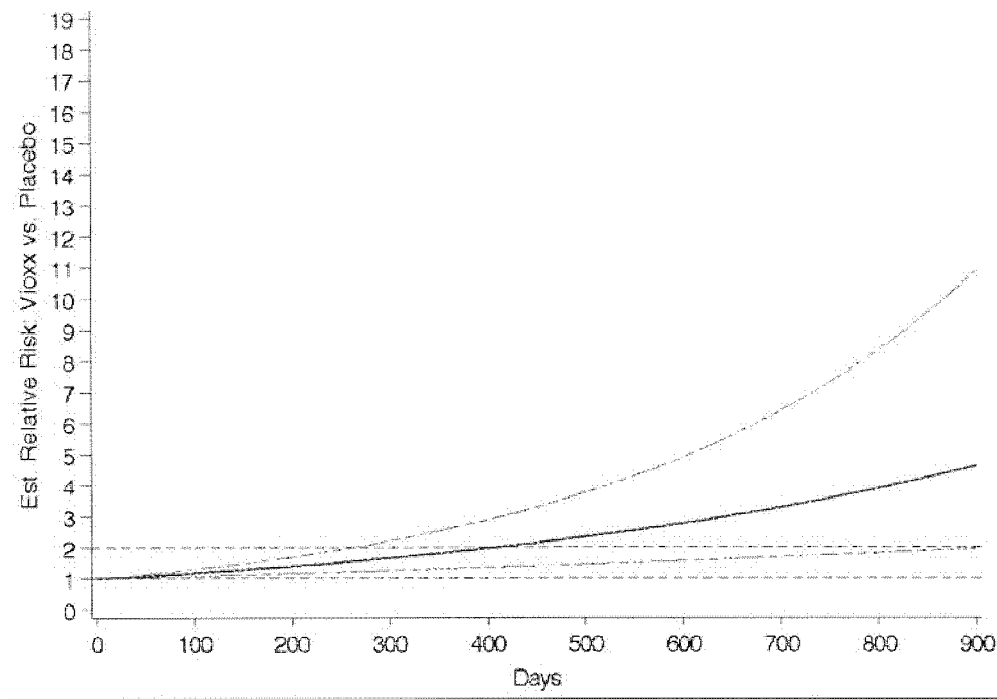
²⁷⁷ 1/27/05 memorandum from J. Ng and H. Wang to R. Bain et al., MRK-AOY0023523, at 28-54. In addition, the Ng Memorandum also included analyses of the data from Merck’s Alzheimer’s Disease Trials (discussed in Appendix Q). These analyses were included for comparison purposes. 1/27/05 memorandum from J. Ng and H. Wang to R. Bain et al., MRK-AOY0023523, at 26-27, 56-72.

Logarithm of Time Test or any other semi-parametric model she tested.²⁷⁸ Figure 7

below shows the hazard ratio plot resulting from that model for the confirmed thrombotic events endpoint in the APPROVe Trial.

Figure 7

Estimated Relative Risk With 95% Confidence Interval for Confirmed Thrombotic Events in APPROVe Trial – Cox Proportional Hazards Model With Treatment*Time As the Factor²⁷⁹



²⁷⁸ 1/27/05 memorandum from J. Ng and H. Wang to R. Bain et al., MRK-AOY0023523, at 25. As described above, the model used in the Linear Time Test included terms for treatment and treatment*time, and the model used in the Logarithm of Time Test included terms for treatment and treatment*log(time). The best-fitting model in Dr. Ng's analysis differed from both of these models because it did not contain a term for treatment (also known as the "main effect"). This model had only one term – the same interaction term from the model used in the Linear Time Test (treatment*time).

²⁷⁹ 1/27/05 memorandum from J. Ng and H. Wang to R. Bain et al., MRK-AOY0023523, at 37.

The solid line in Figure 7 represents the relative risk of experiencing a thrombotic event associated with Vioxx relative to placebo at each point in time shown on the horizontal axis. The two curved dotted lines represent the upper and lower bounds of the 95% confidence interval for the relative risk.²⁸⁰ Although the Ng Memorandum did not state a conclusion with respect to the relative fit of the models used for the Linear Time Test and the Logarithm of Time Test, the fit statistics included in the memorandum indicated that the model used in the Linear Time Test, though not the best fit overall, fit the data better than did the model used in the Logarithm of Time Test.²⁸¹

Although Dr. Quan participated in the modeling discussions with Dr. Ng and apparently received a copy of the Ng Memorandum, his decision about which model to use to test the proportionality of hazards assumption in the course of preparing the APPROVe Trial abbreviated Clinical Study Report and working on the APPROVe article was not influenced by Dr. Ng's independent work. Dr. Ng and Dr. Quan both stated that they did not discuss Dr. Ng's fit analyses or conclusions during the course of her work. Thus, any p-values included in the output from any of the models Dr. Ng fit to the data – p-values that were not included in her memorandum²⁸² – would not have been available to those drafting either document. Finally, the language in the APPROVe manuscript

²⁸⁰ The two straight horizontal dotted lines merely indicate the points along the vertical axis where the risk of Vioxx is equal to that of placebo (at 1) and twice that of placebo (at 2).

²⁸¹ 1/27/05 memorandum from J. Ng and H. Wang to R. Bain et al., MRK- AOY0023523, at 25.

²⁸² Indeed, the only p-values referenced in the Ng Memorandum were calculated based on preliminary data by Dr. Quan before Dr. Ng began her work. According to Mr. Bolognese, the purpose of Dr. Ng's work was not to confirm those results.

identifying the method used to test the proportional hazards assumption did not appear until after Dr. Ng circulated her memorandum. Neither Dr. Quan nor Mr. Bolognese, both recipients of the Ng Memorandum who played a role in drafting the APPROVe manuscript, considered the Ng Memorandum's conclusions as relevant to the drafting of the APPROVe article.

d. Reactions of Merck consultants.

The Company solicited reactions to a near-final draft of the Ng Memorandum from two of its external consultants. On January 15, 2005, Dr. Selwyn forwarded a version of the memorandum to Dr. James Neaton*, a member of the APPROVe External Safety Monitoring Board, and Dr. Scott Zeger*, Chairman of the Biostatistics Department at the Johns Hopkins Bloomberg School of Public Health and a member of Merck's Board of Scientific Advisors.²⁸³ Dr. Selwyn's email explained that the analyses might be presented at the FDA Advisory Committee Meeting in response to potential questions.²⁸⁴

On January 31, 2005, Dr. Zeger* provided his comments on the Ng Memorandum by email soon thereafter,²⁸⁵ stating that "the work [was] thorough and should be highly regarded within the company." He also stated:

1. Most importantly, these data comprising 70 events over 3 years will NEVER ESTABLISH THE "TRUE FORM" of the relative risk model. . . .

²⁸³ 1/15/05 email from M. Selwyn to S. Zeger* and J. Neaton*, MRK-AFO0285777-78.

²⁸⁴ 1/15/05 email from M. Selwyn to S. Zeger* and J. Neaton*, MRK-AFO0285777-78.

²⁸⁵ 1/31/05 email from S. Zeger* to J. Bolognese and J. Neaton*, MRK-AGO0069292.

2. There is reasonably compelling evidence that the relative risk is not constant across all time.

3. The many different parametric, semi-parametric and non-parametric estimates of the relative risk are useful in the background but are too numerous to bring clarity if presented outside of a statistical meeting.

* * *

5. Avoid attempting to conclude that the data prove there is no elevated risk until 18 months. The confidence intervals can not [sic] rule out 1.5-2.0 as early as 6 months.²⁸⁶

According to Dr. Zeger*, the data from the APPROVe Trial could not prove that there was no elevated risk associated with Vioxx during the first 18 months of use, but they did suggest that a change occurred at that time. Dr. Zeger suggested that the Company limit any presentation of the models in the memorandum to two or three different models. He suggested the plot based on the model used in the Linear Time Test as well as the changepoint model that broke the data at 18 months.²⁸⁷

Dr. Neaton* also provided comments on Dr. Ng's Memorandum. He stated that he thought the analyses were "interesting"²⁸⁸ and suggested that Dr. Selwyn and Mr. Bolognese consider pooling the APPROVe Trial data with data from other Vioxx studies:

Given the small [numbers], I would pool the results from Approve and the Alzheimers [sic] studies. Would also

²⁸⁶ 1/31/05 email from S. Zeger* to J. Bolognese and J. Neaton*, MRK-AGO0069292.

²⁸⁷ 1/31/05 email from S. Zeger* to J. Bolognese and J. Neaton*, MRK-AGO0069292.

²⁸⁸ 1/29/05 email from J. Neaton* to M. Selwyn, MRK-AFO0285777. Dr. Neaton* also stated that he preferred, due to their "simplicity," (i) the so-called "changepoint" models, which used different equations before and after selected points in time; and (ii) the version of the Cox proportional hazards model with covariates for treatment and treatment*time. 1/29/05 email from J. Neaton* to M. Selwyn, MRK-AFO0285777.

consider adding other trials. In those models, I would treat study as a stratifying variable in the Cox model, i.e., pool over trials to obtain the estimates for the treatment and treatment* time parameters. If the point is more to assess whether the other trials are consistent with Approve, you might assess that directly with an interaction term for Approve versus other studies.²⁸⁹

Dr. Neaton's* suggestion that MRL pool the data from multiple trials identified two different possible benefits of doing so. First, it would increase the amount of data available for analysis, and, therefore, increase the reliability of the estimates produced by the statistical model. Second, it would permit MRL scientists to test whether the APPROVE Trial was qualitatively different from the other Vioxx trials in which there was no statically significant increased risk on Vioxx as compared to placebo. Testing this proposition would involve including in the Cox proportional hazards model "an interaction term for Approve versus other studies" with treatment. It does not appear that Merck performed the pooled analyses that Dr. Neaton* suggested in connection with its modeling effort.

It does not appear that any of the Ng Memorandum's analyses were presented at the subsequent FDA Advisory Committee Meeting for which they had been conducted.

6. May 2006 Discovery of the Error
in the APPROVE Article and Clinical Study Report.

a. Introduction.

On May 22, 2006, MRL scientists discovered that the article about the cardiovascular data from the APPROVE Trial and the abbreviated Clinical Study Report

²⁸⁹ 1/29/05 email from J. Neaton* to M. Selwyn, MRK-AFO0285777.

for the APPROVe Trial contained an error: although the documents stated that the test of proportionality of hazards was performed using the Logarithm of Time Test, the reported p-value of 0.01 was in fact derived from the Linear Time Test, and the Logarithm of Time Test yielded a p-value of 0.07. This Section discusses: (i) how the error was discovered; (ii) the Company's preliminary investigation into how the error occurred and whether the error should change the Company's conclusions about the APPROVe Trial; and (iii) the Company's May 30, 2006 press release announcing the error.

b. Discovery of the error.

As discussed in Section C of this Appendix, in May 2006, MRL conducted analyses involving recently collected data on the incidence of thrombotic events among patients from the APPROVe Trial occurring more than 14 days after they had discontinued treatment. On May 11, 2006, MRL submitted to the FDA preliminary results of these analyses, and, for the remainder of the month, MRL statisticians, including Mr. Thomas Cook, finalized and checked the accuracy of these results.

To check the analyses that had been performed, Mr. Cook wrote his own programs for analyzing the data, ran his programs on the data, and compared his results to the preliminary results that had been obtained in early May 2006. When Mr. Cook performed the test of the proportionality of hazards on the combined on- and off-drug data from the APPROVe Trial, he noticed that the result he obtained differed from the

one that had been obtained in the course of the preliminary analysis.²⁹⁰ This led Mr. Cook to the discovery that the program he had written and the program that had been used for the preliminary analysis used different methods for testing the proportionality of hazards: while Mr. Cook's program used the Logarithm of Time Test, the program that his MRL colleagues had used for the preliminary May 11, 2006 analysis used the Linear Time Test.²⁹¹

Because MRL had used the same computer program to test the proportionality of hazards on the APPROVe base study data (*i.e.*, the APPROVe Trial data that were published in the APPROVe article and submitted on June 6, 2005 to the FDA), Mr. Cook sought to determine whether the abbreviated Clinical Study Report and the APPROVe article similarly reported the result of the Linear Time Test. To do so, Mr. Cook asked Dr. Beilei Xu, an MRL programmer, to perform both the Linear Time Test and the Logarithm of Time Test on the APPROVe base study data.²⁹² For the confirmed thrombotic endpoint, the Linear Time Test yielded $p=0.01$ (the result that was reported in the abbreviated Clinical Study Report and the APPROVe article), whereas the Logarithm

²⁹⁰ See 5/24/06 email from T. Cook to R. Bain *et al.*, MRK-ARQ0005023 (stating that, on May 22, 2006, all the hazard plots in the summary of results from the APPROVe extension were changed to use the Logarithm of Time Test for proportionality).

²⁹¹ 5/22/06 (11:20 a.m.) email from T. Cook to N. Braunstein, MRK-AIU0331770 ("Following our discussion this morning about the proportional hazards plots, I checked the program and discovered that a non-standard test [*i.e.*, the Linear Time Test] results were included in the hazard plots. They have been rerun with the standard test [*i.e.*, the Logarithm of Time Test]. I will update the summary report and send it to you later today.").

²⁹² 5/25/06 memorandum from J. Bolognese and T. Cook to R. Bain *et al.*, MRK-S0420112192.

of Time Test yielded $p=0.07$.²⁹³ Mr. Cook therefore concluded that the abbreviated Clinical Study Report and the article contained an error.

Subsequently, MRL statisticians determined that the results reported in Dr. Quan's September 13, 2004 report to the APPROVe External Safety Monitoring Board, like the abbreviated Clinical Study Report and the APPROVe article, reported results of the Linear Time Test, not the Logarithm of Time Test.²⁹⁴ Tables 14 and 15 below, respectively, summarize the results of both tests for the confirmed thrombotic and APTC event endpoint and the results of the Linear Time Test for the APTC event endpoint.²⁹⁵ (It appears that the Logarithm of Time Test has not been conducted on the APTC event endpoint.)

²⁹³ 5/25/06 memorandum from J. Bolognese and T. Cook to R. Bain et al., MRK-S0420112192.

²⁹⁴ "APPROVe: Assessment," MRK-AFO0300154, at 54-55 (attached to 6/26/06 letter from P. Kim, "An Open Letter from Merck," MRK-AFO00300152).

²⁹⁵ 5/25/06 memorandum from J. Bolognese and T. Cook to R. Bain et al., MRK-S0420112192 (referring to the $p=0.014$ and $p=0.071$ results for the final APPROVe Trial data); Draft letter to the editor, MRK-AQU0000031, at 32 (referring to the $p=0.048$ result for the preliminary APPROVe Trial data) (attached to 6/8/06 email from J. Bolognese to T. Reiss, MRK-AQU0000030); 9/13/04 Pre-Meeting Report from H. Quan to APPROVe ESMB, MRK-AGO0029517, at 537 (referring to the $p=0.0056$ result for the preliminary APPROVe Trial data).

Table 14

APPROVe Trial – Proportionality P-Values
for Confirmed Thrombotic Events in the APPROVe Trial

	Linear Time Test	Logarithm of Time Test
Preliminary Data	p=0.0056	p=0.048
Final Data	p=0.014	p=0.071

Table 15

APPROVe Trial – Proportionality P-Values
for APTC events in the APPROVe Trial²⁹⁶

	Linear Time Test	Logarithm of Time Test
Preliminary Data	p=0.0235	Unknown
Final Data	p=0.119	Unknown

- c. Investigation of error and its impact on Merck’s conclusions concerning the APPROVe Trial cardiovascular data.

Upon learning of the error on May 22, 2006, Mr. Cook notified Mr. Bolognese and Dr. Braunstein, who in turn notified Dr. Kim, Dr. Honig, Dr. Gertz, Dr. Reiss,

²⁹⁶ 9/13/04 Pre Meeting Report from H. Quan to APPROVe ESMB, MRK-AGO0029517, at 537 (referring to the p=0.0235 result for the preliminary APPROVe Trial data); 3/15/05 APPROVe Trial abbreviated Clinical Study Report, Cardiovascular Safety Report, MRK-I8940100962, at 0994 (referring to the p=0.119 result for the final APPROVe Trial data).

Dr. Bain, Ms. Lahner, and Dr. Huang,²⁹⁷ Messrs. Clark and Frazier were notified on May 22, 2006 as well,²⁹⁸ and the error was disclosed to the full Board of Directors and discussed at the Board's regularly scheduled May 23, 2006 meeting. The Board of Directors decided that the error should be disclosed publicly as soon as the Company had all the relevant facts. Dr. Bain, Dr. Reiss and Mr. Bolognese also notified the external authors of the article at the regularly scheduled May 25, 2006 meeting of the APPROVe Administrative Committee (discussed in more detail in Section D of this Appendix).²⁹⁹

After learning about the error, MRL scientists, including Dr. Reiss, Dr. Huang, Mr. Bolognese, and Mr. Cook, investigated: (i) how the error occurred, and (ii) what effect the discovery of the error had on the Company's conclusions about the APPROVe Trial.

MRL scientists began by contacting Dr. Quan, the statistician who had performed the tests and prepared the abbreviated Clinical Study Report for the APPROVe Trial. (Dr. Quan had left Merck in 2005, and, in May 2006, was working for another pharmaceutical company). Upon being informed of the error, Dr. Quan stated that he believed that he had used the Logarithm of Time Test and that the APPROVe article and the abbreviated Clinical Study Report had described the test accurately. It was therefore

²⁹⁷ 5/22/06 (7:13 p.m.) email from N. Braunstein to P. Kim et al., MRK-AFO0309953 ("URGENT meeting tonight 9PM"; "Please listen to MVX [i.e., voicemail] for details.").

²⁹⁸ 5/23/06 (12:12 a.m.) email from J. Lahner to N. Braunstein et al., MRK-AFO0309953 ("Please call in. Ken Frazier will be joining.").

²⁹⁹ Minutes of 5/25/06 APPROVe Administrative Committee meeting, MRK-ASP0000001.

unclear how the error occurred. As discussed in Section B.3.d of this Appendix, subsequently, however, Dr. Quan recalled the circumstances that likely led to the error. In a June 8, 2006 email to the APPROVe article co-authors (sent a week after the Company issued a press release announcing the error, which is discussed below), Dr. Quan stated that he recalled that the software program that MRL statisticians had written for testing the proportionality of hazards originally used the Logarithm of Time Test but that, after testing the program in the summer of 2004, he determined that the Linear Time Test fit the APPROVe Trial data better – *i.e.*, was more likely to detect a change in the hazard ratio if there was one – than the Logarithm of Time Test.

Specifically, Dr. Quan wrote:

As I could recall, the log (time) was used originally when we prepared the program. But when we tested the program for preparing package for ESMB, I found the linear time fitted the data better and gave result that was more consistent with the results from other analyses [*i.e.*, the Kaplan-Meier plot and six-month interval analyses, Figures 1 and 2 and Tables 2 and 3 above]. The use of linear time is also appropriate. Thus, I switched to linear time in order not to hide any thing from ESMB. Unfortunately, since I could not modify the APPROVe DAP and didn't document this in some document, I forgot all about this when I rushed through the final analysis. Thus, in my mind, I always thought we used log(time).³⁰⁰

³⁰⁰ 6/8/06 email from H. Quan to J. Bolognese *et al.*, MRK-ARQ0007271. Dr. Quan also stated that, when he first ran the Logarithm of Time Test, he obtained a p-value that was greater than 0.05, which was higher than he had expected it to be based on the other analyses he had conducted, most notably the Kaplan-Meier plots, which seemed so clearly to indicate a change in the hazard ratio.

MRL scientists also reviewed the January 27, 2005 Ng Memorandum.³⁰¹ As described more fully in Section B.5 of this Appendix, the modeling described in the Ng Memorandum was an effort to estimate the relative risk of Vioxx versus placebo at each point in time during the course of the APPROVe Trial. In the memorandum, Dr. Ng concluded based on preliminary data that the Cox proportional hazards model including a single term for the interaction of time and treatment (which is distinct from the Linear Time Test) was the most representative of the data from the trial.³⁰² The fit statistics reflected in the body of the Ng Memorandum indicated that the Linear Time Test fit data better than the Logarithm of Time Test.³⁰³

Mr. Cook repeated the modeling work performed by Dr. Ng in late 2004 on the August 2004 data, applying the same models to the final (December 2004) dataset.³⁰⁴ Mr. Cook's analyses confirmed that the Cox model using covariates for treatment and the interaction of treatment and time (*i.e.*, the Linear Time Test) fit the data better than the Cox model using covariates for treatment and the interaction of treatment and the logarithm of time (*i.e.*, the Logarithm of Time Test).³⁰⁵

³⁰¹ 5/23/06 email from J. Bolognese to T. Cook and R. Bain *et al.*, MRK-ARQ0004234 (attaching chart of fit statistics from Ng Memorandum).

³⁰² 1/27/05 memorandum from J. Ng and H. Wang to R. Bain *et al.*, MRK-AOY0023523, at 25.

³⁰³ 1/27/05 memorandum from J. Ng and H. Wang to R. Bain *et al.*, MRK-AOY0023523, at 25.

³⁰⁴ 5/24/06 email from T. Cook to J. Bolognese *et al.*, MRK-ARQ0004233.

³⁰⁵ 5/24/06 email from T. Cook to J. Bolognese *et al.*, MRK-ARQ0004233.

Mr. Cook also concluded that, based on the final (December 2004) data, the best fitting model was not the Cox proportional hazards model including a single term for the interaction of time and treatment (i.e., the model that Dr. Ng had identified as the best fitting model), but rather a changepoint model “that breaks the data at 17 months,”³⁰⁶ i.e., models the first 17 months using one equation and the remainder of the trial using another.

Mr. Cook also calculated the p-value for a similar changepoint model breaking the data at 18 months. The p-value was lower than 0.05 ($p=0.016$), leading Mr. Cook to conclude, as had the APPROVe article and the abbreviated Clinical Study Report, that “the hazard rate ratio in the first 18 months is significantly different than the hazard rate ratio after 18 months.”³⁰⁷

Additionally, at Dr. Huang’s request,³⁰⁸ Mr. Bolognese prepared a memorandum, dated May 25, 2006, regarding the effect, if any, of the error on the Company’s conclusion that, in the APPROVe Trial, the relative risk for confirmed thrombotic events on Vioxx versus placebo changed over time. Mr. Bolognese’s memorandum, prepared with input from Mr. Cook,³⁰⁹ concluded that, notwithstanding the error, the existing data “support[ed] an overall conclusion of non-proportionality of treatment hazards over time

³⁰⁶ 5/24/06 email from T. Cook to J. Bolognese et al., MRK-ARQ0004233.

³⁰⁷ 5/24/06 email from T. Cook to J. Bolognese et al., MRK-ARQ0004233.

³⁰⁸ 5/24/06 email from J. Bolognese to T. Cook et al., MRK-ARQ0004415 (“Phil [Huang] asked me to draft [the memorandum].”).

³⁰⁹ 5/24/06 email from T. Cook to J. Bolognese, MRK-ARQ0004415 (“Nice memo. I suggested a qualification to the first conclusion.”).

in the APPROVe [abbreviated Clinical Study Report confirmed thrombotic event] dataset.”³¹⁰

In the memorandum, Mr. Bolognese noted that the use of the Linear Time Test (the test that produced the $p=0.01$ result reported in the APPROVe article and the abbreviated Clinical Study Report) to assess proportionality of hazards was “appropriate” because (i) “[t]here was no Statistical Analysis Plan (SAP) for the confirmed thrombotic CV data from APPROVe alone”; and (ii) “[a]lthough the SAP for Protocol 203 specified that the primary assessment of proportional hazards would be [the Logarithm of Time Test], a detailed assessment of relationship of hazard ratio with time” – [i.e., the analyses memorialized in the January 27, 2005 Ng Memorandum discussed in Section B.5 of this Appendix] – “revealed a better fit of the [data with] a model including linear time in comparison to one including [logarithm of time].”³¹¹

d. May 30, 2006 press release.

On May 30, 2006, Merck issued a press release entitled “Merck Corrects Description of a Statistical Method Used in APPROVe Study: Study Results Unchanged.”³¹² The press release stated that: (i) Merck was correcting the APPROVe article’s description of the method used to test the proportionality of hazards from the Logarithm of Time Test to the Linear Time Test; (ii) the $p=0.01$ result reported in the

³¹⁰ 5/25/06 memorandum from J. Bolognese and T. Cook to R. Bain et al., MRK-S0420112192.

³¹¹ 5/25/06 memorandum from J. Bolognese and T. Cook to R. Bain et al., MRK-S0420112192.

³¹² 5/30/06 Merck press release, “Merck Corrects Description of a Statistical Method Used in APPROVe Study: Study Results Unchanged,” MRK-ARQ0006011.

article was obtained using the Linear Time Test; and (iii) recent tests had shown that the result of the Logarithm of Time Test on final data for the confirmed thrombotic event endpoint yielded a p-value of 0.07.³¹³

The press release also stated that the correction did not change the results of the APPROVe Trial, “in which an increased relative risk for confirmed thrombotic cardiovascular events for VIOXX compared to placebo was observed beginning after 18 months of continuous daily treatment.”³¹⁴ The press release stated that although the Logarithm of Time Test was “specified” in the Statistical Data Analysis Plan for Protocol 203 as the “primary method” for testing the proportional hazards assumption, that data analysis plan called for use of “numerous statistical and graphical methods . . . to assess whether the relative risk of VIOXX compared to placebo was constant over time.”³¹⁵ The press release further stated that Merck’s use of these additional methods established that the relative risk of confirmed thrombotic events changed over time.³¹⁶ Finally, the press release stated that “[t]he results of diagnostic steps specified in the data analysis plan [i.e., the Statistical Data Analysis Plan for Protocol 203] indicate[d] that the

³¹³ 5/30/06 Merck press release, “Merck Corrects Description of a Statistical Method Used in APPROVe Study: Study Results Unchanged,” MRK-ARQ0006011, at 11.

³¹⁴ 5/30/06 Merck press release, “Merck Corrects Description of a Statistical Method Used in APPROVe Study: Study Results Unchanged,” MRK-ARQ0006011, at 11.

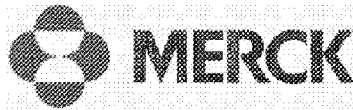
³¹⁵ 5/30/06 Merck press release, “Merck Corrects Description of a Statistical Method Used in APPROVe Study: Study Results Unchanged,” MRK-ARQ0006011, at 11. The press release also attached an excerpt from the Statistical Data Analysis Plan for Protocol 203 where the Logarithm of Time Test was specified. Id. at 14.

³¹⁶ 5/30/06 Merck press release, “Merck Corrects Description of a Statistical Method Used in APPROVe Study: Study Results Unchanged,” MRK-ARQ0006011, at 11-12.

[Linear Time Test] is an appropriate method to assess changes in relative risk over time.”³¹⁷ The press release is reproduced in full below:

³¹⁷ 5/30/06 Merck press release, “Merck Corrects Description of a Statistical Method Used in APPROVe Study: Study Results Unchanged,” MRK-ARQ0006011, at 11.

The press release also stated that Merck “intend[ed] to retain an independent statistical expert to review data and analyses from the APPROVe study.” In early June 2005, Drs. Baron* and Konstam* initiated discussions with Dr. David DeMets*, the chairman of the Department of Biostatistics and Medical Informatics at the University of Wisconsin, regarding the possibility of his performing that role. 6/7/06 email from J. Baron* to D. DeMets*, MRK-AFS0038812. Dr. DeMets* expressed interest in working on this project and suggested that the project would entail: (i) “reproducing the NEJM article”; (ii) “following the Approve Statistical Analysis Plan”; and (iii) “doing the analysis the way [Dr. DeMets*] might do it.” Id.



Press Release

Media Contacts: Kent Jarrell
(202) 230-1833

Investor Contact: Graeme Bell
(908) 423-5185

Casey Stavropoulos
(908) 423-5125

Merck Corrects Description of a Statistical Method Used in APPROVe Study

Study Results Unchanged

WHITEHOUSE STATION, N.J., May 30, 2006 – Merck & Co., Inc. is correcting its prior description of one of the statistical methods used to analyze certain data in the APPROVe study published in 2005, and has notified the study authors, the *New England Journal of Medicine* (*NEJM*) and regulatory authorities. Merck recently discovered the need for this correction while reviewing the preliminary analyses of the off-drug extension data for the APPROVe study. Merck believes that this correction does not change the results of the APPROVe study, in which an increased relative risk for confirmed thrombotic cardiovascular events for VIOXX compared to placebo was observed beginning after 18 months of continuous daily treatment. This correction is unrelated to the recently announced preliminary analysis of the off-drug extension of the APPROVe study.

The VIOXX cardiovascular data analysis plan called for numerous statistical and graphical methods to be used to assess whether the relative risk of VIOXX compared to placebo was constant over time or if it changed over time (see attached). The use of the variable, logarithm of time, was an element in the primary method specified. The reference to logarithm of time in the description of methods published in the *NEJM* and submitted to regulatory agencies was in error. The reported result (p-value = 0.01) came from a statistical model using linear time, not logarithm of time. Recent tests show that the result using logarithm of time has a p-value = 0.07. The results of diagnostic steps specified in the data analysis plan indicate that the linear test is an appropriate method to assess changes in relative risk over time.

As specified in the analysis plan, Merck used additional analytical and graphical methods to evaluate whether the relative risk changed over time. These methods included a

- more -

Appendix R

- 138 -

Kaplan-Meier plot that showed similar curves for placebo and VIOXX during the first 18 months, which then began to separate at approximately 18 months. Relative risks were also calculated over successive six-month intervals in the study. Over the 36-month period of the study, the relative risk was lowest in the first three sets of six-month intervals and highest in the last three sets of six-month intervals, again demonstrating changing relative risk over time. The results of the linear time analysis, the Kaplan-Meier plot, and the interval relative risks together confirm that the relative risk changes over time.

The linear time and logarithm of time analyses only test whether relative risk changes over time, they do not test the overall magnitude of relative risk. The overall magnitude of the relative risks and their associated p-values were described correctly.

Merck believes that today's correction does not change the results of the APPROVe study. Nonetheless, Merck intends to retain an independent statistical expert to review data and analyses from the APPROVe study. The study's authors advised Merck that they intend to submit a correction to the *NEJM*.

About the APPROVe Study

APPROVe (Adenomatous Polyp Prevention on VIOXX) was a multi-center, randomized, placebo-controlled, double-blind study designed to evaluate the efficacy of 156 weeks (three years) of treatment with VIOXX 25 mg in preventing recurrence of colorectal polyps in patients with a history of colorectal adenomas. There was no Statistical Analysis Plan (SAP) for the cardiovascular data from APPROVe alone. Merck planned to combine the cardiovascular data from APPROVe with data from two other placebo-controlled studies, VICTOR and ViP. Given the decision to stop the study early, the APPROVe data were analyzed separately.

About Merck

Merck & Co., Inc. is a global research-driven pharmaceutical company dedicated to putting patients first. Established in 1891, Merck currently discovers, develops, manufactures and markets vaccines and medicines to address unmet medical needs. The Company devotes extensive efforts to increase access to medicines through far-reaching programs that not only donate Merck medicines but help deliver them to the people who need them. Merck also publishes unbiased health information as a not-for-profit service. For more information, visit www.merck.com.

- more -

Appendix R

- 139 -

Forward-Looking Statement

This press release (including the attachment) contains "forward-looking statements" as that term is defined in the Private Securities Litigation Reform Act of 1995. These statements are based on management's current expectations and involve risks and uncertainties, which may cause results to differ materially from those set forth in the statements. The forward-looking statements may include statements regarding product development, product potential or financial performance. No forward-looking statement can be guaranteed, and actual results may differ materially from those projected. Merck undertakes no obligation to publicly update any forward-looking statement, whether as a result of new information, future events, or otherwise. Forward-looking statements in this press release should be evaluated together with the many uncertainties that affect Merck's business, particularly those mentioned in the cautionary statements in Item 1 of Merck's Form 10-K for the year ended Dec. 31, 2005, and in its periodic reports on Form 10-Q and Form 8-K, which the Company incorporates by reference.

#

**Excerpted from the MK-0966 CV Outcomes Data Analysis Plan for
Protocol 203 Combining APPROVe, ViP, and VICTOR
(Jan. 9, 2004)**

Check of Model Assumptions

Analytical and graphical methods will be employed to verify the proportional hazards assumption. The primary method for testing the proportional hazards assumption will be by including the factor $\text{treatment} \times \log(\text{time})$ in the model; nonsignificance ($p > 0.050$) of this factor is not inconsistent with proportionality, i.e., constancy of treatment effect over time. The semi-parametric nature of the proportional hazards model does not require other distributional assumptions.

Further, the log HR will be plotted over time by stratifying time into intervals containing approximately the same number of events within each interval. The log HR within each successive 6-month time interval with confidence limits will also be calculated and plotted. Such plots will provide an indication of any time effect on the HR.

Additional diagnostic steps based on residuals and HR will be performed to examine model fitting if the previously described analyses indicate issues that warrant further assessment. For instance, assessment of the log HR over time may also be made using Schoenfeld's partial residuals [10] with respect to each covariate in the fitted model. Visual and test diagnostics of scaled Schoenfeld residuals [11; 12; 13] may be used here to study the PH assumption for the model covariates. Schoenfeld residuals are defined at event times and compute the difference between the observed covariate (treatment group indicator) and its conditional expectation from the fitted Cox PH model. If the PH assumption holds for a covariate in the Cox model, these scaled residuals when plotted over time should be randomly distributed to either side of "zero" line and should not have any discernible trend/pattern over time. Constancy of relative treatment effect over time may also be assessed for each covariate (discrete or continuous) by using the *zph* test [13]. A *zph* test of treatment effect (based on log-transformation) is essentially the same as testing the significance of adding $\text{treatment} \times \log(\text{time})$ as a time-dependent covariate ([13], pp. 144). Thus, a nonsignificant *zph* test may also be used to validate proportionality assumption for associated covariate.

If the nonproportionality is large and real, various approaches may be explored to investigate the causes of nonproportionality: stratification of covariates, partition of the time axis [14; 15; 16], adding a time-dependent covariate, or using a different model. Some of these approaches are discussed with illustrations in the book by Therneau and Grambsch [13] (Sections 6.5 and Chapter 7). The choice of a non-proportional hazard model, such as a piecewise exponential model, that will be a good fit to the data may be assessed from some of the above diagnostic measures.

- more -

Appendix R

- 141 -

References

10. Schoenfeld D. Partial residuals for the proportional hazard regression model. *Biometrika* 1982;69:239-41.
11. Venables WN and Ripley BD. *Modern applied statistics with S-Plus*, Springer-Verlag Inc (Berlin; New York). 1994.
12. Hess K. Graphical methods for assessing violations of the proportional hazards assumption in Cox regression. *Statistics in Medicine* 1995;14:1707-23.
13. Therneau TM and Grambsch PM *Modeling survival data: extending the cox model*. Springer, New York 2000.
14. Martink DO and Austin H. Exact estimates of a rate ratio. *Epidemiology* 1996;7:29-33.
15. Martin DO and Austin II. An exact method for meta-analysis of case-control and follow-up studies. *Epidemiology* 2000;11:255-60.
16. Zelen M. The analysis of several 2x2 contingency tables. *Biometrika* 1971;58:129-37.

#

e. Additional communications.

On May 30, 2006, in connection with the issuance of the press release, Dr. Kim, Mr. Frazier, and Dr. Bain held a teleconference for pharmaceutical analysts and journalists regarding the APPROVe Trial and the announcement of the error.³¹⁸ On the same day, Dr. Robert Silverman, MRL's regulatory liaison with the FDA, notified the FDA about the correction,³¹⁹ and Mr. Bolognese and Dr. Oxenius notified the New England Journal of Medicine about the need for a correction to the published article.³²⁰ Communications between and among authors of the article and editors of the New England Journal of Medicine about the correction are discussed in Section D of this Appendix.

C. Collection and Analysis of APPROVe Off-Drug Data.

1. Introduction.

The APPROVe Trial Protocol specified, as was standard in Merck trials, that data would be collected on serious adverse events, including cardiovascular events, that

³¹⁸ 5/30/06 Merck media advisory: Conference Call and Audio Web Cast, MRK-AFS0040985; APPROVe Media Call Speaker Notes, MRK-AFS0040987-88.

³¹⁹ 5/30/06 letter from P. Huang to B. Rappaport*, MRK-S0420112190 (“[A] telephone conversation between Dr. Robert Silverman of MRL, and Lisa Malandro of FDA . . . occurred on . . . May 30, 2006, in which notification was provided that MRL was correcting its prior description of one of the statistical methods used to analyze certain data in the APPROVe study published in 2005. MRL recently discovered the need for this correction while reviewing the preliminary analysis of the off-drug extension data for the APPROVe study. MRL believes that this correction does not change the results of the APPROVe study, which found an increased relative risk for confirmed thrombotic cardiovascular events for VIOXX™ compared to placebo beginning after 18 months of treatment.”).

³²⁰ 5/30/06 letter from J. Bolognese and B. Oxenius to the editor of the New England Journal of Medicine. MRK-AQU0000001.