The Placebo Effect in the Treatment of Depression

Irving Kirsch

Harvard Medical School, Beth Israel Deaconess Medical Center, Boston, MA, USA

Summary

Antidepressants are supposed to work by fixing a chemical imbalance, specifically, a lack of serotonin in the brain. However, analyses of the published and the unpublished data that were hidden by the drug companies reveal that most (if not all) of the benefits are due to the placebo effect. Some antidepressants increase serotonin levels, some decrease serotonin, and some have no effect at all on serotonin. Nevertheless, they all show the same therapeutic benefit. Instead of curing depression, popular antidepressants may induce a biological vulnerability making people more likely to become depressed in the future. Other treatments (e.g., psychotherapy and physical exercise) produce the same short-term benefits as antidepressants, show better long-term effectiveness, and do so without the side effects and health risks of the drugs.

© Irving Kirsch, PhD/S. Karger GmbH, Freiburg

Zusammenfassung

Die vorherrschende Annahme zur Wirkung von Antidepressiva ist, dass diese durch die Behebung eines chemischen Ungleichgewichts, speziell eines Serotoninmangels im Gehirn, zustande kommt. Auswertungen der publizierten und nicht publizierten, d.h. von Pharmaunternehmen zurückgehaltenen Daten zeigen jedoch, dass der meiste (wenn nicht gar komplette) Nutzen durch den Placeboeffekt bewirkt wird. Einige Antidepressiva steigern die Serotoninspiegel, manche senken sie, und andere haben überhaupt keine Auswirkung auf Serotonin. Dennoch zeigen sie alle den gleichen therapeutischen Nutzen. Anstatt Depression zu behandeln, könnten gängige Antidepressiva eine biologische Vulnerabilität verursachen, die das Risiko einer zukünftigen Depression erhöht. Andere Behandlungsarten (z.B. Psychotherapie und körperliche Betätigung) erzeugen den gleichen Kurzzeitnutzen wie Antidepressiva, zeigen bessere Langzeitwirkung und tun dies ohne die Nebenwirkungen und Gesundheitsrisiken der Medikamente.

Introduction

On February 26, 2008, an article about antidepressants that my colleagues and I wrote was published in the journal PLoS Medicine [Kirsch et al., 2008]. That morning, I awoke to find that our paper was the front-page story in all of the leading national newspapers in the United Kingdom. A few months later, Random House invited me to expand the article into a book, entitled *The Emperor's New Drugs: Exploding the Antidepressant Myth* [Kirsch, 2009], which has since been translated into French, Italian, Japanese, Pol-

ish, and Turkish. Two years later, the book, and the research reported in it, was the topic of a 5-page cover story in the influential American news magazine, *Newsweek*. Two years after that, it was the focus of a 15-min segment on *60 Minutes*, America's top-rated television news program. Somehow, I had been transformed from a mild-mannered university professor into a media superhero – or super villain, depending on whom you asked. What had my colleagues and I done do warrant this transformation?

To answer that question, we have to go back to 1998, when a former graduate student, Guy Sapirstein, and I published a meta-analysis on antidepressants in an online journal of the American Psychological Association [Kirsch and Sapirstein, 1998]. Meta-

Parts of this article have been published elsewhere.

Irving Kirsch, PhD
Harvard Medical School
Beth Israel Deaconess Medical Center
330 Brookline Avenue, Boston, MA 02215, USA
irvkirsch@gmail.com

analysis is a statistical tool for pooling the results of large numbers of studies on the same topic and analyzing them together. When they were new, meta-analyses were somewhat controversial, but now they are published in all of the major medical journals, where they are widely considered to be the best and most reliable way of making sense of the data from studies with different and sometimes conflicting results.

When Sapirstein and I began our analysis of the antidepressant clinical trial data, we were not particularly interested in antidepressants. Instead, we were interested in understanding the placebo effect. I have been fascinated by the placebo effect for my entire academic career. How is it, I wondered, that the belief that one has taken a medication can produce some of the effects of that medication?

It seemed to Sapirstein and me that depression was a good place to look for placebo effects. After all, one of the prime characteristics of depression is the sense of hopelessness that depressed people feel. If you ask depressed people to tell you what the worst thing in their life is, many will tell you that it is their depression. The British psychologist John Teasdale called this being depressed about depression. If that is the case, then the mere promise of an effective treatment should help to alleviate depression, by replacing hopelessness with hopefulness – the hope that one will recover after all. It was with this in mind that we set out to measure the placebo effect in depression.

Sapirstein and I searched the literature for studies in which depressed patients had been randomized to receive an inert placebo or no treatment at all. The studies we found also included data on the response to antidepressants, because that was the only place one finds data on the response to placebo among depressed patients. I was not particularly interested in the drug effect. I assumed that antidepressants were effective. As a psychotherapist, I sometimes referred my severely depressed clients for prescriptions of antidepressant drugs. Sometimes the condition of my clients improved when they began taking antidepressants; sometimes it did not. When it did, I assumed it was the effect of the drug that was making them better. Given my long-standing interest in the placebo effect, I should have known better, but back then I did not.

Analyzing the data we had found, Sapirstein and I were not surprised to find a substantial placebo effect on depression. What surprised us was how small the drug effect was. 75% of the improvement in the drug group also occurred when people were given dummy pills with no active ingredient in them. Needless to say, our meta-analysis proved to be very controversial. Its publication led to heated exchanges. The response from critics was that these data could not be accurate. Perhaps our search had led us to analyze an unrepresentative subset of clinical trials. Antidepressants had been evaluated in many trials, the critics said, and their effectiveness had been well established.

In an effort to response to these critics, we decided to replicate our study with a different set of clinical trials [Kirsch et al., 2002]. To do this, we used the Freedom of Information Act to request that the Food and Drug Administration (FDA) send us the data that pharmaceutical companies had sent to it in the process of obtaining approval for 6 new generation antidepressants that accounted for the bulk of antidepressant prescriptions being written at the time. There are a number of advantages to the FDA data set. Most important, the FDA requires that the pharmaceutical companies provide information on all of the clinical trials that they have sponsored. Thus, we had data on unpublished trials as well as published trials. This turned out to be very important. Almost half of the clinical trials sponsored by the drug companies have not been published. Only the drug companies and the FDA knew the results of the unpublished trials, and most of them failed to find a significant benefit of drug over placebo. A second advantage of the FDA trials in the FDA data set is that they all used the same primary measure of depression – the Hamilton depression scale (HAM-D). That made it easy to understand the clinical significance of the drug-placebo differences. Finally, the data in the FDA files were the basis upon which the medications were approved. In that sense, they have a privileged status. If there is anything wrong with those trials, the medications should not have been approved in the first place.

In the data sent to us by the FDA, only 43% of the trials showed a statistically significant benefit of drug over placebo. The remaining 57% were failed or negative trials. The results of our analysis indicated that the placebo response was 82% of the response to these antidepressants. Subsequently, my colleagues and I replicated our meta-analysis on a larger number of trials that had been submitted to the FDA [Kirsch et al., 2008]. With this expanded data set, we found once again that 82% of the drug response was duplicated by placebo. More important, in both analyses, the mean difference between drug and placebo was less than 2 points on the HAM-D. The HAM-D is a 17-item scale on which people can score 0–53 points, depending on how depressed they are. A 6-point difference can be obtained just by changes in sleep patterns, with no change in any other symptom of depression. So the 1.8 difference that we found between drug and placebo was very small indeed – small enough to be clinically insignificant. But you don't have to take my word for how small this difference is. The National Institute for Health and Clinical Excellence (NICE), which drafts treatment guidelines for the National Health Service in the United Kingdom, has established a 3-point difference between drug and placebo on the HAM-D as a criterion of clinical significance [NICE, 2004]. Thus, when published and unpublished data are combined, they fail to show a clinically significant advantage for antidepressant medication over inert placebo.

Some have argued that the NICE criterion is arbitrary (e.g., [Turner and Rosenthal, 2008]), and they are correct. It is as arbitrary as using p < 0.05 as a cutoff for statistical significance. However, Joanna Moncrieff and I have discovered a non-arbitrary criterion for clinical significance [Moncrieff and Kirsch, 2015]. In 2013, Stefan Leucht and his colleagues [Leucht et al., 2013] compared ratings on the HAM-D with those made on the Clinical Global Impressions-Improvement (CGI-I) scale [Guy, 1976]. The CGI-I is a 7-point scale, on which clinicians rate patients from 1 (very much improved) through 4 (no change) to 7 (very much worse). Using patient level data from 43 clinical trials involving 7,131 patients, Leucht established that the mean change on the HAM-D for patients rated on the CGI-I as not having changed at all was 3 points,

exactly what NICE had set as a criterion of clinically meaningful improvement. So as it turns out, the problem with the NICE criterion is that it is too lenient. A 3-point difference on the HAM-D is not even detectable by clinicians as any change at all in their patients. A more reasonable criterion would be the HAM-D change that corresponds to a CGI-I rating of 'minimal improvement'. Leucht et al.'s data indicates that a rating of 'minimal improvement' is equivalent to a 7-point decrease in HAM-D scores.

I should mention here the difference between statistical significance and clinical significance. Statistical significance concerns how reliable an effect is. Is it a real effect, or is it just due to chance? Statistical significance does not tell you anything about the size of the effect. Clinical significance, on the other hand, deals with the size of an effect and whether it would make any difference in a person's life. Imagine, for example, that a study of 500,000 people has shown that smiling increases life expectancy – by 5 minutes. With 500,000 subjects, I can virtually guarantee you that this difference will be statistically significant, but it is clinically meaningless.

Our analyses have since been replicated repeatedly [Fountoulakis and Möller, 2011; Fournier et al., 2010; NICE, 2004; Turner et al., 2008]. Some of the replications used our data; others analyzed different sets of clinical trials. The FDA even did its own meta-analysis on all of the antidepressants that they have approved [Khin et al., 2011]. Despite differences in the way the data have been spun, the numbers are remarkably consistent. Differences on the HAM-D are consistently small – always below the level corresponding to a CGI-I rating of 'no change'. Thomas P. Laughren, the director of the FDA's psychiatry products division, acknowledged this on the American television news program *60 Minutes.* He said, 'I think we all agree that the changes that you see in the short-term trials, the difference in improvement between drug and placebo, is rather small.'

It is not only the short-term trials that show a small, clinically insignificant difference between drug and placebo. In their meta-analysis of published clinical trials, NICE (2004) found that the difference between drug and placebo in the long-term trials were no larger than those in the short-term trials. The difference between drug and placebo is small – it is so small that clinicians cannot detect it at all.

Severity of Depression and Antidepressant Effectiveness

Critics of our 2002 meta-analysis have argued that our results were based on clinical trials conducted on subjects who were not very depressed. In more depressed patients, they argued, a more substantial difference would certainly be found. In fact, it was this criticism that led my colleagues and I to reanalyze the FDA data in 2008 [Kirsch et al., 2008]. We categorized the clinical trials in the FDA database according to the severity of the patients' depression at the beginning of the trial, using conventionally used categories of depression. As it turns out, only 1 of the trials was conducted on moderately depressed patients, and that trial failed to show any sig-

nificant difference between drug and placebo. Indeed, the difference was virtually nil (0.07 points on the HAM-D). All of the rest of the trials were conducted on patients whose mean baseline scores put them in the 'very severe' category of depression, and even among these patients, the drug-placebo difference was below the level of clinical significance.

Still, severity did make a difference. Patients at the very extreme end of depression severity, those scoring at least 28 on the HAM-D, showed an average drug-placebo difference of 4.36 points. This is above the criterion for clinical significance proposed by NICE (2004), but it is well below the 7-point difference that corresponds to a CGI-I rating of 'minimal improvement.'

To find out how many patients fell within this extremely depressed group, I asked Mark Zimmerman from the Brown University School of Medicine to send me the raw data from a study in which he and his colleagues assessed HAM-D scores of patients who had been diagnosed with unipolar major depressive disorder (MDD) after presenting for an intake at a psychiatric outpatient practice [Zimmerman et al., 2005]. Patients with HAM-D scores of 28 or above represented approximately 11% of these patients. This suggests that nearly 90% of depressed patients are not receiving a clinically significant benefit from the antidepressants that are prescribed for them.

Yet, this 11% figure may overestimate the number of people who benefit from antidepressants. Antidepressants are also prescribed to people who do not qualify for the diagnosis of major depression. My neighbor's pet dog died; his physician prescribed an antidepressant. A friend in the US was diagnosed with lumbar muscle spasms and was prescribed an antidepressant. I have lost count of the number of people who have told me they were prescribed antidepressants when complaining of insomnia – even though insomnia is a frequently reported side effect of antidepressants. About 20% of patients suffering from insomnia in the United States are given antidepressants as a treatment by their primary care physicians [Simon and VonKorff, 1997], despite the fact that 'the popularity of antidepressants in the treatment of insomnia is not supported by a large amount of convincing data, but rather by opinions and beliefs of the prescribing physicians' [Wiegand, 2008].

Attempts by other researchers to evaluate the association between initial severity and drug-placebo differences have yielded mixed results. Some find the same association that we found [e.g., Fournier et al., 2010; Khin et al., 2011], whereas others find no association between severity and drug-placebo differences [e.g., Fountoulakis et al., 2013; Locher et al., 2015]. However, all meta-analyses find overall drug-placebo differences that are below the NICE criteria for clinical significance. So the question is, is there a subset of very severely depressed patients for whom antidepressants are clinically effective, or do they lack effectiveness at all levels of severity?

Predicting Response to Treatment

Severity of depression is one of the few predictors of response to treatment. Type of antidepressant has little if any impact on treat-

ment response. As summarized in a 2011 meta-analysis of studies comparing one antidepressant to another:

> On the basis of 234 studies, no clinically relevant differences in efficacy or effectiveness were detected for the treatment of acute, continuation, and maintenance phases of MDD. No differences in efficacy were seen in patients with accompanying symptoms or in subgroups based on age, sex, ethnicity, or comorbid conditions...Current evidence does not warrant recommending a particular second-generation antidepressant on the basis of differences in efficacy [Gartlehner et al., 2011].

Although type of medication does not make a clinically significant difference in outcome, response to placebo does. Almost all antidepressant trials include a placebo run-in phase. Before the trial begins, all of the patients are given a placebo for a week or two. After this run-in period, the patients are reassessed, and anyone who has improved substantially is excluded from the trial. That leaves patients who have not benefitted at all from placebo and those who have benefited only a little bit. These are the patients who are randomized to be given a drug or kept on placebo. As it turns out, the patients who show at least a little improvement during the run-in period are the ones most likely to respond to the real drug, as shown not only by physician ratings, but also by changes in brain function [Hunter et al., 2006; Quitkin et al., 1998].

How Did These Drugs Get Approved?

How is it that medications with such weak efficacy data were approved by the FDA? The answer lies in an understanding of the approval criteria used by the FDA. The FDA requires 2 adequately conducted clinical trials showing a significant difference between drug and placebo. But there is a loophole: There is no limit to the number of trials that can be conducted in search of these 2 significant trials. Trials showing negative results simply do not count. Furthermore, the clinical significance of the findings is not considered. All that matters is that the results are statistically significant.

The most egregious example of the implementation of this criterion is provided by the FDA's approval of Viibryd® (Actavis, Inc., Dublin, Ireland) in 2011. Seven controlled efficacy trials were conducted. The first 5 failed to show any significant differences on any measure of depression, and the mean drug-placebo difference in these studies was less than half a point on the HAM-D, and in 2 of the 5 trials, the direction of the difference actually favored the placebo. The company ran 2 more studies and managed to obtain small but significant drug-placebo differences (1.70 points). The mean drug-placebo difference across the 7 studies was 1.01 HAM-D points. This was sufficient for the FDA to grant approval, and the information approved by the FDA for informing doctors and patients reads, 'The efficacy of Viibryd® was established in two 8-week, randomized, double-blind, placebo-controlled trials'. No mention is made of the 5 failed trials that preceded the 2 successful ones.

The failure to mention the unsuccessful trials was not merely an oversight; it reflects a carefully decided FDA policy dating back for decades. To my knowledge, there is only 1 antidepressant in which the FDA included information on the existence of negative trials.

The exception is citalopram, and the inclusion of the information followed an objection raised by Paul Leber, who was at the time the director of the FDA Division of Neuropharmacological Drug Products. In an internal memo dated May 4, 1998, Leber wrote:

> One aspect of the labelling deserves special mention. The (report) not only describes the clinical trials providing evidence of citalopram's antidepressant effects, but make mention of adequate and well controlled clinical studies that failed to do so...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 that the public, or even the majority of the medical community, is aware of this fact. I am persuaded that they not only have a right to know but that they should know. Moreover, I believe that labeling that selectively describes positive studies and excludes mention of negative ones can be viewed as potentially 'false and misleading'.

Hooray for Paul Leber. I have never met or corresponded with this gentleman, but because of this courageous memo, he is one of my heroes.

The Serotonin Myth

Over the years, I have noticed something very strange in the antidepressant literature. When different antidepressants are compared with each other, their effects are remarkably similar. I first noticed this when Guy Sapirstein and I did our 1998 meta-analysis of the published literature. When we first saw how small the actual drug effect was, we thought we might have done something wrong. Perhaps we had erred by including trials that had evaluated different types of antidepressants. Maybe we were underestimating the true effectiveness of antidepressants by including clinical trials of drugs that were less effective than others.

Before submitting our paper for publication, we went back to the data and examined the type of antidepressant used in each trial. Some were selective serotonin reuptake inhibitors (SSRIs), others were tricyclic medications; we lumped together the trials on antidepressant drugs that were neither SSRIs nor tricyclics and called them 'other antidepressants'. And then we noticed that there was a 4th category of drugs in the trials we had analyzed. These were trials in which drugs that are not thought to be antidepressants at all – tranquilizers and thyroid medications, for example – were given to depressed patients and evaluated for their effect on depression.

When we analyzed the drug and placebo response for each type of drug, we found another surprise awaiting us. It did not matter what kind of drug the patients had been given in the trial. The response to the drug was always the same, and 75% of that response was also found in the placebo groups. I recall being impressed by how unusual the similarity in results was, but I have since learned that they are not unusual at all. I have since encountered this phenomenon over and over again. In the STAR*D trial, which, at a cost of $35,000,000, is the most costly clinical trial of antidepressants ever conducted, patients who did not respond to the pre-

scribed SSRI were switched to a different antidepressant. Some were switched to a serotonin-norepinephrine reuptake inhibitor (SNRI), a drug that is supposed to increase norepinephrine as well as serotonin in the brain. Others were switched to a norepinephrine-dopamine reuptake inhibitor (NDRI), which is supposed to increase norepinephrine and dopamine without affecting serotonin at all. And still others were simply given a different SSRI. About 1 out of 4 patients responded clinically to the new drug, but it did not matter which new drug they were given. The effects ranged from 26 to 28%; in other words, they were exactly the same regardless of the type of drug.

The most commonly prescribed antidepressants are SSRIs, drugs that are supposed to selectively target the neurotransmitter serotonin. But there is another antidepressant that has a very different mode of action. It is called tianeptine, and it has been approved for prescription as an antidepressant by the French drug regulatory agency. Tianeptine is an SSRE, a selective serotonin reuptake enhancer. Instead of increasing the amount of serotonin in the brain, it is supposed to decrease it. If the theory that depression is caused by a serotonin deficiency were correct, we would expect tianeptine to make depression worse. But it does not. In clinical trials comparing the effects of tianeptine to those of SSRIs and tricyclic antidepressants, 63% of patients showed significant improvement (arbitrarily defined as a 50% reduction in symptoms), the same response rate that is found for SSRIs, NDRIs, and tricyclics, in this type of trial. It simply does not matter what is in the medication – it might increase serotonin, decrease it, or have no effect on serotonin at all. The effect on depression is the same.

What do you call pills, the effects of which are independent of their chemical composition? I call them 'placebos'.

Antidepressants as Active Placebos

All antidepressants seem to be equally effective, and although the difference between drug and placebo is not clinically significant, it is significant statistically. This leads to the obvious question: What do all of these active drugs have in common that make their effect on depression slightly, but statistically significantly, better than placebo?

One thing that antidepressants have in common is that they all produce side effects. Why is that important? Imagine that you are a subject in a clinical trial. You are told that the trial is double blind and that you might be given a placebo. You are told what the side effects of the medication are. The therapeutic effects of the drug may take weeks to notice, but the side effects might occur more quickly. Would you not wonder to which group you had been assigned, drug or placebo, and noticing one of the listed side effects, would you not conclude that you had been given the real drug? In one study, 89% of depressed patients in the drug group correctly 'guessed' that they had been given the real antidepressant [Rabkin et al., 1986], and in a study examining antidepressants and benzodiazepines in the treatment of panic disorder, 95% of patients in the active drug groups broke blind [Margraf et al., 1991].

It is not only patients who break blind, but also the clinicians who rate them on the HAM-D [Margraf et al., 1991; Rabkin et al., 1986]. In fact, clinicians do better than patients at figuring out the group to which the patient has been assigned. Patients given the real drug do well at 'guessing' that they are in the drug group, but those in placebo arms are much less accurate. In contrast, clinicians doing the ratings are very accurate in identifying the group to which the patient has been assigned for both those in the drug arms and those in the placebo arms. Clinician-rated scales have been designated as the primary outcome in all antidepressant trials submitted to the FDA. It is possible that the use of scales on which patients rate their depressive symptoms would produce more accurate estimates of drug-placebo differences.

In other words, clinical trials are not really double blind. Many patients in clinical trials realize when they have been given the real drug and so do the clinicians who are rating their levels of depression, most likely because of the drug's side effects. What effect is this likely to have in a clinical trial? We do not have to guess at the answer to this question. Bret Rutherford and his colleagues at Columbia University have provided the answer. They compared the responses to antidepressants in studies that did not have a placebo group with those in studies where they did have a placebo group [Rutherford et al., 2009]. The main difference between these studies is that in the first case, patients and raters were certain that the patients were getting an active antidepressant, whereas in the placebo-controlled trials, they knew that they might be given a placebo. Knowing that all patients were getting an active drug boosted the effectiveness of the drug significantly. This supports the hypothesis that the relatively small difference between drug and placebo in antidepressant trials is at least in part due to 'breaking blind' and discerning that the patient is in the drug group, because of the side effects produced by the drug.

What to Do?

To summarize, there is a strong therapeutic response to antidepressant medication. However, the response to placebo is almost as strong. This presents a therapeutic dilemma. The drug effect of antidepressants is not clinically significant, but the placebo effect is. What should be done clinically in light of these findings?

One possibility would be to prescribe placebos, but this entails deception. Besides being ethically questionable, it runs the risk of undermining trust, which may be one of the most important clinical tools that clinicians have at their disposal. Another possibility that has been proposed is to use antidepressants as active placebos. But the risks involved in antidepressant use render this alternative problematic. Among the side effects of antidepressants are sexual dysfunction (which can affect between 70 and 96% of patients on SSRIs [Clayton et al., 2006; Serretti et al., 2009]), long-term weight gain, insomnia, nausea, and diarrhea. Many people who attempt to quit taking antidepressants show withdrawal symptoms [Rosenbaum et al., 1998]. Antidepressants have been linked to increases in suicidal ideation and violent criminal activity among children, ad-

olescents, and young adults [Molero et al., 2015; Stone, 2014; Stone et al., 2009]. Older adults have increased risks of stroke and death from all causes [Andrews et al., 2012]. Pregnant women using antidepressants are at increased risk of miscarriage, and if they do not miscarry, their offspring are more likely to be born with autism, birth malformations, and persistent pulmonary hypertension [Dolmar et al., 2013]. Furthermore, some of these risks have been linked to antidepressant use during the first trimester of pregnancy, when women may not be aware that they are pregnant.

Perhaps the most surprising health consequence of antidepressant use is one that affects people of all ages. Antidepressants increase the risk of relapse after one has recovered. People are more likely to become depressed again after treatment by antidepressants than after treatment by other means – including placebo treatment [Andrews et al., 2012; Babyak et al., 2000; Dobson et al., 2008]. Furthermore, the degree to which the risk of relapse increases depends on the degree to which the particular antidepressant used changes neurotransmission in the brain.

Given these health risks, antidepressants should not be used as a first-line treatment for depression. A better alternative is the use of non-drug treatments. My colleagues and I have conducted a meta-analysis of various treatments for depression, including antidepressants, psychotherapy, the combination of psychotherapy and antidepressants, and 'alternative' treatments, which included acupuncture and physical exercise [Khan et al., 2012]. We found no significant differences between these treatments or within different types of psychotherapy. When different treatments are equally effective, choice should be based on risk and harm, and of all of these treatments, antidepressant drugs are the riskiest and most harmful. If they are to be used at all, it should be as a last resort, when depression is extremely severe and all other treatment alternatives have been tried and failed.

Disclosure Statement

The author declares that there is no conflict of interest concerning this paper.

References

Andrews PW, Thomson JA, Amstadter A, Neale MC: Primum non nocere: an evolutionary analysis of whether antidepressants do more harm than good. Review. Front Psychol 2012;3:117.

Babyak MA, Blumenthal JA, Herman S, et al.: Exercise treatment for major depression: maintenance of therapeutic benefit at 10 months. Psychosom Med 2000;62:633–638.

Clayton A, Keller A, McGarvey EL: Burden of phase-specific sexual dysfunction with SSRIs. J Affect Disord 2006;91:27–32.

Dobson KS, Hollon SD, Dimidjian S, et al.: Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the prevention of relapse and recurrence in major depression. J Consult Clin Psychol 2008;76:468–477.

Domar AD, Moragianni VA, Ryley DA, Urato AC: The risks of selective serotonin reuptake inhibitor use in infertile women: a review of the impact on fertility, pregnancy, neonatal health and beyond. Hum Reprod 2013;28:160–171.

Fountoulakis KN, Möller HJ: Efficacy of antidepressants: a re-analysis and re-interpretation of the Kirsch data. Int J Neuropsychopharmacol 2011;14:405.

Fountoulakis KN, Veroniki AA, Siamouli M, Moller H: No role for initial severity on the efficacy of antidepressants: results of a multi-meta-analysis. Ann Gen Psychiatry 2013;12:26.

Fournier JC, DeRubeis RJ, Hollon SD, et al.: Antidepressant drug effects and depression severity: a patient-level meta-analysis. JAMA 2010;303:47–53.

Gartlehner G, Hansen RA, Morgan LC, et al.: Comparative benefits and harms of second-generation antidepressants for treating major depressive disorder. Ann Intern Med 2011;155:772–785.

Guy W: ECDEU Assessment Manual for Psychopharmacology – Revised. Volume DHEW Publ No ADM 76–338. Rockville, MD, Department of Health, Education, and Welfare 1976, pp. 218–222.

Hunter AM, Leuchter AF, Morgan ML, Cook IA: Changes in brain function (quantitative EEG cordance) during placebo lead-in and treatment outcomes in clinical trials for major depression. Am J Psychiatry 2006;163:1426–1432.

Khan A, Faucett J, Lichtenberg P, et al.: A systematic review of comparative efficacy of treatments and controls for depression. PLoS One 2012;7:e41778.

Khin NA, Chen YF, Yang Y, et al.: Exploratory analyses of efficacy data from major depressive disorder trials submitted to the US Food and Drug Administration in support of new drug applications. J Clin Psychiatry 2011;72:464.

Kirsch I: The Emperor's New Drugs: Exploding the Antidepressant Myth. London, The Bodley Head, 2009.

Kirsch I, Deacon BJ, Huedo-Medina TB, et al.: Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. PLoS Med 2008;5:e45.

Kirsch I, Sapirstein G: Listening to Prozac but hearing placebo: a meta-analysis of antidepressant medication. Prevention and Treatment 1998;1:article 0002a.

Locher C, Kossowsky J, Gaab J, et al.: Moderation of antidepressant and placebo outcomes by baseline severity in late-life depression: a systematic review and meta-analysis. J Affect Disord 2015;181:50–60.

Margraf J, Ehlers A, Roth WT, et al.: How 'blind' are double-blind studies? J Consult Clin Psychol 1991;59:184–187.

Molero Y, Lichtenstein P, Zetterqvist J, Gumpert CH, Fazel S: Selective serotonin reuptake inhibitors and violent crime: a cohort study. PLoS Med 2015;12:e1001875.

Moncrieff J, Kirsch I: Empirically derived criteria cast doubt on the clinical significance of antidepressant-placebo differences. Contemp Clin Trials 2015;43:60–62.

NICE: Depression: Management of Depression in Primary and Secondary Care. Clinical Practice Guideline No 23, 2004. www.nice.org.uk/guidance/cg23 (accessed 24 May, 2005).

Quitkin FM, McGrath PJ, Stewart JW, et al.: Placebo run-in period in studies of depressive disorders: clinical, heuristic and research implications. Br J Psychiatry 1998;173:242–248.

Rabkin JG, Markowitz JS, Stewart JW, et al.: How blind is blind? Assessment of patient and doctor medication guesses in a placebo-controlled trial of imipramine and phenelzine. Psychiatry Res 1986;19:75–86.

Rosenbaum JF, Fava M, Hoog SL, et al.: Selective serotonin reuptake inhibitor discontinuation syndrome: a randomized clinical trial. Biol Psychiatry 1998;44:77–87.

Rutherford BR, Sneed JR, Roose SP: Does study design influence outcome? The effects of placebo control and treatment duration in antidepressant trials. Psychother Psychosom 2009;78:172–181.

Serretti A, Chiesa A: Treatment-emergent sexual dysfunction related to antidepressants: a meta-analysis. J Clin Psychopharmacol 2009;29:259–266.

Simon GE, VonKorff M: Prevalence, burden, and treatment of insomnia in primary care. Am J Psychiatry 1997;154:1417–1423.

Stone M, Laughren T, Jones ML, et al.: Risk of suicidality in clinical trials of antidepressants in adults: analysis of proprietary data submitted to US Food and Drug Administration. BMJ 2009;339:b2880.

Stone MB: The FDA warning on antidepressants and suicidality – why the controversy? N Engl J Med 2014;371:1668–1671.

Turner EH, Matthews AM, Linardatos E, et al.: Selective publication of antidepressant trials and its influence on apparent efficacy. N Engl J Med 2008;358:252–260.

Turner EH, Rosenthal R: Efficacy of antidepressants. BMJ 2008;336:516–517.

Wiegand MH: Antidepressants for the treatment of insomnia: a suitable approach? Drugs 2008;68:2411–2417.

Zimmerman M, Chelminski I, Posternak MA: Generalizability of antidepressant efficacy trials: differences between depressed psychiatric outpatients who would or would not qualify for an efficacy trial. Am J Psychiatry 2005;162:1370–1372.