

## **Industry-corrupted psychiatric trials<sup>1</sup>**

Jay D. Amsterdam<sup>1</sup>, Leemon B. McHenry<sup>2</sup>, Jon N. Jureidini<sup>3</sup>

<sup>1</sup>Depression Research Unit, Department of Psychiatry,  
University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania

<sup>2</sup>Department of Philosophy, California State University, Northridge, California

<sup>3</sup>Critical and Ethical Mental Health Research Group, Robinson Research Institute,  
University of Adelaide, Australia

### **Summary**

The goal of this paper is to expose the research misconduct of pharmaceutical industry-sponsored clinical trials via three short case studies of corrupted psychiatric trials that were conducted in the United States. We discuss the common elements that enable the misrepresentation of clinical trial results including ghostwriting for medical journals, the role of key opinion leaders as co-conspirators with the pharmaceutical industry and the complicity of top medical journals in failing to uphold standards of science and peer review. We conclude that the corruption of industry-sponsored clinical trials is one of the major obstacles facing evidence-based medicine.

**Key words:** antidepressants, citalopram, clinical trials, depression, ghostwriting, key opinion leaders, paroxetine, research misconduct, selective reporting

### **Introduction**

As former editor of the *New England Journal of Medicine*, Marcia Angell reported in 2008:

*Over the past 2 decades, the pharmaceutical industry has gained unprecedented control over the evaluation of its own products. Drug companies now finance most clinical research on prescription drugs,*

---

<sup>1</sup> Editor's note: the first version of this article was published as commentary to Barry Blackwell: Corporate Corruption in the Psychopharmaceutical Industry (<http://inhn.org/controversies/barry-blackwell-corporate-corruption-in-the-psychopharmaceutical-industry/jay-d-amsterdam-leemon-b-mchenry-and-jon-n-jureidinis-commentary-industry-corrupted-psychiatric-trials.html>)

---

*and there is mounting evidence that they often skew the research they sponsor to make their drugs look better and safer [1].*

In this review article we provide an overview of industry-sponsored clinical trials in psychiatry as part of the “mounting evidence” confirming Angell’s conclusion that drug evaluation is a “broken system.” We focus attention on three compelling cases that have demonstrated the extent to which the pharmaceutical industry will corrupt science for profit, SmithKline Beecham’s studies 329 and 352 and Forest Laboratories study CIT-MD-18, and reveal the crucial elements that enabled this state of affairs – industry manipulation of scientific data, ghostwriting clinical trial reports, academic physicians serving the marketing objectives of the sponsor companies and the failure of checks and balances in the peer review system and in regulatory bodies.

### **Medical ghostwriting and data misrepresentation**

It is now well known that pharmaceutical companies launder their promotional efforts through medical communication companies that ghostwrite articles and then pay academic consultants to sign on to the fraudulent articles [2, 3]. These firms are also engaged to “neutralize” academic physicians who have been identified as disloyal to their corporate client’s ineffective or unsafe drugs [4, 5]. What is less clear, however, is the fine detail of the business that is just emerging through litigation.

Pharmaceutical companies seeking to “launch” a new drug on the market or a new indication for a drug already approved for another indication hire a public relations firm and a medical communication company as part of their marketing strategy and publication planning. The firms set up advisory board meetings with key opinion leaders and marketing executives in advance of the clinical trials. Once a trial is complete, the medical ghostwriter who is employed by the medical communications firm produces a draft of a manuscript – from a summary of the Final Study Report of the clinical trial – and seeks feedback from the corporate sponsor. It is at this stage in the manuscript production that misrepresentation of the trial data frequently occurs, since the medical ghostwriter is under the direction of marketing executives to “spin” the data. The medical ghostwriter then revises a number of drafts with input from the external academic “authors” and internal industry scientists, and once the corporate sponsor is satisfied that the final manuscript draft is “on message,” it is submitted by a corporate-designated lead author to a medical journal for peer review. Once the manuscript is submitted, the medical ghostwriter disappears or is acknowledged in the fine print for “editorial assistance” [5].

When a pharmaceutical company engages a ghostwriter from a medical communication firm, a contract between the sponsor and the medical communication firm specifies that the manuscript is the intellectual property of the sponsor. The sponsor also owns the data from the clinical trials that are reported in the paper. When the

manuscript is submitted for publication, copyright is transferred by the sponsor's legal department to the named lead author of the paper, but there is no requirement of this disclosure [6]. In many cases, such as all three of the reports discussed below, the decision about who will be named as an "author" on the publication occurs well after the manuscript has been written by the ghostwriter. Much of the misreporting of the data in ghostwritten manuscripts is a result of the ownership of the data and the manuscript since the sponsor will determine how the data is released to the ghostwriter and to the external, academic "authors" who never review the raw data and are therefore in no position to determine whether the data are accurately reported. None of this critical information is ever made available to peer reviewers engaged by the scientific journals. Since reviewers are completely unaware of the *a priori* study objectives, who vetted the study results, and who actually wrote the article, the peer-review process is *ipso facto* flawed and incomplete.

Ghostwriting in the hands of the pharmaceutical industry has become a major factor in the "crisis of credibility" in academic medicine. Academic authorship is an assertion of intellectual responsibility in that the named authors of scientific reports are collectively responsible for study design, conduct, data analysis and writing. The integrity of science depends on the trust placed in individual clinicians and researchers and in the peer-review system which is the foundation of a reliable body of knowledge. When, however, academic physicians allow their names to appear on ghostwritten articles, they betray this basic ethical responsibility and are guilty of academic misconduct. An annual Harvard University Master Class in Psychopharmacology demonstrates the point. Several presenters – advertised as "world renowned faculty," – have been some of the worst offenders in medical ghostwriting scandals. One academic psychiatrist claims to be "author of over 1,000 scientific articles and book chapters, and the co-editor of *Textbook of Psychopharmacology*" [7]. Among these 1,000 scientific articles is the paroxetine 352 trial discussed below.

Ghostwriting is not limited to drafting a manuscript; rather it is an academic façade for research that has been designed, conducted and analyzed by industry and it is the main vehicle through which the misrepresentation of the data in favor of the study drug is achieved.

The vast majority of ghostwritten publications in medicine will never come to light as ghostwritten. To date, the only cases in which ghostwriting has been exposed to the public are those in which there were damages that resulted in litigation, government inquiries or from physicians who were approached to participate in ghostwriting projects and became whistle blowers [8]. As far as litigation is concerned, only a few select cases will surface since the majority disappear in legal settlement agreements. Incriminating documents remain proprietary information if plaintiffs' attorneys do not seek to remove the confidentiality designation of documents in protective orders.

### Key opinion leaders as co-conspirators

The term “key opinion leader” (KOL) or “thought leader” is a pharmaceutical industry creation for physicians who influence their peers’ medical practice and prescribing behavior [9]. Pharmaceutical companies claim to engage KOLs in the drug development process to gain expert evaluation and feedback on marketing strategy, but in reality these academic physicians are carefully vetted by the industry on the basis of their malleability to the sponsor’s products. KOLs are highly paid “product champions” who are engaged to “defend the molecule” [4]. Few physicians and psychiatrists can resist the flattering offer by industry to become KOLs, but the primary ethical duty to patients is compromised by what David Healy calls “ornamental additions to business” [10].

The industry-academic partnerships that have created the KOL phenomenon are often traced to one of the most influential pieces of legislation in the United States to impact the field of intellectual property law – the Bayh-Dole Act of 1980 [11]. The Bayh-Dole Act created a uniform patent policy that allowed universities to retain ownership to inventions made under federally funded research. The motivation was to speed up the commercialization process of federally-funded research, create new industries and open new markets from the university-patented inventions. The growth of university patents and the commercialization of research that followed Bayh-Dole at first seemed to have nothing but positive effects, but it soon became clear that the legislation had negative results. Universities that were losing government funding found the new source of revenue in the technology transfer to industry, but at the price of a proliferation of conflicts of interest. The most disturbing aspect of these arrangements, however, was the increased motivation to manipulate research results in favor of industry’s products.

In the description below of corrupted psychiatric trials, KOLs engaged by SmithKline Beecham and Forest Laboratories became the named “authors” on the ghostwritten publications that appeared in *The American Journal of Psychiatry* and *The Journal of the American Academy of Child and Adolescent Psychiatry*. Several of these academic KOLs were also on the companies’ Speakers’ Bureaus, Advisory Boards and provided company-sponsored continuing medical education lectures, all designed to promote selective serotonin reuptake inhibitor (SSRI) antidepressants. What passed as “medical education” was carefully disguised drug promotion created by medical communication companies and public relations firms.

### Complicit medical journals

Medical journals are part of the problem rather than the solution to the problem. Instead of demanding rigorous peer review of a submissions and an independent analysis of the data, medical journal editors are pressured to publish favorable articles

of industry-sponsored trials and rarely publish critical deconstructions of ghostwritten clinical trials [12, 13]. As medical journals and their owners have become dependent upon pharmaceutical revenue, the journals fail to adhere to the standards of science. Thus the publication of “positive” studies showing drug safety and effectiveness means more pharmaceutical advertising and more orders of reprints for dissemination by the sales force. In contrast, a “negative” study showing poor tolerability or ineffectiveness results in no such revenue.

### Three case studies

Serious problems with industry-sponsored clinical trials have been clearly identified in the process of peer review of the submitted manuscript yet these manuscripts are published against the reviewers’ negative recommendations [6]. Submissions of deconstructed industry-sponsored clinical trials pass peer review and are rejected by journal editors who override peer review or by attorneys representing the journals’ owners. Moreover, the pharmaceutical and medical device industries manipulate journal editors with threats of libel actions [13]. Finally, when journal editors and their owners such as the American Academy of Child and Adolescent Psychiatry and the American Psychiatric Association are confronted with indisputable evidence of industry fraud published in their journals, they refuse retraction [14]. When the probability of having a ghostwritten, fraudulent, industry-sponsored clinical trial accepted for publication in a high-impact medical journal is substantially higher than the probability of having a critical, deconstruction of the same trial accepted there can be no confidence in the medical literature. In this regard, many medical journals, contrary to common opinion, are not reliable sources of medical knowledge. They are guilty of publishing pseudo-science and have become, in the words of former *British Medical Journal* editor, Richard Smith, “an extension of the marketing arm of pharmaceutical companies” [15].

Few ghostwritten articles of clinical trials in psychiatry have been deconstructed in order to reveal publicly their *sub rosa* research misconduct and misrepresentation of outcome data. Two industry-sponsored, ghostwritten psychiatric articles briefly discussed below have been deconstructed from court documents and have received recent media attention. These two studies have much in common and are both the result of pharmaceutical companies manipulating outcome data in order to facilitate the off-label marketing of SSRI antidepressant medication to children and adolescents. The third deconstruction article has received less attention. It examines an industry-sponsored, ghostwritten article that came to light as part of an academic whistleblower complaint of plagiarism and research misconduct against prominent academic KOLs at medical research universities in the United States and several pharmaceutical company executives. It involved the manipulation of sample size estimates and the misrepresentation of outcome data in adults with bipolar affective disorder.

All three of these deconstructed psychiatric trials were published in a medical journal that does not depend on pharmaceutical industry revenue, the *International Journal of Risk & Safety in Medicine* [16–18]. Previously-confidential industry documents upon which the first two deconstructions are based are posted on the websites Healthy Skepticism and the Drug Industry Document Archive (DIDA).

### SmithKline Beecham Paroxetine Study 329

Study 329 compared the efficacy and safety of paroxetine and imipramine with placebo in the treatment of adolescents with unipolar major depression. The 1993 protocol for the study (and its subsequent amendments) specified two primary outcome measures: change in total Hamilton Rating Scale for Depression (HRSD) score; and proportion of remitters and responders with a change in HRSD score  $\leq 8$  or reduced by  $\geq 50\%$ . The protocol also specified six secondary outcome measures. A total of 275 subjects were enrolled.

The published article was ghostwritten by Sally K. Laden of Scientific Therapeutics Information, Inc., under the direct sponsorship of SmithKline Beecham (SKB) employees and was published by the *Journal of the American Academy of Child and Adolescent Psychiatry (JAACAP)* in July 2001 under the so-called authorship of Keller *et al.* [19, 6]. As *post hoc* deconstruction indicated, the “positive” published results were a gross misrepresentation of the actual negative study results [16]. Deconstruction was made possible by examining approximately 10,000 court documents from a class action lawsuit, *Beverly Smith vs. SmithKline Beecham*. Keller *et al.* claimed that paroxetine was “generally well tolerated and effective for major depression in adolescents” [19], while SKB (SmithKline Beecham) claimed that paroxetine demonstrated “remarkable efficacy and safety” [20]. The *JAACAP* article eventually became one of the most frequently-cited studies in the medical literature in support of SSRI antidepressant use in child and adolescent depression [21]. However, unknown to the *JAACAP* readers, the SKB 329 study was completely negative on all protocol-designated primary outcomes, most secondary protocol-designated outcomes, and that SKB withheld clinically important adverse event information on paroxetine-induced suicidal and manic-like behaviors in children and adolescents.

Initial data analysis showed that there was no significant difference between the paroxetine and placebo groups on any of the eight protocol-specified outcome measures [22]. Undaunted by these disappointing outcomes, SKB and the investigators performed additional, non-protocol designated *post hoc* analyses showing more favorable results for paroxetine. Even then, only two of these *post hoc* comparisons were statistically significant for paroxetine versus placebo by the time study 329 was first ghostwritten for publication in the *Journal of the American Medical Association (JAMA)* [23]. Four of the six negative protocol-specified secondary outcome measures had been removed from the list of secondary outcomes, and the two additional *post hoc* “positive” out-

comes had been added. Thus, overall, 4 of the eight “negative” protocol-designated outcomes were replaced with 4 “positive” outcomes (although many other “negative” measures had been tested and rejected along the way) [16].

The ghostwriter conflated the primary and secondary outcomes as early as the first draft of the manuscript, and all 8 outcomes were described as “primary” in the results section. However, in later ghostwritten drafts, the term “primary” was replaced by the term “depression-related” outcomes [24], whereby later drafts reported that paroxetine was more effective than placebo on 4 of eight outcomes, without disclosing that the original protocol-designated primary and secondary outcomes were really “negative” [6].

In July 2000, after being rejected by peer review from *JAMA* [23], the revised manuscript was submitted to *JAACAP*, where one peer reviewer asked that the primary outcomes be specifically reported [25]. Despite this request, the two original protocol-designated primary outcomes were still not declared, and “authors” continued to claim efficacy for paroxetine based on the conflated outcomes. This conflation extended throughout the remainder of the *JAACAP* peer-review process, obscuring the original “negative” primary outcome results by reporting “positive” outcome results.

Finally, the *JAACAP* article stated that: “Paroxetine was generally well tolerated in this adolescent population, and most adverse effects were not serious” [19]. In contrast, the final unpublished SKB’s study report in November 1998 indicated the presence of many serious and severe adverse events with paroxetine. Specifically, suicidal thoughts and behavior were miscoded under the euphemism “emotional lability” [19, 22]. Moreover, the final study report revealed three additional cases of suicidal ideas or self-harm that were unreported. Thus, at least eight paroxetine subjects reported self-harmed or worsening suicidal ideation compared to only one placebo subject [22].

An internal SKB “Position Piece on the Phase III clinical studies” from 1998 stated that study 329: “failed to demonstrate a statistically significant difference from placebo on the primary efficacy measures,” and set as a target: “To effectively manage the dissemination of these data in order to minimize any potential negative commercial impact.” A cover letter reads: “As you know, the results of the studies were disappointing in that we did not reach statistical significance on the primary endpoints and thus the data do not support a label claim for the treatment of Adolescent Depression” [26]. These documents were disavowed by SKB, but there was certainly more than one person expressing caution at this time. One of those persons who can be cited stated: “Originally we had planned to do extensive media relations surrounding this study until we actually viewed the results. Essentially the study did not really show Paxil was effective in treating adolescent depression, which is not something we want to publicize” [27].

In summary, the results of the paroxetine 329 study were negative for efficacy and positive for harm.

---

### Forest Laboratory Citalopram Study CIT-MD-18

The CIT-MD-18 study was conducted between 1999 and 2002. It was a 9-week, 20-site, randomized, double-blind comparison of the safety and efficacy of citalopram versus placebo in 160 children (age 7–11) and adolescents (age 12–17) with major depressive disorder. CIT-MD-18 was designated a Phase III registration trial supporting an FDA indication for depression in pediatric patients. Forest also parsed out the CIT-MD-18 adolescent results to support an FDA adolescent major depressive disorder indication for escitalopram. The primary efficacy measure was the change from baseline to week 8 on the Children’s Depression Rating Scale-Revised (CDRS-R) total score. Secondary efficacy measures were the Clinical Global Impression severity and improvement subscales, Kiddie Schedule for Affective Disorders and Schizophrenia – depression module, and Children’s Global Assessment Scale [28].

According to court documents made public as part of the *Celexa and Lexapro Marketing and Sales Practices Litigation*, the article was ghostwritten by Natasha Mitchner at Weber Shandwick Communications under instruction from Jeffrey Lawrence (Product Manager Forest Marketing). In an October 15, 2001 email, Mary Prescott of Weber Shandwick makes it explicit that the manuscript was written prior to the selection of Karen Wagner as the lead author and the other so-called academic “authors” [29].

Dr. Wagner’s input was sought only after the first draft of the CIT-MD-18 manuscript was prepared and reviewed by Forest employees. In an email dated December 17, 2001 Mr. Lawrence of Forest wrote to Ms. Mitchner: “Could you do me a favor and finish up the pediatric manuscript? I know you said you only had a bit more to do [...] I took a quick look at it and it looked good so I’d like to get it circulated around here before we send it off to Karen [Wagner]” [30].

Forest control over the manuscript production allowed for a positive spin to the study outcome. The published Wagner *et al.* article concluded that citalopram produced a significantly greater reduction in depressive symptoms than placebo in children and adolescents [31]. This conclusion was supported by claims that citalopram reduced the mean CDRS-R scores significantly more than placebo beginning at week 1 and at every week thereafter (effect size = 2.9); and that response rates at week 8 were significantly greater for citalopram (36%) versus placebo (24%). Wagner *et al.* also claimed comparable rates of tolerability and treatment discontinuation for adverse events (citalopram = 5.6%; placebo = 5.9% [31]).

However, deconstruction of study data and court documents revealed that the claims of Wagner *et al.* were predicated upon a combination of misleading analysis of the primary study outcome, an implausible effect size, introduction of *post hoc* outcomes as if they were primary outcomes, failure to report negative secondary outcomes, inclusion of eight unblinded subjects into efficacy analyses, and misleading analysis and reporting of adverse events. For example, contrary to protocol stipulation, Forest

increased the final sample size by adding back into the primary outcome analysis eight of nine subjects who should have been excluded from the data analysis because they were dispensed unblinded study drug [32].

Forest had performed a primary outcome calculation excluding these subjects [28]. This per protocol exclusion resulted in a “negative” primary efficacy outcome. Ultimately, however, eight of the excluded subjects were added back into the analysis, turning the marginally insignificant outcome ( $p < 0.052$ ) into a statistically significant outcome ( $p < 0.038$ ). The unblinding error was not reported in the published Wagner *et al.* article nor in any of the communications to the medical community including Forest-sponsored posters delivered at medical conferences, press releases or continuing medical education programs.

One revealing internal Forest communication about a letter to the FDA on the unblinding issue made it clear that there was some degree of duplicity in reporting this protocol violation. In a March 14–15, 2000 email exchange between Amy Rubin and Dr. Charles Flicker, Dr. Flicker stated: “Altho ‘potential to cause bias’ is a masterful stroke of euphemism, I would be a little more up front about the fact that the integrity of the blind was unmistakably violated.” Ms. Rubin responded: “Thanks for the compliment. Part of my job is to create ‘masterful’ euphemisms to protect Medical and Marketing” (sic) [33].

Forest also failed to follow their protocol stipulated plan for analysis of age-by-treatment interaction. The primary outcome variable was the change in total CDRS-R score at week 8 for the entire citalopram versus placebo group, using a 3-way ANCOVA test of efficacy [28]. Although a significant efficacy value favoring citalopram was produced after including the unblinded subjects in the ANCOVA, this analysis resulted in an age-by-treatment interaction with no significant efficacy demonstrated in children. This important efficacy information was withheld from public scrutiny and was not presented in the published article, nor did the published article report the power analysis used to determine the sample size. Thus, Forest could not make a claim for efficacy in children (and possibly not even in adolescents). However, if Forest powered the study to make a claim for efficacy in the combined child plus adolescent group, this may have been invalidated as a result of the ANCOVA age-by-treatment interaction and would have shown that citalopram was not effective in children. A further exaggeration of the effect of citalopram was to report an “effect size on the primary outcome measure” of 2.9, which was extraordinary and not consistent with the primary data. This claim was questioned by Martin *et al.* who criticized the article for miscalculating effect size or using an unconventional calculation, which clouded “communication among investigators and across measures” [34]. The origin of the effect size calculation remained unclear even after Wagner *et al.* publicly acknowledged an error and stated that “With Cohen’s method, the effect size was 0.32” [35], which is more typical of antidepressant trials.

Finally, the ghostwritten article failed to mention that five citalopram-treated subjects discontinuing treatment did so due to hypomania, agitation, and akathisia. None of these potentially dangerous states of over-arousal occurred with placebo. Furthermore, citalopram-induced anxiety occurred in one subject severe enough to warrant premature treatment discontinuation; while irritability occurred in three other citalopram (versus one placebo) subject. These adverse events raise concerns about dangers from the activating effects of citalopram that should have been reported in the Wagner *et al.* article. Instead Wagner *et al.* reported “adverse events associated with behavioral activation (such as insomnia or agitation) were not prevalent in this trial” and claimed that “there were no reports of mania” [31].

In summary, the CIT-MD-18 study was negative and therefore not supportive of Forest’s FDA adolescent indication application.

### SmithKline Beecham Paroxetine Study 352

Deconstruction of the SKB paroxetine 352 study was based, in part, upon documents pertaining to: (1) Complaint of Plagiarism of the Nemeroff *et al.* 2001 article published in the *American Journal of Psychiatry* [36], (2) Expert testimony from the *Kilker vs. SmithKline Beecham* trial, October 2009 [37], (3) *Senate Report on Ghostwriting in Medical Literature*, June 24, 2010 [38], (4) SKB Clinical Trials Website Result Summary for Study 29060/352 updated 09 March 2005 [39], (5) SKB Paroxetine Protocol PAR-29060/352 (amended 22 July, 1994), and (6) Complaint of Scientific Misconduct against Dwight L. Evans, Laszlo Gyulai, Charles B. Nemeroff, Gary S. Sachs, Charles L. Bowden *et al.*, July 8, 2011 filed with the Office of Research Integrity (ORI) of the Department of Health and Human Services: ORI 2012-33 [40].

The paroxetine 352 article was ghostwritten by Sally Laden of Scientific Therapeutics Information, Inc. under the sponsorship of SKB employees and was published by the *American Journal of Psychiatry* in June 2001 under the so-called authorship of Nemeroff *et al.* [41]. However, the role of SKB and the ghostwriters was not acknowledged in the article. At least two ghostwritten drafts of the 352 manuscript were produced before the names of any academic authors appeared on the title page. Eventually, prominent academic researchers (with financial ties to SKB) as well as SKB employees, were designated by SKB as “authors” on the 3<sup>rd</sup> draft of the 352 manuscript [17].

According to the Complaint of Research Misconduct on June 25, 2012, the so-called authors were chosen by SKB in consultation with Sally Laden. The so-called authors on the published article had little or no direct involvement in the design, conduct, data analysis, or writing of the manuscript. In fact, the first and second authors on the published article (i.e., Drs. Nemeroff and Evans) were selected for this role by Sally Laden late in the vetting process (after several other authors were moved to less prominent positions in the by-line). SKB had originally selected Dr. Laszlo Gyulai

from the University of Pennsylvania, as the paper's first author. However, Dr. Gyulai was replaced by Dr. Nemeroff. The evidence also indicates that the final SKB-assigned authors on the published article never reviewed or even saw preliminary drafts of the paper, and only saw the final edited manuscript just prior to final acceptance by the *AJP* [17].

The 352 study was designed as an 18-site, 10-week, randomized, double-blind, placebo-controlled comparison of paroxetine versus imipramine in subjects with bipolar type I disorder and was designated a Phase IV (i.e., post-marketing, non-indication) study with a projected duration of 2 years. Its objective was "to compare the efficacy and safety of paroxetine and imipramine to [placebo] in the treatment of bipolar depression in subjects stabilized on lithium therapy." The primary efficacy measures were the change from the baseline HRSD total score, and the change from baseline in the Clinical Global Impression Severity of Illness (CGI/S) score for paroxetine versus placebo and for imipramine versus placebo. The protocol-stipulated secondary outcomes were the proportion of subjects with a final HRSD score  $\leq 7$  or a final CGI/S score  $\leq 2$ . Additional secondary outcomes included the proportion of subjects experiencing adverse events, premature treatment discontinuation, and manic or hypomanic reactions as determined by the DSM-III-R Mania/Hypomania Assessment and the Young Mania Rating Scale (YMRS).

The study population consisted of outpatient subjects  $\geq 18$  years old, with a lifetime diagnoses of bipolar type I disorder and a history of at least one prior manic or major depressive episode within the preceding 5 years who failed to respond to lithium carbonate for  $\geq 7$  weeks at therapeutic lithium levels [41]. The original protocol called for a sample size of 62 subjects per treatment group (or a total of 186 subjects).

The statistical plan called for separate analyses on the entire subject population, and on two subgroups of subjects: (i) those who experienced a manic or hypomanic episode during the study; and (ii) those who did not. The YMRS was to be used to assess severity of manic and/or hypomanic symptoms across treatment conditions, and the relationship between change from baseline in YMRS scores and HRSD scores was to be specifically examined.

Factors that might influence treatment outcome were to be examined via the use of interaction terms in the regression models and those that were not statistically significant (i.e.,  $p > 0.1$ ) in the primary analysis would be dropped from all subsequent analyses. The protocol also stipulated that the comparison of primary interest was paroxetine versus placebo regardless of baseline lithium level stratification. Finally, mania and hypomania were to be analyzed using logistic regression models that included the effect terms of "treatment," "investigator," and "treatment x investigator" interaction. The protocol noted that if the interaction was not significant, it would be dropped from the model. Virtually none of these protocol-designated procedures were followed or reported in the published article [41].

The original sample size estimate of 62 subjects per treatment condition was reduced to 46 per group during the study; the result of low subject enrolment which led SKB to add a 19<sup>th</sup> investigative site. By the time that the study was prematurely terminated by SKB, only 117 (of the originally projected 186) subjects were enrolled, resulting in final sample sizes for paroxetine ( $n = 35$ ), imipramine ( $n = 39$ ), and placebo ( $n = 43$ ). However, by the time the study was published, the declared sample size estimate had mysteriously changed once again, and the authors fictitiously wrote: “the study was designed (sic) to enroll 35 patients per arm, which would allow 70% power to detect a 5-point difference on the Hamilton depression scale score ( $SD = 8.5$ ) between treatment groups” [41].

Although the statistical power was estimated at only 70%, Nemeroff *et al.* failed to inform the reader that this value was unconventionally low, derived *post hoc* (after the analyses had been completed), was not indicative of the original protocol-designated power estimate, and that the original power was based upon 62 subjects per group or that the original value was reduced during the study to 46 subjects per group. Moreover, Nemeroff *et al.* did not inform the reader that the power was further reduced to 35 subjects per group after the data were analyzed. No mention was made that this second *post hoc* power change occurred as an extra-regulatory protocol violation of HHS Good Clinical Practice Guidelines; or, that the second reduction in sample size was made *post hoc* in order to allow the final sample size estimate of 35 subjects per group to comport with the final sample size of the truncated paroxetine enrolment (i.e.,  $n = 35$ ). Nemeroff *et al.* did not acknowledge that the study failed to recruit the originally projected sample size necessary to test the primary study hypothesis, and only hinted by mentioning the low 70% power estimate that the study had insufficient statistical power to adequately test the primary study aims.

Nemeroff *et al.* played down the protocol-designated statistical procedures for the primary efficacy analyses and, instead, emphasized statistical procedures used for analyzing the unnecessary *post hoc* lithium stratification level efficacy analyses. Nemeroff *et al.* did not note that a sample size of 35 subjects per group was insufficient to test for differences among lithium level subgroups. Furthermore, “no adjustments for multiple comparisons were made” [41] which, if properly applied, would have nullified the only “positive” paroxetine finding in the study.

Nemeroff *et al.* failed to note that the YMRS was employed as an outcome measure in the study; and all manic and hypomanic safety ratings obtained with the YMRS were suppressed and omitted from the published article.

SKB conflated primary and *post hoc* analyses to present the only “positive” *post hoc* finding for paroxetine in the entire study as if it was the primary outcome (i.e., a stratified lithium level analysis). However, according to the protocol statistical plan, the *post hoc* lithium level stratification analyses were completely unnecessary. Of more than 30 separate primary, secondary and *post hoc* efficacy analyses reported in the SKB

Clinical Trials Web-site Results Summary, only the *post hoc* comparison of paroxetine versus placebo in subjects with low baseline lithium levels showed a statistically “positive” result for paroxetine. Nemeroff *et al.* attributed the “negative” primary outcome finding of paroxetine versus placebo in all subjects to an excessive placebo response in the “high” lithium level subgroup, although there is no evidence to support this conclusion. Nemeroff *et al.* then emphasized the single “positive” paroxetine efficacy finding as if it was the primary study aim [41].

Finally, Nemeroff *et al.* conflated efficacy and safety data to favor paroxetine by only presenting selected data on treatment-emergent manic and sexual side effect symptoms in subjects taking imipramine and minimized the rate of manic and hypomanic symptoms occurring with paroxetine.

In summary, the paroxetine 352 study was a non-informative trial with insufficient statistical power and inconclusive results. There was no evidence of any paroxetine efficacy in bipolar disorder, and the suppression of safety data from the YMRS outcome measure hid the presence of paroxetine-induced manic induction.

### Concluding remarks

Industry corruption of clinical trials is one of the major obstacles facing evidence-based medicine. As we have demonstrated above, the use of ghostwriters and key opinion leaders in the production of medical journal articles facilitated misrepresentation of the efficacy and safety data in psychiatric clinical trials that have been influential in clinical practice. Unfortunately, there is nothing exceptional about these three cases. All industry-sponsored trials are suspect and should be treated as such. Because so few gain public scrutiny and even fewer are ever formally retracted, it is important to make these articles transparent to correct the scientific record and protect patients from potential harm.

### References

1. Angell M. *Industry-sponsored clinical research: A broken system*. JAMA. 2008; 300(9): 1069–1071. Doi:10.1001/jama.300.9.1069.
2. Healy D, Cattell D. *Interface between authorship, industry and science in the domain of therapeutics*. Br. J. Psychiatry 2003; 183: 22–27.
3. Sismondo S. *Ghost management: How much of the medical literature is shaped behind the scenes by the pharmaceutical industry?* PloS Med. 2007; 4: e286.
4. McHenry L. *On the origin of great ideas: Science in the age of big pharma*. Hastings Cent. Rep. 2005; 35(6): 17–19.
5. McHenry L. *Of sophists and spin-doctors: Industry-sponsored ghostwriting and the crisis of academic medicine*. Mens Sana Monographs. 2010; 8: 129–145.

6. McHenry L, Jureidini J. *Industry-sponsored ghostwriting in clinical trial reporting: A case study*. Account Res. 2008; 15(3): 152–167.
7. Harvard University Master Class in Psychopharmacology. <http://www.hms-cme.net/734280/>.
8. Fugh-Berman A. *The corporate co-author*. J. Gen. Intern. Med. 2005; 20: 546–548.
9. Jureidini J. *Key opinion leaders in psychiatry: A conflicted pathway to career advancement*. Aust. N. Z. J. Psychiatry 2012; 46: 495–497.
10. Healy D. *Let them eat Prozac*. New York: New York University Press; 2004.
11. Krinsky S. *Science in the Private Interest*. Lanham: Roman & Littlefield; 2003.
12. Horton R. *The dawn of McScience*. NY Rev. Books 2004; (March 11): 7–9.
13. Healy D. *Our censored journals*. Mens Sana Monographs. 2008; 6: 244–256.
14. Jureidini J, McHenry L. *Conflicted medical journals and the failure of trust*. Account Res. 2011; 18: 45–54.
15. Smith R. *Medical journals are an extension of the marketing arm of pharmaceutical companies*. PLoS Med. 2005; 2(5): e138.
16. Jureidini J, McHenry L, Mansfield P. *Clinical trials and drug promotion: selective reporting of study 329*. The International Journal of Risk & Safety in Medicine 2008; 20(1–2): 73–81.
17. Amsterdam JD, McHenry L. *The paroxetine 352 bipolar trial: A study in medical ghostwriting*. The International Journal of Risk & Safety in Medicine 2012; 24(4): 221–231.
18. Jureidini J, Amsterdam J, McHenry L. *The citalopram CIT-MD-18 pediatric depression trial: A deconstruction of medical ghostwriting, data manipulation and academic malfeasance*. The International Journal of Risk & Safety in Medicine 2016; 28: 33–43.
19. Keller MB, Ryan ND, Strober M, Klein RG, Kutcher SP, Birmaher B et al. *Efficacy of paroxetine in the treatment of adolescent major depression: a randomized, controlled trial*. J. Am. Acad. Child Adolesc. Psychiatry 2001; 40(7): 762–772.
20. SmithKline Beecham. Hawkins to all sales representatives selling Paxil, Aug 16, 2001. <http://www.healthyskepticism.org/documents/documents/20010816Hawkinstoreps.pdf>.
21. Journal Citation Reports. <http://thomsonscientific.com/products/jcr/>.
22. SmithKline Beecham. A Multi-center, double-blind, placebo controlled study of paroxetine and imipramine in adolescents with unipolar major depression – Acute Phase. Final Clinical Report. SB Document Number: BRL-029060/RSD-100TW9/1/CPMS-329. November 24, 1998. [http://www.gsk.com/media/paroxetine/depression\\_329\\_full.pdf](http://www.gsk.com/media/paroxetine/depression_329_full.pdf).
23. SmithKline Beecham. Draft I, December 18, 1998. <http://www.healthyskepticism.org/documents/documents/DraftI.pdf>.
24. SmithKline Beecham. Draft submitted to JAMA, July 30, 1999. <http://www.healthyskepticism.org/documents/documents/19990726JAMAsubmission.pdf>.
25. SmithKline Beecham. Response to JAACAP Reviewers, November 3, 2000. <http://www.healthyskepticism.org/documents/documents/responsetoJAACAPReviews.pdf>.
26. SmithKline Beecham. Seroxat/Paxil adolescent depression position piece on the Phase III clinical studies, October 1998. <http://www.healthyskepticism.org/documents/documents/19981014PositionPiece.pdf>
27. SmithKline Beecham. E-mail. White to Hood, March 5, 2001. <http://www.healthyskepticism.org/documents/documents/20010305WhitetoHood.pdf>.

28. Forest Laboratories. Forest Research Institute. Study Report for Protocol No. CIT-MD-18. April 8, 2002. <http://industrydocuments.library.ucsf.edu/drug/docs/jpmf0220.pdf>.
29. Forest Laboratories. E-mail re: Pediatric data dated 10/15/2001. <http://industrydocuments.library.ucsf.edu/drug/docs/zymf0220>.
30. Forest Laboratories. E-mail re: Pediatric Manuscript dated 12/17/01. <http://industrydocuments.library.ucsf.edu/drug/docs/yymf0220>.
31. Wagner KD, Robb AS, Findling RL, Jin J, Gutierrez MM, Heydorn WE. *A randomized, placebo-controlled trial of citalopram for the treatment of major depression in children and adolescents*. Am. J. Psychiatry 2004; 161(6): 1079–1083.
32. Jureidini J. Declaration of Dr. Jon Jureidini, November 15, 2012. <http://industrydocuments.library.ucsf.edu/drug/docs/znmf0220>.
33. Forest Laboratories. E-mail re: FDA letter, March 14–15, 2000. <http://industrydocuments.library.ucsf.edu/drug/docs/#id=kjbn0225>.
34. Martin A, Gilliam WS, Bostic JQ, Rey JM. Letter to the editor. *Child psychopharmacology, effect sizes, and the big bang*. Am. J. Psychiatry 2005; 162(4): 817.
35. Wagner KD, Robb AS, Findling RL, Jin J. *Dr. Wagner and colleagues reply*. Am. J. Psychiatry 2005; 162(4): 819.
36. Complaint of Plagiarism of the Nemeroff *et al.* <http://psychrights.org/research/Digest/Science4Sale/110708EthicsComplaintAgainstEvansGyulaiNemeroffSachsBowdenetal.pdf>.
37. Testimony from *Kilker v. SKB trial*, October 2009. <https://www.industrydocumentslibrary.ucsf.edu/drug/docs/#id=rxpd0217>.
38. Senate Report on Ghostwriting in Medical Literature, June 24, 2010. <http://www.grassley.senate.gov/sites/default/files/about/upload/Senator-Grassley-Report.pdf>
39. SKB Clinical Trials Website Result Summary for Study 29060/352 updated 9 March 2005. <file:///C:/Users/j/Documents/Penn%20Items/GSK%20PAR%20352%20Results%20Summary%20Mar%202005.pdf>.
40. Complaint of Scientific Misconduct against Dwight L. Evans, Laszlo Gyulai, Charles Nemeroff, Gary S. Sachs, Charles L. Bowden *et al.*, July 8, 2011. <http://blogs.nature.com/news/files/2012/06/Jay-Amsterdam-today-filed-a-24-page-complaint-with-the-Office-of-Research-Integrity-at-the-US-National-Institutes-of-Health.pdf>.
41. Nemeroff CB, Evans DL, Gyulai L, Sachs GS, Bowden CL, Gergel IP *et al.* *Double-blind, placebo-controlled comparison of imipramine and paroxetine in the treatment of bipolar depression*. Am. J. Psychiatry 2001; 158(6): 906–912.

### Disclosure

*The authors are not members of any industry-sponsored advisory board or speaker's bureau, and have no financial interest in any pharmaceutical or medical device company. Dr. Amsterdam received legal support from Baum, Hedlund, Aristei & Goldman of Los Angeles, California for his Complaint of Scientific Misconduct against Dwight L. Evans, Laszlo Gyulai, Charles Nemeroff, Gary S. Sachs and Charles L. Bowden provided to the United States Department of Health and Human Services Office of Research Integrity in July 2011. Drs. Amsterdam and Jureidini were engaged by Baum, Hedlund, Aristei & Goldman as experts in the Celexa and Lexapro Marketing and Sales Practices Litigation, and Dr. Jureidini in the case of Beverly Smith vs. GlaxoSmith Kline. Dr. McHenry was also engaged as a research consultant in these two cases. Dr. McHenry has been a research consultant for Baum, Hedlund, Aristei & Goldman since 2003.*

***Acknowledgements***

*The authors wish to thank the DIDA website posting the declassified CIT-MD-18 documents on their website. The views expressed herein are those of the authors alone and not necessarily those of any other person, firm or entity. Due to journal limitations of space, this review does not permit a full exposition of the extent of the corruption described in court evidence and testimony related to these studies. The authors warrant that findings have been reported fairly and non-selectively.*

Address: Jay D. Amsterdam  
Depression Research Unit  
Department of Psychiatry  
University of Pennsylvania Perelman School of Medicine  
Philadelphia, Pennsylvania USA  
jamsterd@penmedicine.upenn.edu