13

Is the Efficacy of "Antidepressant" Medications Overrated?

Brett J. Deacon and Glen I. Spielmans

In 1987, the United States Food and Drug Administration (FDA) approved fluoxetine (Prozac) for the treatment of major depression in adults. Fluoxetine quickly became a cultural phenomenon and ushered in the modern "antidepressant era" (Healy, 1997). Manufacturer Eli Lilly marketed fluoxetine as a selective serotonin reuptake inhibitor (SSRI), a depression-specific magic bullet of sorts that purportedly corrected the serotonin imbalance theorized to cause depression. Cover stories in the popular media touted fluoxetine as a "medical breakthrough" (*Newsweek*; Cowley, Springen, Leonard, Robins, & Gordon, 1990) and a "wonder drug" (*New York Magazine*; Schumer, 1989). Peter Kramer's influential *Listening to Prozac* (1993) claimed the drug cured a host of psychological maladies and made some people "better than well." Fluoxetine became one of the first psychotropic medications to earn blockbuster status (Fitzpatrick, 2010), a designation achieved via US$1 billion or more in annual sales. Additional FDA-approved SSRIs such as paroxetine (Paxil, 1991) and sertraline (Zoloft, 1992) followed suit as blockbuster antidepressants and joined Prozac as household names.

The use of antidepressant medications soared following the release of fluoxetine. From 1988–1994 to 2005–2008, the percentage of Americans who took antidepressants increased 400% (National Center for Health Statistics, 2010). By 2005–2008, antidepressants were used by 10.8% of Americans aged 12 and older (Pratt, Brody, & Gu, 2011). Most of these individuals had taken them for more than 2 years, and 13.6% (approximately 3 million Americans) had taken them for 10 or more years. Antidepressants became the third most commonly used class of prescription medication of any kind in the United States, and the most commonly used drug class among adults aged 18–44 years (Pratt et al., 2011). The popularity of antidepressant

Psychological Science Under Scrutiny: Recent Challenges and Proposed Solutions, First Edition.
Edited by Scott O. Lilienfeld and Irwin D. Waldman.
© 2017 John Wiley & Sons, Inc. Published 2017 by John Wiley & Sons, Inc.

medications was accompanied by widespread endorsement of their purported mechanism of action. Consistent with promotion of the serotonin imbalance theory in direct-to-consumer advertisements (Lacasse & Leo, 2005), approximately 90% of Americans came to view depression as the product of a chemical imbalance that should be treated with prescription medication from a psychiatrist (Pescosolido, Martin, Long, Medina, Phelan, & Link, 2010).

Antidepressants are recommended first-line treatments for major depressive disorders in clinical practice guidelines based on reviews of the clinical trials literature. To illustrate, the United Kingdom's National Institute for Health and Care Excellence guideline (2009) recommends antidepressants as an initial treatment for patients with moderate to severe depressive symptoms, as well as for those with mild symptoms who have failed to respond to initial non-drug interventions. The American Psychiatric Association (APA, 2010) practice guideline recommends antidepressants as a first-line treatment for all depressed patients with mild-to-moderate major depression, and states that antidepressants "definitely should be provided for those with severe major depressive disorder" (p. 17). Treatment providers who prescribe antidepressants in accordance with clinical guidelines are engaging in "evidence-based medicine" (Sackett, 2005), which involves the use of evidence from randomized controlled trials in clinical decision-making.

The use of antidepressant medication to correct the presumed chemical imbalance that causes depression has been the dominant approach to the treatment of depression in the United States for more than two decades. This approach is so entrenched that it is difficult to imagine that it could be based on anything less than an unassailable empirical foundation. However, the conventional wisdom about antidepressants has been questioned in recent years by prominent critics armed with scientific data (e.g., Kirsch, 2010; Whitaker, 2010). The compelling nature of these data, and their incompatibility with the standard narrative, has prompted a critical reanalysis of medications marketed as "antidepressants." The purpose of this chapter is to contribute to this reanalysis. As we describe later, the dominant cultural story of antidepressant medications bears little resemblance to the available scientific evidence. Of greater concern is that it never has. Because Eli Lilly's fluoxetine initiated and is synonymous with the antidepressant era, it provides an ideal case study for a critical analysis of antidepressant medications.

Although issues surrounding the science underlying antidepressants are the subjects of this chapter, they are hardly unique. For example, second-generation "antipsychotics" (SGAs, aka atypical antipsychotics) have been similarly overhyped. While these drugs were initially hailed as possessing superior efficacy and safety than older "typical" antipsychotic drugs, such claims were largely derived from studies using biased research designs. For instance, haloperidol was the most common typical antipsychotic to which atypical antipsychotics were compared (Leucht, Kissling, & Davis,, 2009). Haloperidol carries an unusually high risk of causing abnormal movements characterized as extrapyramidal symptoms (EPS), so claiming that an atypical antipsychotic causes substantially lower rates of EPS than "older drugs" based on a comparison with a single drug notorious for causing EPS is

rather dubious. Further, haloperidol was often given in unnecessarily high doses, leading to increased adverse events and likely reduced efficacy, making the atypical antipsychotics appear safer and a bit more efficacious in comparison (Leucht et al., 2009). Results emphasizing superiority of the atypical antipsychotics were trumpeted far and wide, whereas less convenient results were sometimes hidden from public view (Spielmans & Parry, 2010). One team of leading reviewers opined, "Marketing by pharmaceutical companies has often promoted SGAs by smoke and mirrors. Many hopes in the SGAs, such as dramatically better efficacy, compliance, quality of life and no side-effects, have not been fulfilled (Leucht et al., 2009, p. 1600)." In the realm of anxiety treatment, publication bias has been demonstrated for paroxetine (Sugarman, Loree, Baltes, Grekin, & Kirsch, 2014). Turner (2013) summarizes evidence of publication bias for several drugs in the treatment of depression, bipolar disorder, schizophrenia, and autism.

Though likely exacerbated by commercial interests, issues pertaining to inflated psychotropic drug efficacy can be viewed in the broader context of poor replicability across many areas of science, including psychology (Ioannidis, 2012; Makel, Plucker, & Hegarty, 2012; see Chapters 1 and 2). A sobering recent analysis found that, when Food and Drug Administration inspections revealed likely or definite problems with the reliability of data in clinical trials, published versions of the clinical trials in medical journals almost always included these questionable data in their analyses and quite rarely mentioned any violations that FDA inspectors uncovered (Seife, 2015). Although fluoxetine is the target of critical analysis in this chapter, it is hardly an isolated case.

Fluoxetine: Creation of a Blockbuster "Antidepressant"

Prior to its approval by the FDA in 1987, fluoxetine was tested in five double-blind placebo-controlled acute efficacy trials with a total of 1,134 adult patients. These trials were conducted in support of Eli Lilly's goal of obtaining regulatory approval for fluoxetine in the treatment of adult patients with major depressive disorder. As described in Table 13.1, these trials had a number of problematic design features. Although these features are standard practice in industry-sponsored trials (Ioannidis, 2008; Leo, 2006; Safer, 2002; Spielmans & Kirsch, 2014), they compromise scientific integrity by biasing the results in favor of the active medication over placebo. Consistent with their function of serving Eli Lilly's commercial interests, the trials were designed to maximize the probability that fluoxetine would demonstrate a statistically significant advantage over placebo.

How the fluoxetine trials were conducted

Each fluoxetine trial submitted to the FDA included a placebo washout period, after which patients whose symptoms improved on placebo were excluded from the trial. In three trials, investigators also replaced patients who were not responding to

Table 13.1 Problematic design features in antidepressant trials conducted by pharmaceutical companies.

Feature	*Description*	*Effect*
Use of inactive placebos	Patients are randomly assigned to receive inert placebo or the active drug. Unlike inert placebos, drugs produce noticeable side effects. Patients warned about these side effects during informed consent, and researchers who assess patients' symptoms, are likely to guess the condition to which the patient has been assigned at a level that exceeds chance.	The double-blind is likely broken, thereby confounding drug effects with expectancy effects.
Failure to assess the double-blind	Researchers do not assess the extent to which patients and their assessors accurately guess the condition to which patients are assigned. The integrity of the double-blind is unknown.	Researchers cannot rule out the hypothesis that drug effects represent expectancy effects, rendering trial results uninterpretable.
Use of placebo washout period	All patients are assigned to placebo prior to the trial. Patients who improve during the first few weeks are excluded from the trial.	Drug efficacy is inflated to the extent that early placebo response predicts eventual placebo response in the trial.
Use of drug run-in period	Patients assigned to the drug condition who do not respond during the first few weeks, or have negative responses, are excluded from the trial.	Drug safety and efficacy are inflated.
Replacement of non-responders	Following the placebo washout period, all patients are assigned to the drug. Those who fail to improve during the first few weeks are excluded from the trial and replaced with early drug responders.	Drug efficacy is inflated to the extent that early drug response predicts eventual drug response in the trial.
Exclusion of patients with mild-to-moderate depressive symptoms	Patients with mild-to-moderate symptoms are excluded from the trial. Trials are conducted with patients whose average depressive symptoms are in the "very severe" range, among whom antidepressant efficacy is highest.	Drug efficacy is inflated.
Concurrent use of sedative medication	Patients are allowed or encouraged to take sedative medication during the trial to suppress adverse reactions.	Antidepressant effects are confounded with sedative effects. Drug efficacy and safety are inflated.
Problematic efficacy measures	The clinician-rated Hamilton Rating Scale for Depression (HRSD) is the primary outcome measure. The HRSD has problematic reliability and validity. Numerous items assess depression-nonspecific variables (e.g., sleep) that may be improved by drug side effects (e.g., sedation). Validated self-report measures such as the Beck Depression Inventory–II are rarely administered. Self-report measures tend to show lower antidepressant efficacy than clinician rating scales.	Drug efficacy is inflated.
Problematic side effect measures	Side effects are assessed using open-ended or non-specific questions rather than a checklist or direct questioning.	Drug safety is inflated.

fluoxetine after 2 weeks (Kirsch, Moore, Scoboria, & Nicholls, 2002). Thus, treatment outcomes for most acute efficacy trials of fluoxetine were based on data from patients who failed to respond to early placebo treatment and responded positively to early fluoxetine treatment. Under these conditions, even an ineffective antidepressant might demonstrate reliable benefits over inert placebo.

Fluoxetine produces numerous side effects. According to the FDA package insert, a partial list of common reactions (>5% frequency and at least twice that for placebo) includes anorexia, decreased libido, diarrhea, drowsiness, dry mouth, indigestion, impotence, insomnia, nausea, sore throat, rash, sweating, tremor, and weakness. For ethical reasons, participants enrolled in clinical trials are informed of these possible reactions during the informed consent process. Patients who are randomized to fluoxetine and experience the side effects about which they were warned are likely to conclude that they are taking the active medication. This conclusion amplifies the expectation of improvement in a placebo-controlled trial and potentially produces an "enhanced placebo effect" (Kirsch, 2010). Conversely, patients who do not experience the expected side effects are likely to have lowered expectations for improvement due to the perception that they are taking inert placebo. The likelihood that patients (and research personnel who assess them) can accurately guess their assigned treatment based on the perception of side effects is a serious confound in a placebo-controlled trial intended to be double-blind. In the absence of evidence that the double-blind was maintained, it is not possible to determine whether the apparent benefits of medication in a trial reflect the biological effects of the drug or an enhanced placebo effect caused by penetration of the blind. In other words, results of the trial are uninterpretable. A meta-analysis of several fluoxetine clinical trials found a strong correlation between the percentage of fluoxetine participants reporting adverse events and the advantage for fluoxetine over placebo ($r = 0.85$ for clinician-rated depressive symptoms, and $r = 0.96$ for self-reported depressive symptoms; Greenberg, Bornstein, Fisher, Zborowski, & Greenberg, 1994). This lends some credence to the possibility that unblinding due to adverse events may impact ratings of symptom severity.

Investigators did not assess the extent to which patients and/or study personnel were able to accurately guess treatment condition in any fluoxetine trial upon which the drug's regulatory approval was based. Remarkably, the integrity of the double-blind is almost never assessed in antidepressant trials (Even, Siobud-Dorocant, & Dardennes, 2000). When it is, patients and researchers can easily guess which treatment was received (Even et al., 2000). One strategy for maintaining the integrity of the double-blind is the use of an "active placebo," which mimics antidepressant side effects but does not produce therapeutic effects. Despite the appeal of this approach for increasing internal validity, it lacks appeal for commercial purposes because it yields small drug effects (Moncrieff, Wessely, & Hardy, 2004). Because fluoxetine investigators used inert placebo and failed to assess the integrity of the double-blind in each trial, the extent to which differences between fluoxetine and placebo are attributable to biological vs. psychological factors is unknown.

A unique design feature in fluoxetine trials is that patients were permitted to take sedative medication during the trial (Kirsch et al., 2002). This practice was encouraged because sedative medication suppressed the symptoms of "activation" (i.e., a state of extreme inner restlessness known as *akathisia*) evident in many patients. Internal Eli Lilly documents obtained by the *British Medical Journal* revealed that 38% of fluoxetine-treated patients experienced activation, compared to 19% of placebo-treated patients (Lenzer, 2005). These documents also described 12 suicide attempts among fluoxetine patients, compared with only one each among patients given placebo and imipramine (Healy, n.d.a). A 1989 internal Eli Lilly memo noted, "physicians should be advised that, in the absence of sedation, the risk of higher suicidality should be taken into account" (Baum, Hedlund, Aristei, & Goldman, n.d.a). Discovery documents indicate that trial investigators were pressured by company executives to reclassify suicidal events as "overdose" and suicidal thoughts as "depression" (Healy, n.d.b). Concerns about poor efficacy and suicidal events led the German regulatory authority to reject Eli Lilly's application for fluoxetine approval in 1985 (Baum et al., n.d.b).

Results of the fluoxetine trials

Table 13.2 presents results from the five acute efficacy trials of fluoxetine that served as the basis of its approval by the FDA. These data were reported by Kirsch et al. (2002), who obtained them via a Freedom of Information Act request. Three trials yielded a statistically significant advantage of fluoxetine over placebo on Hamilton Rating Scale for Depression (HRSD; Hamilton, 1960) change scores. Two trials did not. One study (Trial 27) included imipramine, a tricyclic antidepressant approved by the FDA in the treatment of depression in 1959. Imipramine produced significantly greater improvement in HRSD scores than fluoxetine. Trial 62, a multiple fixed-dose study conducted on over 700 patients, found no relationship between higher doses of fluoxetine and greater therapeutic benefit. When all trials are combined, mean weighted improvement on the HRSD was 8.3 points for fluoxetine and 7.3 points for placebo. In other words, placebo duplicated 89% of the antidepressant response of fluoxetine.[1]

Eli Lilly's application to the FDA for the approval of fluoxetine in children with major depression included results from two acute efficacy trials. In one trial ($N = 96$), fluoxetine was not significantly more effective than placebo on the pre-specified primary outcome ($p = 0.34$). However, a post-hoc endpoint (>30% reduction in clinician-rated depressive symptoms) achieved significance in favor of fluoxetine ($p = 0.04$). The FDA reviewer noted that the difference between fluoxetine and placebo was not significant if the cutoff point was moved up to 40% or 50% (Center for Drug Evaluation and Research, 2001), and stated that the clinical significance of the 30% post-hoc endpoint "should be a clinical judgment" (p. 36). The reviewer also noted that, in Emslie et al.'s (1997) publication of this trial, the post-hoc endpoint was presented as the primary result. The second trial ($N = 219$) employed an unusual

Table 13.2 Mean improvement (weighted for sample size) for fluoxetine and placebo in trials submitted to the FDA and published versions of the FDA trials.

	Fluoxetine		*Placebo*		
Trial Number	*HRSD Change*	*N*	*HRSD Change*	*N*	*Advantage*
Trials Submitted to the FDA					
19*	12.5	22	5.5	24	7
25	7.2	18	8.8	24	−1.6
27*	11	181	8.4	163	2.6
62 (mild)	5.9	299	5.8	56	0.1
62 (moderate/severe)*	8.8	297	5.7	48	3.1
Published Versions of Trials Submitted to the FDA					
19					
Fabre & Crismon (1985)*	13.3	16	6.5	22	6.8
25					
Rickels, Amsterdam, & Avallone (1986)*	14.6	9	9	12	5.6
27					
Stark and Hardison (1985)*	11	185	8.2	169	2.8
Cohn & Wilcox (1985)*	14.3	54	4.1	57	10.2
Feighner, Boyer, Merideth, & Hendrickson (1989)*	7.9	52	5.8	48	2.1
Byerley, Reimherr, Wood, & Grosser (1988)*	14.4	20	7.6	16	6.8
62 (mild)					
Dunlop, Domseif, Wernicke, & Potvin (1990)	5.9	299	5.8	56	0.1
Fabre & Putman (1987)	N/A	17	N/A	3	N/A
62 (moderate/severe)					
Fabre & Putman (1987)*	14.2	25	−1	2	15.2

* Fluoxetine HRSD change score superior to placebo, $p < 0.05$. Advantage = fluoxetine–placebo difference on HRSD change scores. Only the first author is listed for each publication.

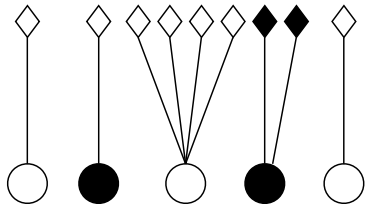
design feature: a drug run-in phase (Leo, 2006). Children assigned to fluoxetine were given 10 mg during the first week, and those who did not respond, or who had negative responses, could be dropped from the trial. The dose was increased to 20 mg at week two, and the authors only reported data from children who had at least 1 week of treatment at this higher dose. As with the first trial, fluoxetine failed to demonstrate a significant advantage over placebo on the pre-specified primary outcome ($p = 0.09$). The FDA reviewer described this study as showing "no evidence of treatment effect." The reviewer concluded, "Overall speaking, the sponsor did not win on these two pediatric depression studies based on the protocol specified endpoint. The evidence for efficacy based on the pre-specified endpoint is not convincing" (Center for Drug Evaluation and Research, 2001, p. 36).

The FDA approved fluoxetine in the treatment of adult depression in 1987. Fluoxetine received FDA approval for depressed children in 2003. Contrary to popular belief, FDA approval only indicates that a rather minimal efficacy standard has been met. Specifically, the FDA guidelines require evidence from two "adequate and well-controlled" trials that medication produces greater improvement than placebo to a statistically (not necessarily clinically) important extent (Spielmans & Kirsch, 2014). There is no limit to the number of trials that can be conducted. Negative trials are ignored. When a significant drug effect is not obtained, as in trials of fluoxetine for depressed children, investigators are sometimes allowed to switch primary outcomes on a post-hoc basis. The clinical significance of symptom improvement is not explicitly considered. Manufacturers are not required to demonstrate efficacy on self-reported symptoms, quality of life, or relevant functional outcomes. Indeed, the FDA has approved antidepressants that demonstrated no advantage over placebo on such measures (Spielmans & Gerwig, 2014). In some cases, results of positive trials for similar drugs are used as evidence of efficacy for the drug under review. Not surprisingly, the FDA has been criticized for setting an unacceptably low bar for drug approval (Spielmans & Kirsch, 2014).

Publications based on the fluoxetine trials

Results of the five fluoxetine trials submitted to the FDA indicate that fluoxetine has limited efficacy in the treatment of depression. However, a different story emerged in the published articles based on these data. Nine scientific papers were published in peer-reviewed scientific journals based on data from the FDA trials. As shown in Figure 13.1, these papers reveal a clear pattern of publication bias. The three trials yielding a statistically significant advantage of fluoxetine over placebo produced six publications. The two non-significant trials yielded three publications, one of which reported fluoxetine to be significantly more effective than placebo. Seven of the nine published articles depicted fluoxetine as significantly more effective than placebo in reducing continuous HRSD scores. An eighth study found a significant advantage of fluoxetine in HRSD response rates.

Results from the published versions of the fluoxetine trials submitted to the FDA are reported in Table 13.2. Trial 25 yielded a non-significant, 1.6-point HRSD advantage of placebo over fluoxetine. However, in the published version of this study (Rickels et al., 1986), dropouts were excluded from the analyses, which reduced the sample size by 50% and produced a statistically significant advantage for fluoxetine of 5.6 points on the HRSD. Data from Trial 27, conducted at six study sites, were published in four separate articles. Although imipramine significantly outperformed fluoxetine in Trial 27, each published article reported a non-significant difference in efficacy between fluoxetine and imipramine. Three articles presented data separately from individual Trial 27 study sites. For example, Byerley et al. (1988) reported results from a study site where fluoxetine was three times more effective, relative to placebo, than the combined multi-site results. A second paper from a Trial 27 study



○	Trial submitted to FDA, fluoxetine superior to placebo ($p < 0.05$)
●	Trial submitted to FDA, fluoxetine not superior to placebo
◇	Published trial, fluoxetine superior to placebo ($p < 0.05$)
◆	Published trial, fluoxetine not superior to placebo

Figure 13.1 Selective and multiple publication of fluoxetine trials submitted to the FDA.

site reported a large 10-point HRSD advantage of fluoxetine over placebo (Cohn & Wilcox, 1985). Fabre and Putman (1987) reported data from a single study site of Trial 62 yielding an extraordinary 15.2-point HRSD change score advantage of fluoxetine over placebo. As shown in Table 13.2, the actual advantage of fluoxetine in this trial, averaged across all study sites, was 3.1 points.

The published fluoxetine trials include a host of problematic reporting features that are common in industry-sponsored trials (Ioannidis, 2008; Safer, 2002; Sismondo, 2007; Spielmans & Kirsch, 2014; Spielmans & Parry, 2010; Turner, Matthews, Linardatos, Tell, & Rosenthal, 2008). These are summarized in Table 13.3. Positive trials were selectively published, often multiple times, and negative results were sometimes spun as positive. Seven of nine studies failed to disclose Eli Lilly's sponsorship of the trial. Self-report measures of depressive symptoms were not reported. Effect size estimates (e.g., Cohen's *d*) were not reported. The integrity of the double-blind was not assessed. In six studies, patients were dichotomized as "responders" or "non-responders" based on whether they evidenced ≥ 50% change in HRSD scores. Two publications (Dunlop et al., 1990; Rickels et al., 1986) obtained a significant advantage of fluoxetine over placebo in responder frequency that was not evident when continuous HRSD scores were analyzed. The most striking result was obtained by Dunlop et al. (1990). This author found a non-significant difference between fluoxetine and placebo of only 0.07 points on the HRSD. However, dichotomizing HRSD change scores produced a significantly higher ($p < 0.05$) proportion of responders on fluoxetine (54%) than placebo (36%). This result exemplifies the "response rate illusion" (Kirsch & Moncrieff, 2007), in which small differences in improvement scores can produce large differences in response rates.

Lessons learned from the fluoxetine trials

A medication that produces an average of one point more improvement on the HRSD than placebo is not a "wonder drug" (*New York Magazine*, Schumer, 1989).

Table 13.3 Problematic reporting features in antidepressant trials conducted by pharmaceutical companies.

Feature	*Description*	*Effect*
Selective publication of trials	Results from trials showing a significant drug effect are almost always published. Results from negative trials are rarely published.	Drug efficacy is inflated.
Spinning negative trials as positive	When results from negative trials are published, they are spun to create the appearance of a significant drug effect.	Drug efficacy is inflated.
Multiple publication of trials	Results from the same trial are published in multiple articles. These may take the form of stand-alone trial reports or pooled analyses in which data from multiple trials are combined. Authors typically fail to disclose that their study reports previously published data.	Drug efficacy is inflated.
Suppression of unfavorable outcomes	Authors exclude data from patients who dropped out due to lack of efficacy or adverse effects, do not report results of non-significant efficacy comparisons, do not report complete data on adverse effects, and hide suicidal events by labeling them as something else.	Drug efficacy and safety are inflated.
Reporting post-hoc outcomes as primary outcomes	Authors cherry-pick outcomes that yield the most favorable results and present them as pre-specified primary outcomes.	Drug efficacy is inflated.
Failure to report clinical significance	Statistics for characterizing the clinical significance of trial results are not reported.	Drug efficacy is inflated.
Dichotomizing continuous outcomes	Participants are dichotomized as "responders" or "non-responders" based on their scores on continuous measures of depressive symptoms. Small, non-significant differences in continuous scores can produce relatively large and statistically significant differences in the proportion of "responders" and "non-responders" to drug and placebo.	Drug efficacy is inflated.

(*Continued*)

Table 13.3 (Continued)

Feature	*Description*	*Effect*
Failure to disclose industry involvement	Authors fail to disclose industry sponsorship of the trial, and/or their financial conflicts of interest with the drug maker.	Published trials have the false appearance of independence from industry influence, lending them undue credibility.
Ghost authorship	Pharmaceutical companies secretly author articles under the byline of academic researchers. Undisclosed authors (i.e., "ghostwriters") hired by drug makers draft scientific manuscripts, and prominent "key opinion leaders" are subsequently added as authors.	Prominent researchers bestow the false appearance of credibility to trials they did not conduct. Researchers are corrupted by financial incentives to endorse the company's product.
Industry ownership of data	Pharmaceutical companies own the data from the trials they sponsor. They have access to accumulating data during the trial, can stop the trial at any time and for any reason, and need to approve the trial manuscript prior to submission for publication. Investigators are constrained from conducting independent data analyses and publishing data regardless of the trial outcome.	Drug efficacy and safety are inflated. Researcher autonomy is constrained.

A medication that is significantly less effective than an antidepressant approved by the FDA when Dwight Eisenhower was president of the United States is not a "medical breakthrough" (*Newsweek*; Cowley et al., 1990). The clinical trials data submitted to the FDA suggest that the efficacy of fluoxetine is small, unreliable, clinically insignificant, and inflated by biased design and reporting practices. Breathless proclamations of fluoxetine's therapeutic benefits in popular media bear little resemblance to the actual clinical trials data. However, these data were hidden from view until 15 years after fluoxetine was approved by the FDA (Kirsch et al., 2002).

The iconic status of Prozac is a product not of its powerful antidepressant properties but rather of Eli Lilly's enormously successful marketing campaign (Healy, 2004). At the center of this campaign were clinical trial results published in prestigious psychiatry journals. They told the story of a new, safe, and highly effective "antidepressant." This story was repeated in the media, advertised directly to consumers, and conveyed to physicians by sales representatives. Published clinical trials were the Trojan horses (Healy, 2012) in which Eli Lilly inserted marketing appearing as science inside the gates of the peer-reviewed scientific community. Medical journals became, in the words of former *British Medical Journal* editor Richard Smith (2005), "an extension of the marketing arm of pharmaceutical companies" (p. 364). Misleading publications featuring design and reporting flaws would eventually become the basis for clinical practice guidelines recommending antidepressants such as fluoxetine as a first-line treatment for depression (e.g., APA, 2010). These guidelines were authored by psychiatrists who had extensive and ongoing financial relationships with the pharmaceutical companies whose products they reviewed. To illustrate, first author of the APA (2010) guidelines Alan Gelenberg disclosed the following conflicts (among others not listed here): (a) consulting for Eli Lilly, Pfizer, AstraZeneca, Wyeth, Novartis, Forest, GlaxoSmithKline, ZARS Pharma, Jazz Pharmaceuticals, Lundbeck, Takeda, and Dey Pharma; (b) serving on speakers bureaus for Pfizer, GlaxoSmithKline, and Wyeth; and (c) receiving grant funding from Eli Lilly, Pfizer, and GlaxoSmithKline.

Several conclusions can be drawn from Eli Lilly's creation of fluoxetine as a blockbuster antidepressant. First, the acute efficacy trials were designed to suppress the placebo effect and inflate the apparent efficacy of fluoxetine. Although this is standard practice in industry-sponsored trials, it severely limits their validity and generalizability. Second, fluoxetine is not particularly efficacious. It was significantly less effective than imipramine and failed to significantly outperform placebo in numerous trials designed to show an advantage of fluoxetine. This is not to say that patients taking fluoxetine did not experience symptom reduction – they did. However, the magnitude of this reduction was not significantly larger with fluoxetine than placebo to a degree that is clinically meaningful (Kirsch et al., 2008). Third, the FDA's approval of fluoxetine exemplifies Spielmans and Kirsch's (2014) contention that "The FDA's framework for evaluating clinical trials allows drugs with minimal efficacy in terms of symptomatic improvement – and no benefit in terms of quality of life or social functioning – to enter the marketplace of approved treatments" (p. 760). This observation is particularly applicable to the FDA's approval

of fluoxetine in depressed children based on two clinical trials that failed to demonstrate a significant advantage of fluoxetine over placebo on the pre-specified primary outcome measure. Fourth, Eli Lilly's pattern of selective publication, spin, and suppression of negative outcomes indicates that published versions of the trials submitted to the FDA are marketing masquerading as science. Evidence-based medicine founded on results of industry-sponsored trials may be more accurately construed as "marketing-based medicine" (Spielmans & Parry, 2010). The published literature overestimates the efficacy of fluoxetine (Turner et al., 2008) and compromises the ability of patients, treatment providers, scientists, policy-makers, and other stakeholders to make accurate assessments about its clinical effects and informed decisions about its use. Fifth, fluoxetine's reputation as an extraordinarily effective treatment for depression is based largely on a scientific myth. By the time this myth began to unravel (Kirsch et al., 2008; Turner et al., 2008), one in 10 Americans aged 12 and older were taking antidepressant medications (Pratt et al., 2011). Fluoxetine's patent exclusivity had expired, and, having made billions of dollars from its "wonder drug," Eli Lilly had moved on to the creation of other controversial blockbusters – such as the antipsychotic olanzapine (Zyprexa), which was later rebranded as a "broad spectrum psychotropic" (Spielmans, 2009).

The Efficacy of "Antidepressant" Medications

Our critical analysis of fluoxetine is broadly applicable to newer-generation antidepressants as a group. In terms of efficacy in depressed adults, meta-analytic studies have consistently reported small drug effects. For example, an analysis of clinical trials data submitted to the FDA for six newer-generation antidepressants yielded a statistically significant but small average drug effect of 1.8 points on the HRSD (Kirsch et al., 2002). This difference falls short of clinical importance according to the National Center for Health and Care Excellence (2010). Although there is no gold standard for defining a clinically important antidepressant effect (Spielmans & Kirsch, 2014), changes of three points or less on the HRSD correspond to ratings of "no change" on clinician-rated global symptom severity (Leucht, Fennema, Engel, Kaspers-Janssen, Lepping, & Szegedi, 2013). A follow-up meta-analysis (Kirsch et al., 2008) found that antidepressant efficacy increased significantly as a function of baseline severity. Clinically significant antidepressant efficacy was only evident in studies of patients who had, on average, baseline depressive symptoms in the upper end of the "very severe" range on the HRSD. Different antidepressants had statistically equivalent efficacy, and placebo duplicated 82% of the improvement observed in the drug groups.

One limitation to the studies analyzed by Kirsch and colleagues (2002, 2008) is that all but one study was conducted with patients whose average baseline depression score was severe. To address this limitation, Fournier et al. (2010) conducted a meta-analysis of patient-level data in clinical trials that included a broader range of baseline symptom severity. A clinically significant drug effect was evident only

among very severely depressed patients with HRSD scores ≥25. The authors concluded, "True drug effects (an advantage of ADM over placebo) were nonexistent to negligible among depressed patients with mild, moderate, and even severe baseline symptoms" (p. 51). Fournier et al. (2010) also noted that the apparent efficacy of antidepressants is largely based on studies of very severely depressed patients. Although such studies create the perception that antidepressants are efficacious, they are unlikely to provide clinically meaningful benefit over placebo for the vast majority of depressed individuals who take them.

Publication bias severely compromises the validity of the published antidepressant literature and confounds attempts to draw meaningful conclusions from it via systematic review and meta-analysis. This reality was laid bare in a seminal article published in the *New England Journal of Medicine* by Turner et al. (2008). The authors obtained results of 74 trials of 12 antidepressant medications submitted to the FDA. Corresponding publications based on these trials were located, and their results were compared with the submitted trials. Approximately half (51%) of the FDA trials yielded a significant drug effect; of these, 97% were published. Among the trials submitted to the FDA with negative or questionable outcomes, 61% were not published, and 30.6% were published but portrayed as positive. Separate meta-analyses of the FDA and published data showed that the efficacy of the 12 antidepressants, collectively, was inflated by 32% in the published literature. Because Turner and colleagues (2008) considered only the first publication of a given FDA trial, their analysis excluded the many duplicate and pooled publications of antidepressant trials identified by previous investigators (Melander, Ahlqvist-Rastad, Meijer, & Beermann, 2003). Accordingly, the extent to which the published literature inflates the actual efficacy of antidepressants likely exceeds the 32% figure reported by Turner et al. (2008).

The efficacy of antidepressants in children is particularly tenuous (Leo, 2006). A recent meta-analysis (Spielmans & Gerwig, 2014) found no significant differences between newer-generation antidepressants and placebo on self-reported depressive symptoms ($p = 0.36$) or measures of quality of life, global mental health, self-esteem, and autonomy ($p = 0.13$). In contrast, meta-analytic reviews typically report a small but statistically significant advantage of antidepressants over placebo on clinician-rated symptom measures (e.g., effect size of $d = 0.20$; in Bridge et al., 2007). Whittington, Kendall, Fonagy, Cottrell, Cotgrove, and Boddington (2004) concluded that published and unpublished data together show an unfavorable risk–benefit profile for paroxetine, sertraline, citalopram, and venlafaxine. Only fluoxetine was deemed to have a positive risk–benefit profile; however, fluoxetine's apparently superior efficacy among antidepressants in youth is not due to a greater drug response but lower rates of placebo response than those observed for other drugs (Bridge, Birmaher, Iyengar, Barbe, & Brent, 2009).

An influential meta-analysis published in *JAMA* (Bridge et al., 2007) reported a small but statistically significant advantage of antidepressants over placebo in the reduction of clinician-rated depressive symptoms in children. This meta-analysis included trials with serious methodological flaws in which negative outcomes were suppressed (Leo, 2006). One of these trials involves the paroxetine study 329 (Keller

et al., 2001), which has been the subject of calls for retraction for data manipulation, ghostwriting, misleading reporting, and undisclosed conflicts of interest (1 Boring Old Man, 2011; Healy, 2006; McHenry & Jureidini, 2008); a book by an investigative journalist (Bass, 2008); and a lawsuit for consumer fraud against GlaxoSmithKline filed by former New York attorney general Eliot Spitzer. In 2012, GlaxoSmithKline agreed to a US$3 billion settlement with the United States Department of Justice for (among other alleged crimes) off-label promotion and failure to disclose safety data for Paxil. The criminal plea agreement (United States Department of Justice, 2012) alleged, in reference to study 329, "… GSK participated in preparing, publishing and distributing a misleading medical journal article that misreported that a clinical trial of Paxil demonstrated efficacy in the treatment of depression in patients under age 18, when the study failed to demonstrate efficacy." GlaxoSmithKline's resolution was the largest health care fraud settlement in US history and the largest fine ever paid by a pharmaceutical company.

Conclusion

The title of this chapter asks the question "Are antidepressants overrated?" If "overrated" is defined as a discrepancy between their reputation and the available scientific evidence, the answer is an unequivocal "yes." The popularity of antidepressants in clinical practice and popular culture is belied by an uninspiring and misleading clinical trials literature. The industry-sponsored trials that dominate the scientific literature are designed to minimize the placebo response and maximize the drug response. Biased trial design and reporting practices further manufacturers' commercial interests but compromise scientific integrity. Despite stacking the deck in favor of the antidepressant, approximately half of all industry-sponsored trials fail to produce a statistically significant drug effect. On average, trial results reveal a small and likely clinically meaningless advantage of antidepressants over placebo for all but the most severely depressed patients. This result is similar across different antidepressants and is independent of their dose. Published versions of industry-sponsored trials systematically exaggerate antidepressant efficacy and minimize their adverse effects. Until recently, these trials were perceived as credible by a naïve scientific community. The validity of the published literature is severely compromised by pharmaceutical company marketing masquerading as science. Meta-analytic reviews and clinical guidelines based on the published literature are similarly compromised (Whittington et al., 2004). Even reviews of the unpublished literature are threatened by hidden data manipulation (Healy, n.d.b, 2006) and the suppression of negative outcomes (Leo, 2006) in the original trials.

Modern antidepressant medications such as fluoxetine were not so much discovered as invented (Healy, 1997, 2004). Our critical analysis of fluoxetine illustrates how a minimally efficacious drug became a cultural icon through a marketing campaign based on selectively published clinical trials data. The marketing of other newer-generation antidepressants followed a similar pattern. In the case of

paroxetine, this marketing was so egregious that GlaxoSmithKline was found guilty of health care fraud by the United States Department of Justice.

Although antidepressants are the primary subject of this chapter, problems associated with flawed clinical trial design and reporting practices, inconsistent clinical trial results, and exaggerated efficacy and safety claims also apply to SGAs (Leucht et al., 2009; Spielmans & Parry, 2010; see Chapters 1–5). Poor replicability across clinical studies appears to be heavily influenced by competing commercial interests. To illustrate, Heres, Davis, Maino, Jetzinger, Kissling, and Leucht (2006) reported that the overall outcome favored the sponsor's drug in 90% of head-to-head trials of SGAs. Sponsor-friendly outcomes were influenced by sources of bias including "doses and dose escalation, study entry criteria and study populations, statistics and methods, and reporting of results and wording of findings" (Heres et al., 2006, p. 185). The poor replicability of clinical trial results for both antidepressant and newer-generation antipsychotic medications is consistent with broader concerns about poor replicability of psychological research in general (Ioannidis, 2012). However, commercial interests provide a uniquely powerful incentive for biased research due to their ability to facilitate FDA approval and lucrative marketing campaigns.

Like antidepressants, SGAs owe their popularity in part to aggressive marketing based on selectively published data from clinical trials with biased methodology (Spielmans & Parry, 2010). These marketing campaigns have often earned pharmaceutical manufacturers large government fines for allegations of healthcare fraud. According to the United States Department of Justice (2007), Bristol-Myers Squibb engaged in illegal marketing tactics for aripiprazole that included paying kickbacks to physicians and promoting the drug for off-label use among children and nursing home residents suffering from dementia. These allegations resulted in a US$515 million settlement in 2007. Financial settlements were subsequently reached with the US Department of Justice for alleged illegal marketing of others SGAs, including Eli Lilly's olanzapine (Zyprexa; US$1.4 billion in 2009); Pfizer's ziprasidone (Geodon; US$2.3 billion in 2009), AstraZeneca's quetiapine (Seroquel; US$520 million in 2010), and Johnson & Johnson's risperidone (Risperdal; US$2.2 billion in 2013). At the time of this writing, the best-selling drug in America is the SGA aripiprazole (Abilify; Michaelson, 2014). Sales of aripiprazole from April 2013 through March 2014 totaled US$6.9 billion, an amount that exceeded sales of all antidepressant medications combined. This 1-year sales figure is more than 13 times greater than the financial settlement Bristol-Myers Squibb reached with the Justice Department in 2007 for illegally marketing aripiprazole.

In closing, the dominant cultural story of antidepressant medications is, in the words of eminent scholar John Ioannidis, "an evidence myth constructed from a thousand randomized trials" (2008, p. 1). Now that the myth has been exposed (e.g., Ioannidis, 2008; Kirsch, 2010; Whitaker, 2010), critical public dialogue on the safety and efficacy of antidepressants is taking place. It is our hope that this chapter will advance this critical dialogue, so the clinical management of depressed patients reflects their best interests rather than the commercial goals of pharmaceutical companies seeking to invent the next "blockbuster" antidepressant.

Endnote

1 As noted by Kirsch et al. (2002), standard deviations are not reported in most clinical trial summaries obtained from the FDA. These data are also absent from most published trials of fluoxetine described in the section titled "Publications Based on the Fluoxetine Trials." The absence of standard deviations precludes calculation of traditional effect size estimates. However, since the HRSD was used as the primary outcome measure in each trial, it is possible to combine results across studies on this measure without reference to standard deviations. The relative efficacy of fluoxetine vs. placebo can thus be described in terms of differences in HRSD raw change scores, or the percentage overlap in HRSD change scores. These indices are arguably more readily interpretable than effect size estimates based on standardized mean differences.

References

1 Boring Old Man (May 7, 2011). *Retract study 329 ….* Retrieved December 29, 2014, from http://1boringoldman.com/index.php/2011/05/07/retract-study-329/

American Psychiatric Association. (2010). *Practice guideline for the treatment of patients with major depressive disorder* (3rd edn). Arlington, VA: Author.

Bass, A. (2008). *Side effects: A prosecutor, a whistleblower, and a bestselling antidepressant on trial*. Chapel Hill, NC: Algonquin Books.

Baum, Hedlund, Aristei, & Goldman, P.C. (n.d.a). *A quest for justice*. Retrieved December 29, 2014, from http://www.baumhedlundlaw.com/06.pdf

Baum, Hedlund, Aristei, & Goldman, P.C. (n.d.b). *A quest for justice*. Retrieved December 29, 2014, from http://www.baumhedlundlaw.com/03.pdf

Bridge, J. A., Birmaher, B., Iyengar, S., Barbe, R. P., & Brent, D. A. (2009). Placebo response in randomized controlled trials of antidepressants for pediatric major depressive disorder. *American Journal of Psychiatry*, *166*, 42–49.

Bridge, J. A., Iyengar, S., Salary, C. B., Barbe, R. P., Birmaher, B., Pincus, H. A., et al. (2007). Clinical response and risk for reported suicidal ideation and suicide attempts in pediatric antidepressant treatment: A meta-analysis of randomized controlled trials. *Journal of the American Medical Association*, *297*, 1683–1696.

Byerley, W. F., Reimherr, F. W., Wood, D. R., & Grosser, B. I. (1988). Fluoxetine, a selective serotonin uptake inhibitor, for the treatment of outpatients with major depression. *Journal of Clinical Psychopharmacology*, *8*, 112–115.

Center for Drug Evaluation and Research. (2001). *Statistical review of fluoxetine for pediatric OCD and depression*. Retrieved December 29, 2014, from http://www.accessdata.fda.gov/drugsatfda_docs/nda/2003/18936S064_Fluoxetine%20Pulvules_statr.pdf

Cohn, J. B., & Wilcox, C. (1985). A comparison of fluoxetine, imipramine, and placebo inpatients with major depressive disorder. *Journal of Clinical Psychiatry*, *46*, 26–31.

Cowley, G., Springen, K., Leonard, E. A., Robins, K., & Gordon, J. (March 26, 1990). The promise of fluoxetine. *Newsweek*, 38–41.

Dunlop, S. R., Domseif, B. E., Wernicke, J. F., & Potvin, J. H. (1990). Pattern analysis shows beneficial effect of fluoxetine treatment in mild depression. *Psychopharmacology Bulletin*, *26*, 173–180.

Emslie, G. J., Rush, A. J., Weinberg, W. A., Kowatch, R. A., Hughes, C. W., Carmody, T., et al. (1997). A double-blind, randomized, placebo-controlled trial of fluoxetine in

children and adolescents with depression. *Archives of General Psychiatry*, *54*, 1031–1037.

Even, C., Siobud-Dorocant, E., & Dardennes, R. M. (2000). Critical approach to antidepressant trials. Blindness protection is necessary, feasible and measurable. *British Journal of Psychiatry*, *177*, 47–51.

Fabre, L. F., & Crismon, L. (1985). Efficacy of fluoxetine in outpatients with major depression. *Current Therapeutic Research*, *37*, 115–123.

Fabre, L. F., & Putman, H. P. III (1987). A fixed-dose clinical trial of fluoxetine in outpatients with major depression. *Journal of Clinical Psychiatry*, *48*, 406–408.

Feighner, J. P., Boyer, W. F., Merideth, C. H., & Hendrickson, G. G. (1989). A double-blind comparison of fluoxetine, imipramine and placebo in outpatients with major depression. *International Clinical Psychopharmacology*, *4*, 127–134.

Fitzpatrick, L. (January 7, 2010). A brief history of antidepressants. *Time*. Retrieved December 29, 2014, from http://content.time.com/time/health/article/0,8599,1952143,00.html

Fournier, J. C., DeRubeis, R. J., Hollon, S. D., Dimidjian, S., Amsterdam, J. D., Shelton, R. C., et al. (2010). Antidepressant drug effects and depression severity: A patient-level meta-analysis. *Journal of the American Medical Association*, *303*, 47–53.

Greenberg, R. P., Bornstein, R. F., Fisher, S., Zborowski, M. J., & Greenberg, M. D. (1994). A meta-analysis of fluoxetine outcome in the treatment of depression. *The Journal of Nervous and Mental Disease*, *182*, 547–551.

Hamilton, M. A. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, *23*, 56–61.

Healy, D. (n.d.a). *A quick guide to the suicide data on Prozac from October 86*. Retrieved December 29, 2014, from http://www.healyprozac.com/Trials/CriticalDocs/suicattempt031086.htm

Healy, D. (n.d.b). *A quick guide to the suicide data on Prozac from October 86*. Retrieved December 29, 2014, from http://www.healyprozac.com/Trials/CriticalDocs/cbouchy131190.htm

Healy, D. (2004). *Let them eat Prozac: The unhealthy relationship between the pharmaceutical industry and depression*. New York, NY: New York University Press.

Healy, D. (2006). Manufacturing consensus. *Culture, Medicine, and Psychiatry*, *30*, 135–156.

Healy, D. (2012). *Pharmageddon*. Los Angeles, CA: University of California Press.

Heres, S., Davis, J., Maino, K., Jetzinger, E., Kissling, W., & Leucht, S. (2006). Why olanzapine beats risperidone, risperidone beats quetiapine, and quetiapine beats olanzapine: An exploratory analysis of head-to-head comparison studies of second-generation antipsychotics. *American Journal of Psychiatry*, *163*, 185–194.

Ioannidis, J. P. A. (2008). Effectiveness of antidepressants: An evidence myth constructed from a thousand randomized trials? *Philosophy, Ethics, and Humanities in Medicine*, *3*, 1–9.

Ioannidis, J. P. A. (2012). Why science is not necessarily self-correcting. *Perspectives on Psychological Science*, *7*(6), 645–654.

Keller, M. B., Ryan, N. D., Strober, M., Klein, R. G., Kutcher, S. P., Birmaher, B., et al. (2001). Efficacy of paroxetine in the treatment of adolescent major depression: A randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*, 762–772.

Kirsch, I. (2010). *The emperor's new drugs: Exploding the antidepressant myth*. New York, NY: Basic Books.

Kirsch, I., Deacon, B. J., Huedo-Medina, T. B., Scoboria, A., Moore, T. J., & Johnson, B. T. (2008). Initial severity and antidepressant benefits: A meta-analysis of data submitted to the FDA. *PLoS Medicine*, *5*, 0260–0268.

Kirsch, I., & Moncrieff, J. (2007). Clinical trials and the response rate illusion. *Contemporary Clinical Trials*, *28*, 348–351.

Kirsch, I., Moore, T. J., Scoboria, A., & Nicholls, S. S. (2002). The emperor's new drugs: An analysis of antidepressant medication data submitted to the U.S. Food and Drug Administration. *Prevention & Treatment*, *5*, article 23.

Kramer, P. (1993). *Listening to Fluoxetine: The landmark book about antidepressants and the remaking of the self*. New York, NY: Penguin Books.

Lacasse, J., & Leo, J. (2005). Serotonin and depression: A disconnect between the advertisements and the scientific literature. *PLoS Medicine*, *2*, 1211–1216.

Lenzer, J. (2005). FDA to review "missing" drug company documents. *British Medical Journal*, *330*, 7.

Leo, J. (2006). The SSRI trials in children: Disturbing implications for academic medicine. *Ethical Human Psychiatry and Psychology*, *8*, 29–41.

Leucht, S., Fennema, H., Engel, R., Kaspers-Janssen, M., Lepping, P., & Szegedi, A. (2013). What does the HAMD mean? *Journal of Affective Disorders*, *148*, 243–248.

Leucht, S., Kissling, W., & Davis, J. M. (2009). Second-generation antipsychotics for schizophrenia: Can we resolve the conflict? *Psychological Medicine*, *39*, 1591–1602.

Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in psychology research: How often do they really occur? *Perspectives on Psychological Science*, *7*(6), 537–542.

McHenry, L. B., & Jureidini, J. N. (2008). Industry-sponsored ghostwriting in clinical trial reporting: A case study. *Accountability in Research*, *15*, 152–167.

Melander, H., Ahlqvist-Rastad, J., Meijer, G., & Beermann, B. (2003). Evidence b(i)ased medicine – selective reporting from studies sponsored by pharmaceutical industry: Review of studies in new drug applications. *British Medical Journal*, *326*, 1171–1173.

Michaelson, J. (November 9, 2014). Mother's little anti-psychotic is worth US$6.9 billion a year. *The Daily Beast*. Retrieved December 29, 2014, from http://www.thedailybeast.com/articles/2014/11/09/mother-s-little-anti-psychotic-is-worth-6-9-billion-a-year.html

Moncrieff, J., Wessely, S., & Hardy, R (2004). Active placebos versus antidepressants for depression. *Cochrane Database of Systematic Reviews*, *1*, 1–27.

National Institute for Health and Care Excellence. (2009). *Depression in adults: The treatment and management of depression in adults*. Clinical practice guideline No. 90. London: National Institute for Health and Care Excellence.

National Center for Health & Clinical Excellence. (2010). *Depression: The treatment and management of depression in adults (updated edition)*. London: The British Psychological Society and The Royal College of Psychiatrists.

National Center for Health Statistics. (2010). *Health, United States, 2010: With special feature on death and dying*. Table 95. Hyattsville, MD.

Pescosolido, B. A., Martin, J. K., Long, J. S., Medina, T. R., Phelan, J. C., & Link, B. G. (2010). A disease like any other? A decade of change in public reactions to schizophrenia, depression, and alcohol dependence. *American Journal of Psychiatry*, *167*, 1321–1330.

Pratt, L. A., Brody, D. J., & Gu, Q. (2011). *Antidepressant use in persons aged 12 and over: United States, 2005–2008*. NCHS Data Brief, No 76. Hyattsville, MD: National Center for Health Statistics.

Rickels, K., Amsterdam, J. D., & Avallone, M. F. (1986). Fluoxetine in major depression: A controlled study. *Current Therapeutic Research*, *39*, 559–563.

Sackett, D. L. (2005). *Evidence-based medicine*. New York, NY: Wiley.

Safer, D. J. (2002). Design and reporting modifications in industry-sponsored comparative psychopharmacology trials. *The Journal of Nervous and Mental Disease*, *190*, 583–592.

Schumer, F. (December 18, 1989). Bye-bye blues: A new wonder drug for depression. *New York Magazine*, 46–53. Retrieved December 29, 2014, from https://books.google.com.au/books?id=NugCAAAAMBAJ&pg=PA46&lpg=PA46&dq=%22Bye-bye+blues:+A+new+wonder+drug+for+depression%22&source=bl&ots=DFvPI1UVPr&sig=c7pBV7lp8JcrpCQ32ALZ154Yjmo&hl=en&sa=X&ei=-9egVKOHMcW1mwXg6oLoCQ&ved=0CCAQ6AEwAQ#v=onepage&q=%22Bye-bye%20blues%3A%20A%20new%20wonder%20drug%20for%20depression%22&f=false

Seife, C. (2015). Research misconduct identified by the US Food and Drug Administration: Out of sight, out of mind, out of the peer-reviewed literature. *JAMA Internal Medicine*, *175*, 567–577.

Sismondo, S. (2007). Ghost management: How much of the medical literature is shaped behind the scenes by the pharmaceutical industry? *PLoS Medicine*, *4*, e286.

Smith, R. (2005). Medical journals are an extension of the marketing arm of pharmaceutical companies. *PLoS Medicine*, *2*, e138.

Spielmans, G. I. (2009). The promotion of olanzapine in primary care: An examination of internal industry documents. *Social Science & Medicine*, *69*, 14–20.

Spielmans, G. I., & Gerwig, K. (2014). The efficacy of antidepressants on overall well-being and self-reported depression symptom severity in youth: A meta-analysis. *Psychotherapy and Psychosomatics*, *83*, 158–164.

Spielmans, G. I., & Kirsch, I. (2014). Drug approval and drug effectiveness. *Annual Review of Clinical Psychology*, *10*, 741–766.

Spielmans, G. I., & Parry, P. I. (2010). From evidence-based medicine to marketing-based medicine: Evidence from internal industry documents. *Bioethical Inquiry*, *7*, 13–29.

Stark, P., & Hardison, C. D. (1985). A review of multicenter controlled studies of fluoxetine vs. imipramine and placebo in outpatients with major depressive disorder. *Journal of Clinical Psychiatry*, *46*, 115–123.

Sugarman, M. A., Loree, A. M., Baltes, B. B., Grekin, E. R., & Kirsch, I. (2014). The efficacy of paroxetine and placebo in treating anxiety and depression: A meta-analysis of change on the Hamilton rating scales. *PLoS ONE*, *9*(8), e106337.

Turner, E. H. (2013). Publication bias, with a focus on psychiatry: Causes and solutions. *CNS Drugs*, *27*, 457–468.

Turner, E. H., Matthews, A. M., Linardatos, E., Tell, R. A., & Rosenthal, R. (2008). Selective publication of antidepressant trials and its influence on apparent efficacy. *New England Journal of Medicine*, *358*, 252–260.

United States Department of Justice. (September 28, 2007). *Bristol-Myers Squibb to pay more than $515 million to resolve allegations of illegal drug marketing and pricing*. Retrieved December 29, 2014, from http://www.justice.gov/archive/opa/pr/2007/September/07_civ_782.html

United States Department of Justice. (July 2, 2012). *GlaxoSmithKline to plead guilty and pay $3 billion to resolve fraud allegations and failure to report safety data*. Retrieved December 29, 2014, from http://www.justice.gov/opa/pr/glaxosmithkline-plead-guilty-and-pay-3-billion-resolve-fraud-allegations-and-failure-report

Whitaker, R. (2010). *Anatomy of an epidemic: Magic bullets, psychiatric drugs, and the astonishing rise of mental illness in America*. New York, NY: Crown.

Whittington, C. J., Kendall, T., Fonagy, P., Cottrell, D., Cotgrove, A., & Boddington, E. (2004). Selective serotonin reuptake inhibitors in childhood depression: Systematic review of published versus unpublished data. *The Lancet*, *363*, 1341–1345.