How useful are unpublished data from the Food and Drug Administration in meta-analysis?

Catherine H. MacLean, Sally C. Morton, Joshua J. Ofman, Elizabeth A. Roth, Paul G. Shekelle, for the Southern California Evidence-Based Practice Center

Abstract

The goals of this systematic review and meta-analysis were to ascertain whether studies of nonsteroidal anti-inflammatory drugs (NSAIDs) summarized in the FDA reviews are ultimately published, to compare the methodologic and population characteristics of studies summarized in the FDA reviews with those reported in peer reviewed literature, and to compare the pooled relative risk of dyspepsia from NSAIDs in each data source. Summary measures of risk difference were calculated with a random effects model; meta-regression was used to assess the effect of study covariates. Among 37 studies described in the FDA reviews, one was published. Sample size, gender distribution, indication for drug use, and methodologic quality did not vary significantly between the published and FDA data. The pooled risk ratio for dyspepsia obtained using published data (1.21) or FDA data (1.07) did not differ significantly or practically. Data from FDA reviews may be a viable data source for systematic reviews and meta-analyses but only after being subjected to the same methodologic scrutiny as published data. © 2003 Elsevier Science Inc. All rights reserved.

Keywords: Unpublished data; Systematic review; Meta-analysis; Non-steroidal anti-inflammatory drugs; Dyspepsia; Gastrointestinal toxicity

1. Introduction

The inclusion of unpublished data in systematic reviews or meta-analyses is controversial [1,2]. Proponents argue that exclusion of unpublished data may systematically alter the conclusions of meta-analyses, and therefore attempts should be made to include all available evidence. Opponents argue that unpublished data have not gone through the crucible of the peer-review process and are therefore more suspect than published data. One large survey concluded that most experts believe unpublished data should be included as long as the studies can be subjected to the same scrutiny as published data [1]. However, finding unpublished evidence presents a challenge because there does not exist a systematic way of accessing it.

One potentially large and easily accessible systematic repository of unpublished data on the efficacy and safety of pharmaceuticals is the United States Food and Drug Administration (FDA), which produces publicly available reviews of all studies submitted as part of the New Drug Application (NDA) process. However, the content and methodologic quality of these data have not been evaluated, nor has there been an assessment of how many studies in the FDA reviews also appear in the peer-reviewed literature (ie, the extent to which these data are “unpublished”). We offer the first such look, which we undertook as part of a meta-analysis assessing the complications of nonsteroidal anti-inflammatory drugs (NSAIDs).

2. Methods

As part of the regulatory process to approve drugs for use in the United States, sponsors must submit a NDA to the FDA. NDAs contain extensive information about studies conducted to support claims about the safety and efficacy of the drug for its proposed indications. FDA medical officers review these data and prepare reports summarizing that information that are subsequently used to evaluate the merits of the NDA. The original study material is considered proprietary and not available to the public. However, the FDA
reviews are available to the public under the Freedom of Information Act.

This study compares the quantity and quality of data obtained from the FDA with that of published data on the use of selected NSAIDs in adults and compares pooled estimates of the risk of dyspepsia associated with NSAID use based on the FDA data with those based on the published data.

This work was sponsored by Merck & Company through a contract to RAND that specified that the RAND investigators had control over data collection, analysis, report writing, and decisions regarding publication.

2.1. Selection of published articles

As part of a larger study that was designed to evaluate the gastrointestinal toxicities of NSAIDs, we identified randomized controlled trials (RCTs) of NSAIDs through an unrestricted search of MEDLINE, EMBASE, HEALTHSTAR, and BIOSIS covering the years 1966 through 1997 (Table 1). English and non-English language titles, abstracts, and reports were reviewed to identify published studies in which any oral NSAID was administered for more than 4 days, included a placebo group, enrolled subjects 18 years of age or greater, and reported gastrointestinal side-effects by subject. Studies that reported dyspepsia were included in this analysis. We limited this analysis to dyspepsia because this outcome is frequently reported in randomized controlled trials of NSAIDs and hence would permit comparison between published trials and FDA reports that primarily report clinical trial data.

2.2. Selection of studies from NDAs

We searched for all FDA reviews of all NDAs and supplements to NDAs for the following NSAIDs: naproxen, ibuprofen, diclofenac, etodolac, and nabumetone. These drugs were chosen because they are the most commonly prescribed NSAIDs in the United States, collectively comprising about 70% of the US prescription NSAID market [3]. Each FDA review was hand-searched to identify reports of RCTs meeting the same inclusion and exclusion criteria we used to identify the published clinical trials. FDA reviews were obtained through the Freedom of Information Act.

Each FDA review consisted of a stack of approximately 200 to 500 photocopied pages. Some reviews were of poor copy quality, and some had a few pages missing or out of order. Some reviews lacked a table of contents, which made hand-searching more difficult than with published journals. Even when a table of contents was present, it did not list specific clinical trials, but rather directed the reader to the section of the report that included clinical trials. Clinical trials could be summarized on a single page, on several continuous pages, or on discontinuous pages. In essence, this meant that each page had to be individually assessed.

2.3. Assessment of overlap between published and FDA data

We compared each published study that met our inclusion criteria with each FDA study that met our inclusion criteria to determine whether an individual study was reported in both domains. Specifically, we compared the drug and dose characteristics of the published and FDA studies. We then hand-reviewed all studies that “matched” on these criteria to determine whether they represented the same study based upon drug indication and duration of therapy.

2.4. Assessment of population and dosing characteristics

The age and gender of study participants, indications for drug use, and drug dosages used in published studies were compared with those reported in the FDA studies. We determined the indication for drug use and assigned the populations studied to one of the following disease/symptom categories: normal, arthritis, or “other.”

Because NSAIDs have a wide range of dosing possibilities and because it is believed that complications and efficacy increase with higher doses, we compared the total daily NSAID dose used in published studies with those used in FDA studies. We classified each total daily dose of each NSAID used in each study as being low, moderate, or high (Table 2).

2.5. Assessment of methodologic attributes of studies

We compared the methodologic attributes of peer-reviewed published studies with those of the studies reported in the FDA reviews. Specifically, we determined whether or not descriptions of the following were present in each study or FDA summary of a study: randomization, appropriateness of randomization, blinding, appropriateness of blinding, withdrawals and dropouts, and concealment of allocation. Each of these items has been validated as an in-

Table 1
Search strategy

<table>
<thead>
<tr>
<th>Major exploded subject headings</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Nonsteroidal anti-inflammatory agents (NSAIDS)</td>
<td></td>
</tr>
<tr>
<td>Or Prostaglandin synthase inhibitor/s and cyclooxygenase inhibitor/s</td>
<td></td>
</tr>
<tr>
<td>And gastrointestinal diseases or esophageal diseases or digestive system diseases</td>
<td></td>
</tr>
<tr>
<td>2. Digestive disease/s, digestive complication/s, gastrointestinal complication/s, dyspepsia, peptic or esophageal or stomach ulcer/s, gastrointestinal or digestive hemorrhage or bleeding or perforation</td>
<td></td>
</tr>
<tr>
<td>And adverse toxic-side effect/s, risk, chemically induced, contraindications</td>
<td></td>
</tr>
<tr>
<td>And human only</td>
<td></td>
</tr>
<tr>
<td>Not neoplasms, cancer, malignancy, carcinoma, adenocarcinoma/s, case reports, case studies, case series, veterinary, animal experiment/s, letters</td>
<td></td>
</tr>
</tbody>
</table>
2.6. Assessment of pharmaceutical sponsorship of studies

Because bias in research sponsored by pharmaceutical companies has been reported [8–10], we assessed whether the included studies were sponsored by a pharmaceutical company. By definition, all studies submitted to the FDA were sponsored by a pharmaceutical company. To determine if the published studies were sponsored by a pharmaceutical company we abstracted (i) whether the manuscript explicitly stated that the study was sponsored by a pharmaceutical company, (ii) whether one or more of the authors were employed by a pharmaceutical company, (iii) whether employees of a pharmaceutical company were acknowledged, and (iv) whether the drugs used in the study were provided by a pharmaceutical company.

2.7. Data abstraction

All published titles, abstracts, and articles were reviewed independently by two trained physician-reviewers. Data elements, including items pertaining to study methodologic attributes, study population, dosing, outcome measures, and study statistics were abstracted by these same reviewers. All disagreements were resolved by consensus. The reviewers were not blinded to study outcomes.

The FDA study data were extracted using a two-step process. In the first step, each FDA study was reviewed by one of two trained nurse-reviewers, who abstracted data elements identical to those abstracted for the published data. Both nurses reviewed a 10% sample of studies to determine reliability of the abstraction. Any variable for which there was not 95% absolute agreement or for which Cronbach’s kappa [11] was not $\geq 0.8$ between nurse reviewers was re-abstracted for all studies by a pair of physician-reviewers. In addition, the physician pair abstracted all Jadad scoring criteria and concealment of allocation.

2.8. Definition of dyspepsia

The NSAID literature search identified a large number of terms used to report upper abdominal symptoms. Whereas most terms are reported across studies, the terms themselves are overlapping. We reviewed the gastroenterology literature for definitions of dyspepsia and found that all definitions included epigastric pain/discomfort as part of the symptom complex, but they differed as to whether they included various other symptoms, such as nausea, vomiting, and heartburn [12–14]. Thus, we developed a definition of dyspepsia for the meta-analysis that included any outcome terms (including the term “dyspepsia”) relating to epigastric or upper abdominal pain/discomfort but did not include nausea, vomiting, or heartburn. Because nearly all studies reported adverse outcomes by symptom and not by patient and because some studies included as outcomes more than one symptom meeting our definition of dyspepsia (ie, “stomach ache” and “epigastric distress”), we analyzed our data in three ways that correspond to conservative and liberal estimates of the risk of dyspepsia: (i) assuming that patients could report more than one symptom, such that the number of persons with dyspepsia was equal to the sum of all the reported outcomes, and for those studies that had disagreement between the single most frequent symptom occurring in the NSAID-treated group and the control group, we chose the outcome that minimized the risk ratio (conservative estimate); (ii) in the same circumstance, we chose the outcome that maximized the risk ratio (liberal estimate); and (iii) assuming all the symptoms represented unique persons, such that the number of persons with dyspepsia was equal to the sum of all the reported symptoms meeting our definition of dyspepsia (liberal estimate). Because the number of studies that were affected by disagreements between outcomes was low ($n = 5$) and because treating each symptom as a unique patient increased the percentage reporting dyspepsia but not the risk ratio, our results differed little among these methods. We present here only the conservative analyses.

2.9. Individual study statistics

For each study, we determined the percentage of subjects in each study arm who had the outcome, and we estimated the risk difference and risk ratio to assess the treatment effect. Studies that did not report symptoms consistent with our definition of dyspepsia or that did not state in the Methods section that these symptoms would be measured were not included in the dyspepsia analysis.

Most crossover studies did not report results separately for the study phase before the crossover. We considered three different methods of incorporating these data into our analysis, and we conducted a sensitivity analysis of the manner in which the crossover data were incorporated into our analysis, and the conclusions did not change regardless of the method chosen. Therefore, for studies that did not report results separately by study phase, we averaged the results before and after the crossover.

If zero outcomes were reported for a study’s control group, to ensure that all study statistics were defined, we ap-

---

**Table 2**

<table>
<thead>
<tr>
<th>NSAID</th>
<th>Total daily dose (mg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diclofenac</td>
<td>200</td>
</tr>
<tr>
<td>Etodolac</td>
<td>1000</td>
</tr>
<tr>
<td>Ibuprofen</td>
<td>1200</td>
</tr>
<tr>
<td>Nabumetone</td>
<td>3200</td>
</tr>
<tr>
<td>Naproxen</td>
<td>1500</td>
</tr>
</tbody>
</table>
plied the standard contingency table adjustment by adding a 0.5 to the treatment and control group numbers of outcomes and a 1 to the numbers of patients in the treatment and control groups.

2.10. Analysis of study characteristics

We conducted chi-squared tests to assess differences between the categorical characteristics of published and FDA studies. To assess the differences between the two groups of studies in mean quality scores and sample sizes, we conducted $t$ tests of the differences between means.

2.11. Pooled estimates

We pooled the treatment and control group percentages, risk differences, and risk ratios using a random effects model, which accounts for within and between study variance, using standard formulas for the estimates and confidence intervals [15,16]. Along with the pooled estimates, we report the result of the chi-squared test of heterogeneity, which tests the null hypothesis that all studies are estimating the same effect.

2.12. Meta-regression analysis

We also conducted a meta-regression analysis [17] to determine whether the effect of NSAIDs varied by whether the study was published or was an FDA study after adjusting for study differences. The following covariate-by-treatment interactions were tested: age, patient type, whether the study was published or an FDA study, whether the study excluded patients with previous gastrointestinal complaints, study quality, dose, and duration.

3. Results

3.1. Studies identified

We identified 4881 published titles that included the name of an NSAID or pertained to NSAIDs as a group (Figure 1). Among these, 2704 were rejected because the title clearly did not describe studies of NSAIDs among adults. An additional 377 articles were rejected after reviewing the abstract because they did not describe studies of NSAIDs among adults. Among the remaining 1800 articles, 1730 (96%) full text articles were obtained and screened. Among the articles screened, 1397 were rejected because they met one or more of the following exclusion criteria: (1) not a study (eg, review, etc.), (2) did not report outcome of interest (ie, gastrointestinal toxicity), (3) study population not human adults, and (4) duration of NSAID exposure less than 5 days. Among the remaining 333 articles, 55 were placebo-controlled, randomized clinical trials of NSAIDs. Among these, 20 studied the use of ibuprofen, naproxen, etodolac, diclofenac, or nabumetone; 15 reported dyspepsia as an outcome [18–32].

We identified 33 FDA reviews of NDAs or supplements to NDAs for ibuprofen, naproxen, etodolac, diclofenac, or nabumetone (Fig. 1). Among these, it was clear from the NDA listing that 19 were for pediatric, ophthalmic, or combination drug preparations. We obtained the reviews of the remaining 14 NDAs. Among these, two contained no information on studies of efficacy or toxicity but rather described only bioequivalence reports or legal correspondence. The remaining 12 reviews summarized 141 studies of efficacy or toxicity. Among these studies, 115 were RCTs, 25 were nonrandomized controlled trials, and one was an uncontrolled trial. A placebo arm was used in 72 of the RCTs and none of the nonrandomized clinical trials. Thirty-seven of the RCTs met our inclusion criteria. Eleven of these studies reported dyspepsia as an outcome.

Among the studies that met our inclusion criteria, one was described in both an FDA review and the published literature [23]. The results and conclusions were the same in both reports. However, the published report specified the number of patients in treatment and control groups with dyspeptic symptoms, whereas the FDA review did not. Data from the published report were included in the meta-analysis. The investigators listed for this study in the FDA review were completely different from the authors listed on the published report.

3.2. Population and dosing

Table 3 describes characteristics of the populations and drug indications used in the published and FDA data. Other than age, there were no statistically significant differences between published and FDA studies. Relative to published studies, FDA studies had a greater proportion that assessed women, patients with arthritis, and low or variable doses of NSAIDs. Because single studies contained multiple treatment arms with more than one dose level, the dosage categories are not mutually exclusive and thus could not be compared statistically.

3.3 Methodologic attributes

All published and all FDA studies were randomized. The majority of studies reported in either the published literature or the FDA reviews were blinded and described withdrawals and dropouts (Table 4). Although descriptions of methods of randomization or concealment of allocation were found for few of the published or FDA studies, they were found more often for the published studies ($P = 0.031$ and $P = 0.001$, respectively). The percentage of studies that described an appropriate method of blinding also differed significantly between the groups, with descriptions in 60% of published studies and 3% of FDA studies ($P < 0.001$). The average Jadad scores for the published and FDA studies were 3.7 and 2.9, respectively ($P = 0.006$). Jadad scores of 3 or higher occurred among 90% of the published and 81% of the FDA studies ($P = 0.271$).

3.4. Pharmaceutical sponsorship of studies

By definition, all studies submitted to the FDA were sponsored by a pharmaceutical company. Among the 15 published studies, 8 reported that the study was sponsored in total or in part by a pharmaceutical company [21,23–
Five of these listed one or more employees from the sponsor company as authors [23,25,28,31,32]. An additional three studies listed employees of pharmaceutical companies as authors but did not specify the source of funding for the study [19,20,29]. Two studies acknowledged assistance from employees of pharmaceutical companies and did not specify the source of funding [19,22], although, in one of these studies, the trial drugs were provided by a pharmaceutical company [18]. The other two studies did not specify the source of funding and did not list as authors or acknowledge employees of pharmaceutical companies.

3.5. Association of NSAIDs with dyspepsia

The pooled risk ratio for dyspepsia among NSAIDs relative to placebo users obtained when data from the published literature and the FDA studies were combined was 1.14 (95% confidence interval [CI] 0.85–1.53) (Table 5). The stratified analysis produced a pooled risk ratio for dyspepsia among NSAIDs relative to placebo users of 1.21 (95% CI 0.81–1.81) for the published studies and 1.07 (95% CI 0.70–1.63) for FDA studies (Table 5).

In the meta-regression analysis, treatment effect varied significantly only by NSAID dose ($P = 0.037$), with high dose being associated with more than twice the rate of dyspepsia as low or medium doses. Treatment effect did not vary significantly by whether the study was published or an FDA study ($P = 0.728$).

4. Discussion

Our results provide some measure of reassurance to people who are concerned about the methodologic quality of
manufacturer-supported studies submitted to the FDA as part of the NDA process. We found no meaningful difference between published studies (most of which were manufacturer sponsored) and manufacturer-sponsored unpublished studies on the most widely accepted criteria for methodologic quality of controlled trials, the Jadad scale. Although more published than FDA studies reported concealment of allocation, the low proportion of studies from either data source fulfilling this criterion makes its use as a measure of methodologic quality less meaningful. We did find a large difference between published and FDA studies in the reporting of appropriate methods of randomization and blinding. Whether this is due to poorer methodologic quality of manufacturer-supported studies submitted to the FDA, omission from the FDA reviews of methods that were described in the materials submitted to the FDA, or greater attention to reporting of these criteria by authors and editors in peer-reviewed journals is unknown.

Our results also provide some measure of reassurance to people who are concerned that efficacy and toxicity estimates reported in existing systematic reviews and meta-analyses have been systematically biased because they did not include unpublished data residing in the FDA reviews. We found no statistical or practical differences between published and FDA studies in terms of pooled estimates of the risk of dyspepsia. Consequently, the risk ratio for dyspepsia calculated from published and FDA data does not differ significantly from the risk ratio calculated using published data only.

Among the studies reported in the FDA reviews of NDAs, only one was identified in the published literature. The reasons for this are unknown, but it reinforces the widely held opinion that a large number of clinical trials go unpublished. We can conjecture that the FDA studies went unpublished because the manufacturer did not support publication (for proprietary reasons or otherwise), or the principal investigators decided not to submit the reviews for publication (due to loss of interest or lack of time), or the results were submitted but failed to be accepted (although we could not detect many meaningful differences in study characteristics or quality between published and FDA studies). Further work is needed to assess the reasons why studies submitted to the FDA are not published.

What can this study tell us about the role unpublished data from the FDA should play in systematic reviews and meta-analyses? For the selected NSAIDs we assessed, we found the FDA reviews contained a large number of previously unpublished studies on efficacy and risks that were not meaningfully different from published studies in terms of methodologic quality or reporting of dyspepsia. Until additional empirical data are available to support different conclusions, we suggest that inclusion of FDA data in systematic reviews and meta-analyses should be considered in the following situations: (i) when there is a paucity of published data and (ii) when there is an a priori reason to suspect that FDA data may be systematically different than published data.

In both situations, any FDA data that are used must be subjected to the same methodologic scrutiny as published data and analyses to assess whether and to what degree inclusion of FDA data affect the results. In cases where differences in results between published and FDA data are apparent, thoughtful analyses of the differences in study characteristics between the two data sources could potentially produce use-

---

Table 3
Characteristics of studies reported in published literature and FDA reviews

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Published studies (n = 20)</th>
<th>FDA studies (n = 37)</th>
<th>Significance of comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean sample size (SD)</td>
<td>114 (89)</td>
<td>150 (107)</td>
<td><em>P</em> = 0.41</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td><em>P</em> = 0.012</td>
</tr>
<tr>
<td>18–64 y</td>
<td>6 (30%)</td>
<td>4 (11%)</td>
<td></td>
</tr>
<tr>
<td>18 to &gt;65 / ≥ 65 y</td>
<td>14 (70%)</td>
<td>22 (59%)</td>
<td></td>
</tr>
<tr>
<td>Not described</td>
<td>0</td>
<td>11 (30%)</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td><em>P</em> = 0.105</td>
</tr>
<tr>
<td>Male only</td>
<td>1 (5%)</td>
<td>1 (3%)</td>
<td></td>
</tr>
<tr>
<td>Female only</td>
<td>0</td>
<td>8 (22%)</td>
<td></td>
</tr>
<tr>
<td>Mixed</td>
<td>18 (90%)</td>
<td>24 (65%)</td>
<td></td>
</tr>
<tr>
<td>Not described</td>
<td>1 (5%)</td>
<td>4 (11%)</td>
<td></td>
</tr>
<tr>
<td>Drug indication</td>
<td></td>
<td></td>
<td><em>P</em> = 0.084</td>
</tr>
<tr>
<td>Normals</td>
<td>4 (20%)</td>
<td>1 (3%)</td>
<td></td>
</tr>
<tr>
<td>Arthritis</td>
<td>11 (55%)</td>
<td>23 (62%)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>5 (25%)</td>
<td>13 (35%)</td>
<td></td>
</tr>
<tr>
<td>Dose used <em>a</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>5</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td>8</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>5</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>3</td>
<td>17</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: SD, standard deviation.

_a_ Chi-squared test used for categorical variables; _t_ tests of the differences between means used for continuous variables.

_b_ Categories are not mutually exclusive because a single study might contain more than one dose level in different treatment arms.

Table 4
Methodologic attributes of studies reported in published literature and FDA reviews

<table>
<thead>
<tr>
<th>Jadad score component</th>
<th>Number (%) of published studies (n = 20)</th>
<th>Number (%) of FDA studies (n = 37)</th>
<th>Significance of comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomized <em>b</em></td>
<td>20 (100%)</td>
<td>37 (100%)</td>
<td></td>
</tr>
<tr>
<td>Appropriate randomization described <em>b</em></td>
<td>6 (30%)</td>
<td>3 (8%)</td>
<td><em>P</em> = 0.031</td>
</tr>
<tr>
<td>Blinded <em>b</em></td>
<td>18 (90%)</td>
<td>36 (97%)</td>
<td><em>P</em> = 0.239</td>
</tr>
<tr>
<td>Appropriate blinding described <em>b</em></td>
<td>12 (60%)</td>
<td>1 (3%)</td>
<td><em>P</em> &lt; 0.001</td>
</tr>
<tr>
<td>Withdrawals and dropouts described <em>b</em></td>
<td>17 (85%)</td>
<td>29 (78%)</td>
<td><em>P</em> = 0.545</td>
</tr>
<tr>
<td>Studies with Jadad score ≥ 3</td>
<td>18 (90%)</td>
<td>29 (81%)</td>
<td><em>P</em> = 0.271</td>
</tr>
<tr>
<td>Concealment of allocation reported</td>
<td>5 (25%)</td>
<td>0</td>
<td><em>P</em> = 0.001</td>
</tr>
</tbody>
</table>

_Ch_ Chi-squared test used.

_b_ Jadad score component
can be considered a viable source for systematic reviews and meta-analyses but only after being subjected to the same methodologic scrutiny as published data.

4.1. Limitations

This study has several limitations. First, this study is limited to one drug class, and, hence, its findings may not be generalizable to other drug classes. However, the class assessed was large and permitted evaluation of a large number of studies. Second, the NDA material obtained through the Freedom of Information Act is a summary of the original data reported by the manufacturer and is therefore subject to all the errors of omission and commission inherent in summarized reports. However, published studies themselves are also summaries of original studies and are themselves subject to possible differences between what was reported and what was actually done. For published and unpublished studies, meta-analysts in most cases have available only the reports of original data, not the original data itself. Third, we used a method to assess the methodologic quality of the published and unpublished data that has been previously validated only for published studies. Fourth, the findings we report here are based on an assessment of a common adverse event and as such cannot be generalized to rare events or efficacy estimates. Last, because the published studies we report were, for the most part, sponsored by the pharmaceutical industry (as were all of the unpublished studies), comparisons between the methodologic quality of published, nonindustry-sponsored research and unpublished industry-sponsored research could not be made. However, given that the methodologic quality of the unpublished industry-sponsored research reported here was generally high, it is unlikely these studies would be of lower methodologic quality than published, nonindustry-sponsored research.

5. Conclusions

Data from FDA reviews should not be excluded from systematic reviews and meta-analyses based solely on concerns about their methodologic integrity. Rather, these data can be considered a viable source for systematic reviews and meta-analyses but only after being subjected to the same methodologic scrutiny as published data.

References

[18] Berry H, Bloom B, Hamilton EBD. A comparative study of piroxi-


