DAVID HEALY – JULY 30TH
APPENDIX 1: ZOLOFT EFFICACY

FDA

The process of gaining FDA approval for a new drug involves several steps. First, the company must conduct laboratory testing in animals to determine whether the drug will be safe and, to some extent, effective. If relatively safe, the company submits an investigational new drug ("IND") application to FDA to test the product with human subjects.

The human studies involve clinical trials carried out in three phases—Phase I, II, and III studies. Phase I studies are conducted in healthy volunteers. Phase II and III studies test for safety and effectiveness of the drug for specific indications and patient populations.

After the clinical trials are completed, the company compiles the data in a new drug application ("NDA"). FDA reviews the: (1) safety and effectiveness in the drug's proposed use; (2) appropriateness of the proposed labeling; and (3) adequacy of manufacturing methods to assure the drug's strength, quality, and identity. Although the FDA reviews the NDA, the company remains responsible for ensuring that the drug is manufactured, promoted, and labeled correctly.[1]

When a drug is approved by the FDA, it means the drug manufacturer satisfied the regulatory requirements set forth in the Food Drug and Cosmetic Act ("FDCA").

The FDA does not conduct its own clinical trials, and can only deny an application if it finds the application lacks "substantial evidence that the drug will have the effect it purports or is represented to have[.]" *Id.* at § 355(d)(5). The FDCA mandates that the FDA approve an application *unless* it finds the application lacks substantial evidence of efficacy. "Substantial evidence" is defined under 21 U.S.C. § 355(d) as

> [E]vidence consisting of adequate and well-controlled investigations … on the basis of which it could fairly and responsibly be concluded by such experts that the drug will have the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the labeling or proposed labeling thereof. If the Secretary determines, based on relevant science, that data from *one adequate and well-controlled clinical investigation and confirmatory evidenc*e (obtained prior to or after such investigation) are sufficient to establish effectiveness, the Secretary may consider such data and evidence to constitute substantial evidence[.]

Zoloft

Zoloft (sertraline) is a selective serotonin reuptake inhibitor ("SSRI") in the same class of drugs as Prozac (fluoxetine) and Paxil (paroxetine). There is no evidence that these drugs correct an abnormality in people who are depressed and considerable evidence that they may make brains more rather than less abnormal.

In the early 1970s, Pfizer's psychoactive compounds were based on a thiothixene structure. From this they produced tametraline, a dopamine reuptake inhibitor. The initial research into

[1] *See Wyeth v. Levine*, 555 U.S. 555, 570 (2009) (holding that, regardless of any FDA approval, pharmaceutical manufactures bear sole responsibility for the sufficiency of a drug label).

tametraline was halted because of stimulant effects in animals. In 1977 a series of drugs were derived from tametraline, one of which was a serotonin reuptake inhibitor, later named sertraline.

In the 1980s, sertraline underwent human clinical trials for the treatment of depression, and in 1991 FDA approved sertraline for the treatment of major depressive disorder ("MDD"). It was released by Pfizer under the brand name Zoloft in 1991. In later years, Zoloft gained approval for promotion for obsessive-compulsive disorder, panic disorder, post-traumatic stress disorder, premenstrual dysphoric disorder, and social anxiety disorder.

Effective?

Under the 1962 statute a drug has to be demonstrated to be effective. This in practice means it must demonstrate that it is superior to placebo, in a double-blind randomized placebo-controlled trial (RCT). RCT participants are divided (unbeknownst to them) into a treatment group, where the participants receive the drug, or a control group, where they receive a placebo. If there is no meaningful difference between the treatment and control group, then the drug is not considered effective.

Unlike other ailments, where objective measurements are obtainable through blood and tissue samples, there does not exist a physiological test for determining the extent of a person's depression or other nervous disorder. Researchers could rely on the number of dead people in each group, or a functional measure such as return to work or on the reports of patients concerning their depression. This latter is done using rating scales completed by doctors designed to measure the severity of a patient's depression.

But changes in a rating scale measure do not necessarily demonstrate effectiveness. For example, in a double blind trial alcohol might well beat placebo on a depression rating scale measure as key elements of the measure involve the person being less anxious and getting to sleep quicker.

Furthermore, a change on a rating scale in trials that last 6 weeks, which was the average length of Zoloft trials, demonstrates a treatment effect rather than efficacy or effectiveness.

It is not clear that a decision to license an antidepressant on this basis is in fact legal. It is not clear that the licensing of Zoloft as an antidepressant has conferred any benefits beyond those that might have been obtained had alcohol been so licensed.

The consequences of this licensing are pursued below. This report will lay out minimal criteria for efficacy and make it clear that Zoloft did not reach these criteria.

ZOLOFT'S LACKLUSTER EFFICACY DATA

Quite aside from the above, Pfizer knew long before it was approved by the FDA, that Zoloft had a problem meeting even minimal treatment effect criteria. In a March 8, 1984 memorandum,[2] a Pfizer employee, discussing a Zoloft efficacy report that needed to be submitted to the FDA, stated that the data "are not in favor of sertraline as the placebo group has the most beneficial response in all of these instances. As for the HAM-D [a depression rating scale used in clinical trials to measure changes in depression] . . . placebo still seems to be the most effective group in this subset of patients[.]"

Most of the early Zoloft studies proved to be negative, or "failed". In the majority of these

[2] All Pfizer documents referenced here have been produced in litigation and unsealed, and are available from D Healy.

studies, there was no significant difference between Zoloft and placebo in relieving depression. In some studies, placebo outperformed Zoloft in treating depression. In other studies, Zoloft was shown to be less effective at treating depression than non-drug therapies such as St. John's wort or daily exercise.

Because of Zoloft's lackluster efficacy data, Pfizer had difficulties in gaining approval from European regulators. In an April 18, 1991 memorandum, a Pfizer employee expressed "serious concerns regarding the approval of sertraline in key European countries." According to an April 11th 1991 memorandum, Pfizer "received an unfavorable review in a number of countries. The common key issue is that regulators are not convinced of sertraline's efficacy versus placebo." Pfizer recognized that an "analyses of [placebo controlled U.S. studies] strongly indicate that they are not highly convincing of sertraline efficacy versus placebo and will not provide the strong database required to overcome regulatory obstacles." Pfizer, therefore, decided to create a "strongly positive, placebo controlled study . . . to ensure regulatory success." This was to be accomplished by designing a study "to enhance the probability of success drawing on the knowledge gained from past trials[.]"

In the United Kingdom (UK), Pfizer improperly sought help in obtaining approval of sertraline from a consultant to the regulators reviewing Zoloft, Dr. Stuart Montgomery. According to an April 24, 1989 memorandum, Dr. Montgomery told Pfizer he would remain a "disinterested party" at the U.K. Regulatory Agency until Pfizer appealed any negative decision and, thereafter, "he would be happy to act as an advisor to Pfizer and declare an interest." In addition, according to an April 26th 1989 memorandum, Dr. Montgomery provided Pfizer with inside information that the UK regulators' decision on Zoloft was "borderline" at that time and that the company "should be wary of providing them with efficacy data against placebo[.]"

As part of Pfizer's NDA for Zoloft, Pfizer submitted five placebo-controlled trials.[3] These included: Study 101, Study 103, Study 104, Study 310, and Study 315.

Of the five studies submitted in support of Zoloft's efficacy, Studies 101, 310, and 315 were either negative or failed trials, Study 104 was statistically positive, and Study 103 was mixed. None of the studies offered even modestly persuasive information that Zoloft in fact is effective *clinically*. A typical depressed patient on the Hamilton Rating Scale might have a score of 30 points. A patient who appears much better, restored to normal, would score around 5. The average change in studies in which Zoloft "beat" placebo is of the order of 2-3 points. The overall benefit for Zoloft across all studies is even less.

In Study 101, *placebo outperformed Zoloft* in treating depression, albeit not to a statistically significant degree. On average, patients taking placebo were 3.5 points better on the Hamilton Rating Scale for Depression ("HAM-D")[4] scale compared to those taking 50mg,

[3] In addition to the five efficacy trials examined by the FDA, Pfizer submitted the data from a relapse prevention trial they had conducted—Study 320. Relapse prevention trials are not designed to establish initial efficacy and are not required for drug approval. They are intended to evaluate the benefit of continuing patients on the medication after an initial positive response. In study 320, all patients were given Zoloft for eight weeks, with no placebo control group. Those who responded were then either kept on Zoloft or switched to placebo. The purpose was to compare the rate of relapse on Zoloft to that on placebo, following an initial positive response to Zoloft. FDA evaluators expressed grave concerns about the methodology of this trial and concluded that it did not provide support for the long term efficacy of Zoloft.

[4] The HAM-D scale is a multiple item questionnaire used to measure a person's perception of depression. It is usually composed of 17-29 questions where the patient rates specific areas on a 0-5 point scale. A person's HAM-D scale rating can be anywhere

100mg, 200mg, and 400mg doses of Zoloft.

In Study 310, *placebo again outperformed Zoloft*. The benefit in HAM-D scores of placebo patients over Zoloft-treated patients was on average 0.8 points.

In Study 315, Pfizer compared amitriptyline (Elavil), an older Merck antidepressant with Zoloft. On average, patients taking amitriptyline outperformed placebo on average by 3.8 points on the HAM-D scale. Zoloft was *not better than placebo* either statistically or clinically. The results indicated that patients taking Zoloft outperformed placebo on average by only 1.2 points on the HAM-D scale.

Pooling together the data from Studies 101, 310, and 315, the three clearly negative Zoloft trials, on average people taking placebo outperformed those taking Zoloft by 0.85 points on the HAM-D scale.

Study 103 had mixed results. Study 103 contained three treatment groups (50mg, 100mg, and 200mg) and one control group. Zoloft was statistically superior to placebo at the lowest dose (50mg), with a mean difference of 2.7 HAM-D points. The results for the other doses, however, were not statistically significant. At 100mg, the mean difference from placebo was 2.1 points, and at 200mg, the difference was 1.5 points.[5] Pfizer combined the results of all the treatment groups to find a positive outcome, a problem that was remarked upon by at least one FDA reviewer: "[Pfizer] relied upon an analysis which combined data from these three separately randomized groups and treated them as one group." Thus, it is unclear if Study 103 was statistically positive since it only provided statistically positive results for the 50mg Zoloft treatment group. However, regardless of the mixed statistical results, the results of Study 103 failed to demonstrate clinical significance at any dosage.

In the ordinary course of events if a drug helps at one dose it might be expected to help more at a higher dose and in fact FDA would have licensed Zoloft had the 100mg and 150mg doses consistently outperformed the 50mg dose.

Study 104 was the only study that showed consistent statistically positive results for Zoloft compared to placebo. On average, patients taking Zoloft experienced an additional 3.5 point[6] increase on the HAM-D scale than those taking placebo. The 3.5 difference barely satisfied the criterion for clinical significance, but the overall effect size of the study was insufficient.

Pooling together the data from Study 104 and Study 103, the only two clinical trials demonstrating at least some support for Zoloft over placebo, reveals a small treatment effect that is below the criteria of clinical significance. The mean difference between Zoloft and placebo in these two studies was 2.83 points on the HAM-D.

---

between 0-70 depending on the scale used.

[5] The pattern of findings suggests that the higher the dose, the worse the effect of Zoloft. A similar pattern was observed in Study 310.

[6] It should also be noted that a 3.5 HAM-D difference is still a rather small difference in clinical terms. Scores on the 17 item HAM-D range from 0 to 53. A 6 point difference can be obtained merely by changes in sleep patterns, without changes in any other symptoms. Thus, even with a 3.5 point difference over placebo, these results are marginal at best.

THE FDA'S APPROVAL OF ZOLOFT

On November 19, 1990, the Psychopharmacological Drugs Advisory Committee[7] ("PDAC") for the FDA convened to discuss Pfizer's Zoloft NDA. The focus of the meeting was to discuss Zoloft's efficacy and safety. However, prior to the meeting, according to an internal Pfizer document dated September 5, 1990, Dr. Paul Leber, Director of the FDA's Division of Neuropharmacological Drug Products, informed Pfizer that the he would be "solicit[ing] the support of the Advisory Committee that the two key pivotal depression studies are adequate and well controlled." He pointed out problems with the studies Pfizer had submitted to prove Zoloft's efficacy and safety, but assured Pfizer that he thought he could convince the Committee that the studies were sufficient to recommend approval.

The Zoloft PDAC consisted of: Chair, Daniel E. Casey, M.D., a psychiatrist of the Veteran's Affairs Medical Center in Portland, Oregon; John M. Davis, M.D., a psychiatrist of the Illinois State Psychiatric Institute; Markku I. Linnoila, M.D., PhD, the Chief of Clinical Studies at the National Institute of Alcohol Abuse and Alcoholism; Robert F. Prien, Ph.D., the Director of Clinical Psychopharmacology, Division of Clinical Research of National Institute of Mental Health; Javier I. Escobar, M.Sc, M.D., a professor of psychiatry, University of Connecticut; Jeffrey A. Liebernan, M.D.; a physician and researcher for the Long Island Jewish Medical Center and Albert Einstein College of Medicine; Robert Mark Hammer, Ph.D., a psychiatry and a statistics professor at the Medical College of Virginia; Linda Frances Hezel, Ph.D., a professor of nursing at the University of California, Kansas City; Carol Ann Tamminga, M.D., a professor of psychiatry at the University of Maryland; Paul Leber, M.D., the Director of the FDA Division of Neuropharmacological Drug Products; and Thomas P. Laughren, M.D., a Group Leader for Psychopharmacology at the FDA.

The meeting began with presentations from FDA personnel about Zoloft's efficacy and safety data. The efficacy data was presented by Dr. J. Hillary Lee. During the efficacy presentation, Dr. Lee discussed Study 103 and Study 104 in detail, and presented slides about the results of the trials. Dr. Lee expressed some concerns with Study 103 since there were large numbers of "drop-outs" during the trial. Dr. Lee concluded her presentation by stating that "the sponsor has provided two studies, one of which is more consistent than the other, to demonstrate the effectiveness of Sertraline [Zoloft]."

Following Dr. Lee's presentation, Dr. Ed Nevius presented a statistical review of clinical trial data related to efficacy. Dr. Nevius pointed out that Study 103 and Study 320 had several "problems with regard to the design of two of the studies[.]" In discussing Study 103, Dr. Nevius commented that Pfizer had "combined data from these three separately randomized groups and treated them as one group." Dr. Nevius gave some general background about what statistical issues the FDA was looking for in evaluating Study 103 and Study 104. After Dr. Nevius, Dr. Japobrata Choudhury presented his statistical findings regarding Study 103.

Next, Dr. Ryder, on behalf of Pfizer, gave a presentation. During the presentation, in response to a declaration that the clinical evidence clearly supported efficacy, Dr. Prien interrupted, stating:

> In terms of clinical results from studies 103 and 104, we have been provided

[7] The PDAC of the FDA reviews and evaluates data concerning the safety and effectiveness of marketed and investigational human drug products for use in the practice of psychiatry and related fields and makes appropriate recommendations to the FDA Commissioner. Recommendations from the PDAC are generally adopted by the FDA Commissioner. The PDAC is comprised of 11 voting members including the Chair that are selected by the Commissioner or designee from among authorities knowledgeable in the fields of psychopharmacology, psychiatry, epidemiology or statistics, and related specialties.

> with mean change scores from baselines and the outcome of statistical analyses for the difference in mean change scores between drug and placebo. However, it is difficult to determine the clinical significance of the statistically significant differences in mean change scores, particularly when you are looking at differences in change of only 2 or 3 points on the HAM-D total score . . . which is not a tremendous difference.

The committee members began to engage in a discussion on the topic, but Dr. Leber interrupted and asked the committee to wait until the presentations were completed. Immediately before opening up the discussion to the PDAC members, Dr. Leber explained their limited role in evaluating the Zoloft NDA:

> I think you have to understand that when we face an application from a regulatory perspective, we are asked to face what the law requires us to do . . . We are required to approve the drug unless we find that the tests submitted failed to contain substantial evidence of efficacy. That means more than one investigation which is adequate and well controlled which would allow experts — experts by experience, training and background — to reach a conclusion that the drug is effective.

Throughout the discussion, various committee members raised serious concerns about Zoloft's efficacy. In fact, nearly the entire discussion centered on whether Zoloft was proven to be effective in treating depression. When the PDAC members asked what they should consider as proof of efficacy, Dr. Leber explained that he had "no idea what constitutes proof of efficacy, except on the basis of what we, as a Committee, agree on an as *ad hoc* case as there needs to be. You can be guided by the past but the inference is an abstraction – what is an antidepressant?"

Notwithstanding, PDAC members criticized Zoloft's efficacy data. For instance, Dr. Hammer, a statistician, highlighted issues related to the two "positive" trials in light of the negative ones:

> Can I toss a few statistical issues as the statistician here? In a sense, what we have is not just an effect size problem but, in a sense, what we are trying to wrestle with is sort of an intuitive multiple comparison problem. If all we had was the two out-patient studies and they fairly clearly showed some Sertraline effect, we would have what I interpret as the criteria necessary for us to say go ahead and approve the drug. That is, we have more than one well-documented study that demonstrates an effect.
>
> So the question is how do we interpret these two positive results in the context of several more studies that fail to demonstrate that effect? I am not sure I have an answer to that but I am not sure that the law requires me to have an answer to that – fortunately or unfortunately. That would mean, in a sense, that [Pfizer] could just do studies until the cows some home until he gets two of them that are actually statistically significant by chance alone, walks them out and says that he has met the criteria.

Another major concern among the PDAC members was whether the observed benefit seen in Zoloft was not just statistically significant, but also clinically significant. As Dr. Prien explained:

> In terms of the regulatory charge outlined by Paul that the sponsor provided evidence from more than one adequate study that Sertraline is effective for the treatment of depression, this group seems to agree that for outpatients we

> have established efficacy. I am not so sure. I would just like to throw this out because, as Paul states, there are no set boundaries, no standards for what constitutes a clinically significant difference or treatment outcome.
>
> I am somewhat less convinced about study 103, where I believe that the differences in mean change scores between drug and placebo are marginal perhaps, at best, in looking at all the outcome criteria scales...
>
> But I think this is something that the Committee has to deal with, that is, whether they actually think that the efficacy data are sufficient to classify it as more effective than placebo, ***not only statistically but clinically***.

Dr. Hammer explained the difference between clinical and statistical significance:

> With sufficient subjects, a very small difference between several groups may well be statistically significant but for the size of the affect the difference between groups may be so small as to be unimportant in terms of treatment. Suppose I were working on some sort of a drug that influenced survival in cancer patients and I could demonstrate with sufficient subjects that the drug increased life expectancy from 435 days to 435.2 days. That might be statistically significant but it would be clinically unimportant.

In an effort to assuage the PDAC members, Dr. Leber stated:

> We are always in a position of trying to make a fair judgment, knowing that we have to weigh the requirements of the law, the expectations of a public that wants freer access to new and effective drugs, even if they are not necessarily as potent on a milligram basis or even in terms of the size of the treatment effect as others. They may have other advantages. It is very hard to judge a drug on a single dimension….
>
> At the end of that, we have to go back to the regulatory charge that I raised. It is not that they are entitled to every claim, every superlative ever made, but is the application, as submitted, such that we have a right to conclude it does not have evidence of safety for use; it does not have evidence of efficacy or it is inadequately labeled. I am not counting now the chemistry requirements.
>
> If we can reach those conclusions, we can reject the application. If we cannot reach those conclusions, ***you have to approve the application***. That is where we are.

In response, Dr. Prien stated:

> Let me throw this back to you, Paul. What would you do if, say, you get a statistically significant different with a difference in a HAM-D of perhaps one-tenth or three-tenths of a point; a CGI score that is just slightly different? What would you do with this, Paul? Are we going to fall back to the position that statistical significance is enough to demonstrate efficacy or are you going to leave this in the hands of a committee to try to judge whether, in their opinion, this is a clinically effective difference?

After a lengthy discussion, the PDAC moved to vote on whether the Pfizer had met the minimum regulatory requirements. The PDAC voted on the question, “has the sponsor provided evidence from more than one adequate and well-controlled clinical investigation that supports the conclusion that Sertraline is effective for the treatment of depression[?]”

The PDAC voted six in favor, one against, with two abstaining.

Pfizer itself found it odd that the FDA did not question Pfizer about efficacy. According to a February 25, 1991 internal Pfizer email between two of its employees: "There were no questions on efficacy . . . I find it odd that FDA not at all questioning efficacy and there are significant questions raised by several European companies."

In an August 1991 FDA memorandum obtained through the Freedom of Information Act, Dr. Leber wrote about his recommendation that Zoloft be approved for marketing, stating that, "[i]n recommending this action, I have considered the fact that the evidence marshaled to support sertraline's efficacy as an antidepressant is not as consistent or robust as one might prefer it to be."

In another FDA memorandum, in December 1991, Dr. Leber acknowledged that other foreign regulatory agencies were not willing to allow Zoloft marketing in their countries due to Pfizer's inability to prove efficacy. Dr. Leber went on to state, "I do not believe we can successfully introduce similar, more demanding requirements domestically, at least until there is significant 'sea change' in our society's collective attitude towards Federal regulation of new drug approvals." Dr. Leber warned that the FDA's approval of Zoloft was likely to be challenged because the FDA is not "as demanding as it ought to be in regard to its standards for establishing the efficacy of antidepressant drug products."

Dr Leber left FDA some years later and sent up a consulting company which initially had Pfizer as its only client.

PFIZER "MARKETING" OF ZOLOFT

Pfizer's marketing of Zoloft involved portraying it as highly effective and suppressing all hint that the clinical trial data showed debatable effectiveness.

The Drug Label

Zoloft's drug label has never disclosed the clinical trial data required to properly understand Zoloft's efficacy in treating depression. The label lacks significant clinical trial information that consumers and prescribing healthcare professionals would need to determine, for themselves, if purchasing or prescribing Zoloft is worth the risks. When Zoloft first entered the market, its product label stated:

> ZOLOFT (sertraline hydrochloride) is indicated for the treatment of depression. The efficacy of ZOLOFT in the treatment of a major depressive episode was established in six to eight week controlled trials of outpatients whose diagnoses corresponded most closely to the DSM-III category of major depressive disorder.

This suggests that all clinical trials performed on Zoloft supported efficacy when, in fact, there were at least three negative or failed efficacy trials that indicated Zoloft could not outperform placebo.

Moreover, this disclosure of efficacy data is misleading because it suggests that numerous clinical trials support efficacy when there were, arguably, only two, at best and the experts were split on this issue

Zoloft's drug label is inherently misleading because it does not mention or discuss the numerous clinical trials in which Zoloft was shown to be no more effective than placebo.

Moreover, the drug label fails to disclose that the two clinical trials that supposedly demonstrated Zoloft's efficacy were clinically insignificant. Instead, Pfizer concocted a description which, through the artful omission of material facts, conveys the impression that Zoloft is a very effective treatment for depression.

Doctors and consumers deserve to know what Zoloft's efficacy truly is and decide, in light of Zoloft's substantial negative efficacy information, whether purchasing Zoloft is worth the risks. Absent having knowledge of Zoloft's negative efficacy information, no reasonable consumer or prescriber can make an informed decision about whether to purchase or prescribe Zoloft for depression. In other words, by omitting this material information, Pfizer has robbed consumers and prescribing healthcare professionals of having sufficient information to properly decide whether to purchase or prescribe Zoloft.

Requiring the disclosure of negative efficacy information on a label has been endorsed by the FDA. The omission of negative trial data is material since physicians need to evaluate the positive data in relation to the negative data in weighing whether the prescription is suitable for a patient. As Dr. Leber stated in a memorandum to the Department of Health and Human Services regarding the disclosure of negative trials associated with the antidepressant Celexa (generically known as citalopram)[8]:

Leber:

> "One aspect of the labeling deserves special mention. The Clinical Efficacy Trials subsection within the Clinical Pharmacology section not only describes the clinical trials providing evidence of citalopram's antidepressant effects, but makes mention of adequate and well controlled clinical studies that failed to do so. I am mindful, based on prior discussions of the issue, that the Office Director is inclined toward the view that the provision of such information is of no practical value to either the patient or prescriber. I disagree. I believe it is useful for the prescriber, patient, and 3rd party payer to know, without having to gain access to official FDA review documents, that citalopram's antidepressants effects were not detected in every controlled clinical trial intended to demonstrate those effects. I am aware that clinical studies often fail to document the efficacy of effective drugs, but I doubt the public, or even the majority of medical community, are aware of this fact. ***I am persuaded they not only have a right to know, but should know. Moreover, I believe that labeling that selectively describes positive studies and excludes mention of negative ones can be viewed as being potentially "false and misleading."***

Selective Publication of Clinical Trial Data

Pfizer's drug label is just one component of a larger marketing scheme designed to mislead consumers and healthcare professionals about Zoloft's efficacy. Specifically, Pfizer has engaged in selective and biased publication of Zoloft's clinical trials with the aim of promoting favorable studies and suppressing negative ones. Dr. Marcia Angell, the former editor of the New England Journal of Medicine and currently a senior lecturer in the department of social medicine at Harvard Medical School, has commented that "[m]any drugs that are assumed to be effective are probably little better than placebos, but there is no way to know because negative results are hidden." Marcia Angell, *Drug Companies & Doctors: A Story of Corruption*, N.Y. REV. OF BOOKS, (January 15, 2009) *available at* http://www.nybooks.com/articles/archives/2009/jan/15/drug-companies-doctorsa-story-of-corruption/.

Selective publication of clinical trial data allows Pfizer to promote its drug in a scientific

[8] Celexa was approved by the FDA in 1998 and its label includes a limited discussion of Celexa's negative efficacy information.

vacuum. Since the only data available to consumers and prescribers are positive, the perceived efficacy of Zoloft is one-sided and overstated. A study published in the New England Journal of Medicine exposed this practice. *See* Erick H. Turner, et al., Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy, 358 NEW ENG. J. MED. 252, 252-60 (2008). The study found "a bias toward the publication of positive results" and that a survey of published literature indicates that ninety-four (94) percent of clinical trial studies were positive, whereas only fifty-one (51) percent of the studies actually submitted to the FDA were positive. The authors state that "[s]elective reporting deprives researchers of the accurate data they need to estimate effect size realistically" and that, by "altering the apparent risk-benefit ratio of drugs, selective publication can lead doctors to make inappropriate prescribing decisions that may not be in the best interest of their patients and, thus, the public health." More importantly, an estimate of how much the impression of each drug's effectiveness was inflated by not publishing unfavorable studies revealed that Zoloft was the second to worst drug of those evaluated, with an estimated 64% inflated efficacy.

To gauge the extent of the problem here, consider this. There are a number of groups independent of pharmaceutical companies who attempt to establish what the evidence shows. The best known of these is the Cochrane Collaboration. In 2009, they analyzed the available publications on current antidepressants and concluded that Zoloft was the best or one of the best antidepressants to use (Cipriani A et al, Comparative efficacy and safety of 12 new generation antidepressants. Lancet 373, 746-758 (2009) . This result is achieved on the basis of a distortion of data in ghostwritten articles and the suppression of negative trials. When this is undone, as in the Turner analysis above, it becomes clear that Zoloft is the least "effective" of then available antidepressants – if it is effective at all.

Several of the most prestigious medical journals in the world, including the New England Journal of Medicine, the Lancet, and the Journal of the American Medical Association, recognized the problem of pharmaceutical control over study design and publication:

> We are concerned that the current intellectual environment in which some clinical research is conceived, study subjects are recruited, and the data are analyzed and reported (or not reported) may threaten this precious objectivity.
>
> [C]corporate sponsors have been able to dictate the terms of participation in the trial – terms that are not always in the best interests of academic investigators, the study participants, or the advancement of science generally.
>
> Investigators may have little or no input into trial design, no access to the raw data, and limited participation in data interpretation. These terms are draconian for self-respecting scientists, but many have accepted them because they know that if they do not, the sponsor will find someone else who will.

Editorial, *Sponsorship, Authorship, and Accountability*, 345 NEW ENGL. J. MED. 825-27 (2001).

Pfizer's Ghostwriting Program

Doctors rely heavily on published papers and the peer review process to gain insight into new and potentially life-saving advances in medicine, as well as to learn about potential dangers that were previously unknown. *See* Puneet Manchanda & Elizabeth Honka, *The Effects and Role of Direct-to-Physician Marketing in the Pharmaceutical Industry: An Integrative Review*, 5 YALE J. HEALTH POL'Y L. & ETHICS 785, 796 (2005). The success of this system is predicated on integrity and transparency. Doctors, regulatory officials, researchers, and even consumers make important healthcare and research decisions based on the belief that these studies are unbiased and accurate.

Pfizer understood that the best way to ensure the success of Zoloft was to convince the scientific and medical community that Zoloft was safe and effective by cultivating a body of "peer-reviewed research" to enhance Zoloft's credibility. To that end, Pfizer created a large-scale ghostwriting program. Pfizer would author, or have a medical communications company author, a study specifically designed to promote a marketing message, i.e., Zoloft's efficacy. Then, Pfizer would pay "key opinion leaders" ("KOLs") to put their names on the article and get the article published in specifically targeted medical journals. When the article appeared in the journal, there would be no indication of Pfizer's involvement.

Pfizer had an entire team devoted to the publication of positive medical journal articles about Zoloft. In addition, Pfizer worked with outside medical ghostwriting vendors to create a steady stream of Zoloft-positive medical journal articles.

Healy and Cattell conducted an analysis of Zoloft articles that were "coordinated" by a medical communications company called Current Medical Direction ("CMD"). Pfizer had hired CMD to promote Zoloft in the 1990s. David Healy & Dinah Cattell, *Interface Between Authorship, Industry and Science in the Domain of Therapeutics*, 183 BRITISH J. OF PSYCH. 22-27 (2003). According to the study, CMD coordinated over 85 medical journal articles about Zoloft during a three-year period. By early 2001, 55 of these articles had been published in journals such as the *New England Journal of Medicine*, *Journal of the American Medical Association* (*JAMA*), *Archives of General Psychiatry* and the *American Journal of Psychiatry*. Interestingly, all of the clinical trial results were favorable to Zoloft. The analysis found that "the CMD-linked articles report[ed] universally positive results" and that there were "significant discrepancies between published data and the raw data from the actual clinical trials." Most of the 85 articles appeared to have been authored by CMD and, in a number of instances, the authors were listed in Pfizer internal memos as "TBD" (i.e., "to be determined"). The study concluded that:

> The combination of distinguished journal, distinguished author, an efficient distribution system and sponsored platforms appears to have led to an impact on the therapeutics domain greatly in excess of 50% of the impact of the rest of the literature on sertraline. The impact of this literature on third-party payers and other interested parties is at present unquantifiable. The question of literature impact would seem to be tied closely to the nature of ghostwriting. Authorship lines from perceived opinion-leaders with minimal company representation and non-declaration of other non-academic authorship inputs increase the likelihood that these articles will be influential with prescribers and purchasers.

Corroborating the study's finding, an internal company document shows that CMD kept a "Zoloft publications scorecard" for Pfizer which contained a running list of in-progress medical journal articles with details of the status, the names of the designated authors (KOLs or "thought leaders" in their respective fields), and the ghostwriting vendor to be used.

An internal PowerPoint presentation prepared by Pfizer in 2000 states that the purpose of the publications program was to, among other things, "promote efficacy[,]" "[h]ighlight drug's superiority to a competitor(s)[,]" "[l]everage good will with academic investigators[,]" "[i]ncrease media and public perception of the drug and Pfizer[,]" and "[p]rovide tools for sales force to drive prescriptions based on data[.]" The PowerPoint explains that the "bottom line" in publication projects is to "optimize our ability to sell Zoloft[.]"

What proportion of the articles on Zoloft linked to Pfizer are likely to have been ghost-written? Close to 100%.

Paid-For Opinions and Conflicts of Interest

Studies have demonstrated that "[a]uthors who had financial relationships with pharmaceutical companies were significantly more likely to reach supportive conclusions than authors without such industry affiliations" and "reports suggest that industry may alter, obstruct, or even stop publication of negative studies." Bekelman J et al Scope and Impact of Financial Conflicts of Interest in Biomedical Research, JAMA 289, 454-465 (2003).

Dr. Marcia Angell, in reviewing a book about how conflicts of interest impact drug development, explains that "highly influential faculty physicians – referred to by the industry as 'thought-leaders' or 'key opinion leaders' (KOLs) . . . write textbooks and medical journal papers, issue practice guidelines (treatment recommendations), sit on FDA and other governmental advisory panels, head professional societies, and speak at the innumerable meetings and dinners that take place every year to teach clinicians about prescription drugs ..." Marcia Angell, *Drug Companies & Doctors: A Story of Corruption*, N.Y. Rev. of Books, (January 15, 2009) *available at* http://www.nybooks.com/articles/archives/2009/jan/15/drug-companies-doctorsa-story-of-corruption/. Dr. Angell explains that "[c]onflicts of interest affect more than research. They also directly shape the way medicine is practiced, through their influence on practice guidelines issued by professional and governmental bodies, and through their effects on FDA decisions[.]"

ZOLOFT or ALCOHOL

Based on the precedent set by Zoloft it would be entirely possible to get alcohol on the market as an antidepressant.

The prime factors that prevent companies bringing alcohol on the market as an antidepressant are first their inability to take out a patent on it for this purpose and second consumer familiarity. Consumer familiarity provides a source of competing information that companies cannot control. In contrast, the SSRIs were and still are unknown quantities to most people making it possible to manage the views of doctors and patients more readily.

The analysis presented below is designed to bring out exactly what is meant by the developments above. It has already been published in two academic publications (Healy D in Le Meilleur des Marches. In Borch-Jacobsen M (ed) Big Pharma, Les Arenes Paris, pp 399-426 (2013); & Healy D. Challenging the dominance of the pharmaceutical industry in psychiatry, In Higgins A, McDaid S (eds). Mental Health in Ireland, Gill MacMillan, Dublin, pp 251-269 (2014)) and been presented in over 20 academic conferences, including the American Psychiatric Association and others (list available from author) without at present any significant argument against its validity. All of the points listed below are supported by peer-reviewed publications.

The key point to bear in mind is that Alcohol is a psychotropic drug with a significant anxiolytic effect. Alcohol is very like benzodiazepine drugs which have already been put through "antidepressant" trials with in many instances better results than the Zoloft results.

There is no suggestion here that alcohol is an antidepressant. The point is that the procedures used for Zoloft would allow alcohol to be licensed as an antidepressant and majority opinion is almost certain to be that the mass and indiscriminate use of alcohol for this purpose would not be a good idea, might even be a public health disaster.

Getting on the Market

As the example of Zoloft above demonstrates, the regulatory requirements regarding clinical trials allow us multiple opportunities to get a positive result for alcohol. Only approximately

one third of the first 18 or so Zoloft trials were positive. Ultimately in the case of alcohol it appears we might only need to have 1 of 5 trials "positive" - two out of ten.

In any alcohol trials we run, we can use as our yardstick of "effectiveness" not lives saved or people returning to work or people objectively performing better or people in their own estimation performing better but rather a 2-3 point change in the score on rating scales where sick patients often score 30 points and recovered patients score 5.

There are in fact more dead bodies in the Zoloft arm of Zoloft trials than in the placebo arm. There is no evidence of a superior rate of return to work on Zoloft. In fact as with alcohol there is good reason to think that Zoloft may impair cognitive function and produce less efficient working should the patient return to work.

In the case of the rating scales used, scales such as the Hamilton Rating Scale for depression are sensitive to the side effects of the drug so that simply taking the drug may produce a benefit on the scale whether or not there is a benefit for the underlying condition. The anxiolytic or sedative effects of alcohol would produce substantial benefits on the Hamilton rating scale for depression[9].

There are in fact four classes of rating scale. First, disease specific scales rated by doctors, like the Hamilton Rating Scale. Second, disease specific scales rated by the patient, like the Beck Depression Inventory. Third, global functioning scales rated by the doctor such as quality of life scales. Fourth, global functioning scales rated by the patient – again quality of life scales[10].

If we use rating scales as our measure of efficacy, an effective treatment should produce positive changes in all four domains. Not a single SSRI does.

Pfizer commissioned Robert Spitzer to produce a patient rated Quality of Life scale. This was developed and put into Zoloft trials. The results remain almost completely unpublished, as do the results for similar scales used by other SSRI producing companies – perhaps 100 trials or so where the data remain unpublished[11].

In the case of alcohol therefore we can simply depend on clinicians to find some benefit from the treatment for us on a rating scale that has an unproven relationship to the real world. We do not have to demonstrate consistent benefits across a range of domains and a range of perspectives.

We can compare alcohol to placebo rather than compare it against a treatment known to work for depression such as the older tricyclic antidepressants. In all trials where SSRI drugs where compared to older antidepressants the older drugs have proven substantially more effective – as in the trial comparing Zoloft to amitriptyline above. This is not just an anomaly of the Zoloft data, it is true of all SSRIs. Prozac, Paxil and Celexa have all been

[9] Healy D (2000). The assessment of outcome in depression. Measures of social functioning. Reviews in Contemporary Pharmacotherapy 11, 295-301.

[10] Healy D (2000). The assessment of outcome in depression. Measures of social functioning. Reviews in Contemporary Pharmacotherapy 11, 295-301.

[11] Healy D (2000). The assessment of outcome in depression. Measures of social functioning. Reviews in Contemporary Pharmacotherapy 11, 295-301.

demonstrably inferior to tricyclic antidepressants in trials that have been done[12]

We can also compare alcohol against placebo in a set of mild mood disorders rather than in a set of severe disorders. If the treatment were truly effective you would want to pick a severe group of conditions to demonstrate its superiority to placebo. But whatever about a rating scale benefit in mild depression, just as we would not expect alcohol to work for severe depression, neither Zoloft nor other SSRIs work for severe depression.

When running trials, we can improve the profile of alcohol by screening out anyone showing a good response to placebo or a bad reaction to alcohol during the first week of the study. These extraordinary developments that have crept into practice and been accepted by regulators place since Zoloft was first licensed make it almost impossible not to succeed in demonstrating alcohol is antidepressant.

Only some of the studies we undertake have to show a benefit for alcohol over placebo. If there are a lot of studies, perhaps even a preponderance of studies, in which alcohol fails to beat placebo, these can be discounted. Regulators, as demonstrated above, will conceal the fact that they have seen studies where alcohol has failed to beat placebo. As a result patients and doctors will not be aware of negative studies.

Although in our trials placebo may have effects approaching 80-90% of those found for alcohol, regulators and the academics we have co-opted to sell alcohol, will happily give doctors and the public the impression that 100% of the apparent benefits of alcohol for depression stem from the alcohol and none from placebo factors.

In a proportion of our alcohol studies, investigators may find out later that not all the patients actually existed – as has happened for Zoloft and other SSRI trials. Non-existent patients conveniently do not have troublesome side effects[13]. The trend towards non-existent patients is likely increasing as clinical trials more recently have been outsourced to Mexico, Eastern Europe, India and elsewhere. But even if this comes to light, it will make no difference. Once on the market, our license to sell alcohol as an antidepressant will not be revoked.

In one important study of aripiprazole's use for maintenance treatment of bipolar disorder (a particularly lucrative indication), it proved no better than placebo in c. 30 North American study centers, but turned out to be dramatically better than placebo in 2 Mexican centers. When all the centers were added together, the overall results for aripiprazole were marginally superior to placebo. The regulators approved this study. The published account of the study gave no indications that aripiprazole only "worked" in Mexico[14].

Given this license to help us achieve the right result in an alcohol trial, there seems little reason to think we could not achieve the outcome required.

In summary, we can therefore do studies in which more people die on alcohol than placebo, fail to get back to work on alcohol compared to placebo, do better on alcohol than placebo in perhaps no more than 33% of cases on our chosen rating scale that is almost certain to produce the right results, and we will still be able to market alcohol as an "antidepressant".

[12] Healy D. Let Them Eat Prozac. New York University Press, New York.
[13] Healy D. Let Them Eat Prozac. New York University Press, New York
[14] Rosenlicht N, Tsai AC, Parry PI, Spielmans G, Jureidini J, Healy D (2011). Aripiprazole in the maintenance treatment of bipolar disorder: A critical review of the evidence and its dissemination into the scientific literature. PLoS Medicine, 8, e10000434

Making the Market

After approval, in order to make our market, we need only publish the trials in which there were positive findings. But we can publish these multiple times, giving the impression that there were far more positive trials than in fact there were. We can aim at having up to 50 publications for each trial[15].

Our ghostwriters can even take a negative study and polish the results to make it look positive. The ghostwritten articles will never mention studies that have failed to show efficacy.

In due course when it comes to shaping the marketing campaign for alcohol, the data generated by these studies is almost free-floating content that can be molded into almost any shape we might wish. For instance, if an opportunity arises in the painkiller market, because another compound like Vioxx has run into trouble, some minimal benefits that may have been registered in the trial, in terms of feeling slightly better in painful situations, can be polished by ghostwriters into a series of articles that trumpet the analgesic qualities of alcohol in order to take advantage of any opportunity that has opened up.

This is almost close to exactly what has happened in the case of Cymbalta marketing for pain[16], and to Zoloft marketing for PTSD. Something similar could be engineered for alcohol depending on market conditions at the time of launch.

While we are busy getting wine on the market as an antidepressant, several other companies can file for product patents on whiskey, gin, brandy, beer, and port. The combined marketing of both our and other companies can encourage doctors to put patients on combinations of whiskey, gin, brandy and port, or even combinations of Scotch, and to keep patients on these combinations for extended or indefinite periods of time on the basis that all these antidepressants have been shown to "work" and what could be wrong with having a treatment resistant patient on several drugs all of which work.

Based on the published trials, guidelines will have to endorse alcohol for use in nervous disorders and perhaps have it as a first line therapy based on its excellent safety profile. Depending on how clever companies are with the trials they run, or choose to publish, the guideline makers in all seriousness may put whiskey forward a first line treatment with brandy second line and gin third line. Based on the evidence, this is more likely to happen if whiskey makers have a greater presence than gin-makers in the country in which the guideline is being written.

Endorsement by a guideline makes it almost mandatory for doctors to use alcohol. Furthermore we can likely engineer it, while a guideline recommending alcohol first may only do so for a 3 month period – in line with the evidence, the involvement of our consultants and marketing efforts mean that later guidelines and position statements will successively extend the recommended treatment period so that an increasing number of patients being treated by doctors who adhere to guidelines are likely to have alcohol for life.

We will also be able to engineer it that professional associations such as the American Association of Obstetricians and Gynecologists issue position statements in support of the need to ensure that women of child-bearing years are screened for depression and put on alcohol[17].

[15] Healy D (2012). Pharmageddon. U California Press, Berkeley.

[16] Spielmans G (2008). Duloxetine does not relieve painful physical symptoms in depression. Psychotherapy and Psychosomatics 77, 12-16.

[17] Healy D (2012). Pharmageddon. U California Press, Berkeley

Staying on the Market

When it comes to the side-effects of alcohol, our ghostwriters can hide these under terms such as ‘failure of response’ or perhaps list an initial side effect such as ‘nausea’ when in fact the individual had nausea, vomiting, followed by an epileptic convulsion. They can also simply fail to mention problems by saying they have only included those problems that appeared at a 10% rate or more.

When patients have an adverse effect on alcohol, such as a convulsion, we can dismiss this as anecdotal – not evidence based. In contrast, we can write up any dramatic improvement on alcohol during its early period on the market in both the academic and mainstream media, even featuring it on television and radio, under headings like “alcohol saved my life” with celebrity endorsements of the wonders of this new antidepressant.

The trials we have undertaken to bring alcohol to market only last for six to eight weeks. This is particularly helpful in terms of adverse events in that few of the problems that might be expected from alcohol emerge in a six to eight week period. There will not for instance be the doubling of suicidal or homicidal acts in alcohol trials seen in Zoloft and other SSRI trials.

In the case of any problems that emerge outside this time frame, we can argue that no placebo controlled data support the claimed adverse event, and that in the interests of being responsible both we and the doctors who use our drugs have to operate only on the basis of the scientific evidence.

If there is an increase in epileptic convulsions on alcohol compared to placebo in the course of our clinical trials but this is not statistically significant, we can rely on journals, regulators and academics to say there is no evidence for any increase in the rate of convulsions as has happened with the increased rates of suicidal acts linked to SSRIs in clinical trials.

The bias of doctors, helped by us, means that a culture will emerge early on in the use of alcohol that will attribute any of the difficulties people may have stopping alcohol to the nervous problem that was being treated in the first place rather than to dependence and withdrawal. We can be sure that twenty years after alcohol is first marketed that a majority of doctors will fail to recognize that it causes dependence. They will instead be likely to explain to patients that it’s just like insulin – their bodies are not producing enough alcohol and they need to continue treatment for life.

More astonishingly perhaps, regulators aware of this will put forward serious proposals to allow a licensing of alcohol or Zoloft on the basis of trials in which patients feel worse when the drug is stopped.

In the case of pregnancy, this bias and our marketing means that we should be able to make alcohol one of the most commonly prescribed drugs in pregnancy within a few years. And indeed compared with SSRI antidepressants, a glass of wine per day appears relatively safe. We can educate doctors to tell women who avoid coffee, soft cheeses etc. that leaving their nerves untreated will harm their babies.

Based on the data from antidepressant marketing, it should only take about a decade to

move from a less than 1% use of antidepressants during pregnancy to a 10-15% use[18].

Finally, we know from past experience with other drugs that in a few years' time alcohol is likely to be linked to suicide and perhaps violence. We have a number of academics who we can enlist to produce graphs to show that as alcohol consumption has gone up that suicide and violence rates have fallen in countries such as the United States. This is in fact the case but no-one at present puts this drop down to alcohol use. We can get our experts to make this link. And we can depend on the editors of leading journals to refuse to publish any correspondence that might be critical of studies like this[19].

We can organize for cost utility analyses as thick as telephone books to demonstrate that the cost of alcohol is minimal compared to the quality of life gained. Provided the analysts and economists stick to the published ghostwritten papers and do not attempt to access any sequestered data, we can show that if governments pay for widespread access to alcohol that there will be a net benefit to society.

As an antidepressant, alcohol will be made available on prescription only. This will transform it from a stranger into a neighbor. Taking an antidepressant becomes taking the advice of our doctor. We are in general wary of strangers and comfortable with neighbors. We neglect the fact that we are most likely to be abused or harmed by neighbors or relatives. Neighbors and relatives are familiar and we think we can manage the risks.

We have a feel for the traditional risks of alcohol but far from treating therapeutic alcohol or other new drugs as strangers and regarding them as dangerous and risky, mediated through doctors, we treat prescription only drugs as safer than drugs like alcohol or nicotine, even though prescription-only drugs are so precisely because we have every reason to think they will be riskier than drugs like alcohol.

For instance, we now regard prescribed amphetamines as safe to give to children with ADHD, even toddlers, while the authorities still jail others for possession of street amphetamines on the basis of the risks they pose.

For the patient, getting treatment from a doctor suspends the natural caution that our consumers might feel about taking our new chemical.

For the doctors, now that prescribed alcohol has been tested in protocols in which it looks safer than and as effective as SSRIs, even though doctors know what the risks of traditional alcohol are, they are it seems prepared to act as though prescribed on-patent alcohol comes risk free. This is partly because in the case of prescribed alcohol in contrast to traditional alcohol, doctors never get a hangover from prescribed alcohol and never crash their automobiles because of it.

It is becoming increasingly clear that making a new drug like alcohol available through doctors is a way to hide hazards such as liver failure or lung cancer on average for 10 to 15 years from the time that people in the street have begun to claim that their liver failure or lung cancer stems from our drug.

Not only can the medical profession be depended on to deny a link while patients are reporting a link but even after regulators put black box warnings on alcohol about a risk,

---

[18] Healy D, Mangin D, Mintzes B (2010). The ethics of randomized placebo controlled trials of antidepressants with pregnant women. Internat J of Risk and Safety in Medicine 22, 7-16 DOI 10.3233/JRS-2010-0487.

[19] Healy D (2012). Pharmageddon. U California Press, Berkeley

even if that risk is a lethal one, most doctors will still deny that this risk happens.

Doctors finally provide us with significant insurance against product liability. In the event that a doctor testifies that he would have given alcohol no matter what the warning on it, we are legally immune to any product liability actions stemming from its use.

Conclusions

This is a serious effort to look at whether Zoloft has in fact been proven to be effective. An effective treatment would show a clear benefit for the drug in all trials done. This has not been the case with Zoloft.

An effective drug would show a benefit in terms of lives saved or a restoration of function. This has not been the case with Zoloft. The opposite appears to be the case.

If rating scales are used, an effective drug would show a benefit across all four major domains of measurement. This has not been the case with Zoloft.

It is difficult to imagine clinical trial data being sequestered for a truly effective drug. The Zoloft clinical trial data is sequestered.

Zoloft like alcohol has an anxiolytic effect. A good doctor could use this effect of Zoloft just as they could use a comparable effect of alcohol to good therapeutic effect. But this is not the same as demonstrating that the treatment is effective in general.

There is ***no data*** on which to make a case that the public health consequences of licensing Zoloft are likely to be less bad in terms of dependence, birth defects, suicidality and other medical problems arising from chronic Zoloft use than a licensing of alcohol would have produced. This gives a measure of the problems we face.

DAVID HEALY
July 30th 2014