Special Articles

Antidepressants and the placebo response

IRVING KIRSCH

University of Hull (United Kingdom)

SUMMARY. **Aims** – To evaluate new generation antidepressants in relation to the placebo response. **Methods** – I review meta-analyses in which response to antidepressant medication and response to placebo were calculated. **Results** – All but one of these meta-analyses included unpublished as well as published trials. Most trials failed to show a significant advantage of SSRIs over inert placebo, and the differences between drug and placebo are not clinically significant for most depressed patients. Documents obtained from the U.S. Food and Drug Administration (FDA) revealed an explicit decision to keep this information from the public and from prescribing physicians. **Conclusions** – Because they do not incur drug risks, exercise and psychotherapy, which show at benefits at least equal to those of antidepressants, may be a better treatment choice for depressed individuals.

Declaration of Interest: The author has not in the last 2 years received any support, including that from drug companies and honoraria for lectures and consultancies, from interests potentially in conflict with this work.

Received 09.05.2009 – Accepted 10.05.2009.

INTRODUCTION

The meta-analyses reviewed in this article were initiated not because of any interest in evaluating the effects of antidepressants, but because of my long-standing interest in the effects of response expectancy (Kirsch, 1985). Response expectancies are anticipations of automatic subjective reactions, like changes in depression, anxiety, pain, etc. I have argued that response expectancies are self-confirming. The world in which we live is ambiguous, and one of the functions of the brain is to disambiguate it rapidly enough to respond quickly. We do this, in part, by forming expectations. So what we experience at any given time is a joint function of the stimuli to which we are exposed and our beliefs and expectations about those stimuli (Kirsch, 1999).

This response expectancy hypothesis has been the focus of most of my research. The particular topic areas (hypnosis, psychotherapy, placebo effects, etc.) were chosen merely because they provided a convenient opportunity for examining expectancy effects. It seemed to me that depression ought to be particularly responsive to expectancy effects. This is because hopelessness is a central feature of depression (Abramson *et al.*, 1978), and hopelessness is an expectancy. Specifically, it is the expectancy that a negative state of affairs will not get better, no matter what one does to alleviate it.

If you asked depressed people what the worst thing in their lives is, many will tell you that it is their depression. They believe that their depression will continue, no matter what they do – a very depressing thought indeed. As John Teasdale (1985) noted, these people are depressed about their depression. If this is the case, then the expectancy of improvement should produce improvement. The belief that one will improve is the opposite of the hopelessness that may be maintaining the depression or, at the very least, it is an important component of it. For this reason there ought to be a substantial placebo effect associated with the treatment of depression.

Address for correspondence: Professor I. Kirsch, Department of Psychology, University of Hull, Cottingham Road, Hull, HU6 7RX (United Kingdom).

E-mail: i.kirsch@hull.ac.uk

LISTENING TO PROZAC BUT HEARING PLACEBO

In 1998, Guy Sapirstein, and I undertook a meta-analysis, the purpose of which was to evaluate the placebo effect in depression (Kirsch & Sapirstein, 1998). We searched the literature for studies in which depressed patients had been randomized to receive antidepressant medication, an inert placebo, psychotherapy, or no treatment at all. We included studies of psychotherapy, because those were the only ones in which patients had been randomized to a no-treatment control condition, and we needed that condition to evaluate the placebo effect. The response to a placebo is not the same as the effect of the placebo. The placebo response (as opposed to the placebo effect) may at least in part be due to the passage of time, spontaneous remission, the natural history of the disorder, and regression to the mean. Just as the difference between the drug response and the placebo response is deemed to be the drug effect, so the difference between the placebo response and improvement in a no-treatment control group can be interpreted as the placebo effect.

The results of our meta-analysis indicated substantial improvement among patients given medication (Standardized Mean Difference = 1.55) or psychotherapy (SMD = 1.60). However, patients given placebos also improved (SMD = 1.16), whereas those in no-treatment control groups showed relatively little improvement (SMD = 0.37). This meant that approximately 25% of the improvement in the drug group would have occurred without any treatment whatsoever, 50% was a placebo effect, and only 25% was a true drug effect.

Our surprise at the outcome of our analysis led us to wonder whether it may have been due to the diversity of antidepressants in the clinical trials we had analyzed. Perhaps some of the medications were very effective and others not, leading us to underestimate the drug effect. To assess this possibility, we returned to our data set and classified the various studies in terms of the type of medication evaluated. We categorized them into four types: tricyclic medications, SSRIs, miscellaneous other antidepressants, and other medications. The consistency was remarkable. Regardless of the type of medication studied, 75% of the response to the active drugs was duplicated by placebo, leaving a true drug effect of only 25% in each case. What makes this particularly surprising is the response to what we have labelled 'other medication'. These are active drugs that are not regarded as antidepressants (e.g., lithium, barbiturates, and thyroid medication given to depressed patients who were not suffering from depression). They too produced substantial improvement in depression, as great as that produced by tricyclics, SSRIs, and other antidepressants.

The finding of equivalent antidepressive effects of all of these different drugs led us to search for their commonality. One thing they have in common is that they all produce side effects. Placebos can also produce side effects, but they do so to a much lesser degree than active medications (Philipp *et al*., 1999). Why is this important? Imagine that you are recruited to a clinical trial for an antidepressant medication. As this is a double blind trial, you are told that you may receive medication or you may receive placebo. You are also told that the active medication has been reported to produce a number of side effects, such as dry mouth and drowsiness, and you are told that the therapeutic effect may not become evident for some weeks. You are likely to wonder to which group you have been assigned, the active drug group or the placebo control group. You notice that your mouth has become dry and that your feel drowsy. At this point, you are likely to conclude that you have been assigned to the drug condition. Indeed, data indicate that about 80% of patients assigned to the active drug condition in clinical trials of antidepressants break blind and conclude that they are in the active drug condition (Rabkin *et al*., 1986). Being more certain that you have been assigned to the drug group, you will have a stronger expectancy for improvement, which according to the response expectancy hypothesis should produce greater improvement. In other words, it is possible that the superiority of active antidepressant to inert placebo is due to the breaking of blind by patients in the active drug condition. Rather than being a true drug effect, it is an enhanced placebo effect.

THE EMPEROR'S NEW DRUGS

Needless to say, our first meta-analysis proved to be quite 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 and their effectiveness had been well established.

In an effort to respond to these critics, we decided to replicate our study with a different set of clinical trials (Kirsch *et al*., 2002). 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 six 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. First, 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. Second, the same primary outcome measure – the Hamilton depression scale (HAM-D) – was used in all of the trials. That made it easy to understand the clinical significance of the drug-placebo differences. Third, these were the data on the basis of which the medications were approved. In that sense they have a privileged status. If there is anything wrong with them, the decision to approve the medications in the first place can be called into question.

In the data sent to us by the FDA, only 43% of the trials showed a statistically significant benefit of drug over placebo. 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 found 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 two points on the HAM-D. The National Institute for Clinical Excellence (NICE), which drafts treatment guidelines for the National Health Service in the United Kingdom, has a established as 3 point difference between drug and placebo on the HAM-D as a criterion of clinical significance (National Institute for Clinical Excellence, 2004). Thus, when published and unpublished data are combined, they fail to show a clinically significant advantage for antidepressant medication over inert placebo.

At roughly the same time as our second meta-analysis of the FDA data set was done, Corrado Barbui and his colleagues (Barbui *et al*., 2008) analyzed the data on paroxetine that had been reported on the GlaxoSmithKline (GSK) website. As part of the settlement of a lawsuit brought against GSK by the State of New York for withholding data showing negative results, the company is required to post summary data from all of its clinical trials of antidepressants, including those that have not been published (Spitzer, 2004). Unlike the FDA files, which are limited to pre-approval trial, the GSK website includes post-marking trials as well. Barbui *et al*. (2008) found 40 placebo-controlled studies of Seroxat for the treatment of major depression, including the 16 that had been sent to the FDA. Although they analyzed response rates rather than mean symptom change, the results of their analysis of these 40 studies were virtually identical to the results of our analysis of the studies that had been sent to the FDA. In their analysis, the placebo was 83% as effective as the real drug. Thus the failure to find a clinically significant difference between drug and placebo holds for post-marketing as well as pre-marketing trials.

There are two types of design that were used in the clinical trials submitted to the FDA. The most common involved allowing prescribing physicians to adjust the dose as needed during the course of the trial. In addition, approximately 1/4 of trials used a fixed dose design, in which patients were randomized to receive particular doses of the medication. Thus we were concerned that the data we had analyzed might have included patients who were assigned to receive an inadequate or sub-clinical dose of the medication. If this were the case, then we might have underestimated the drug effect.

To check out this possibility, we performed an additional analysis on the fixed-dose clinical trials. Specifically, we compared improvement among patients given the lowest dose used in the trial with those with improvement among patients given the highest dose. We found that improvement at the lowest dose (9.57 points on the HAM-D) was virtually identical to improvement at the highest dose (9.97 on the HAM-D). Nor was there any apparent advantage for mid range doses. In fact, out of approximately 40 comparisons of different doses of the same antidepressant, only one significant difference was reported. In a study of fluoxetine in moderately to severely depressed patients, the two lower doses were significantly more effective than the high dose, which was not significantly more effective than placebo.

THE "DIRTY LITTLE SECRET"

Whereas the response to our earlier meta-analysis was incredulity, the response to our analysis of the FDA data indicated unanimous acceptance among 12 groups of independent scholars-some of them clinical trialists who had carried out evaluations of antidepressants for pharmaceutical companies-who had been invited to comment on the paper. As one group of commentators put it, "many have long been unimpressed by the magnitude of the differences observed between treatments and controls, what some of our colleagues refer to as the 'dirty little secret' in the pharmaceutical literature" (Hollon *et al*., 2002).

Perhaps the most disturbing aspect of the keeping of this secret is the complicity of the FDA. Among the data we received using our Freedom of Information request were copies of internal memos. One of these, written by the Director of the Division of Neuropharmacological Drug Products includes the following revealing information.

"The Clinical Efficacy Trials subsection within the Clinical Pharmacology section 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. I am mindful, based on prior discussions of the issue, that the Office Director is inclined toward the view that the provision of such information is of no practical value to either the patient or prescriber. I disagree. I believe it is useful for the prescriber, patient, and 3rd party payer to know, without having to gain access to official FDA review documents, that citalopram's antidepressants effects were not detected in every controlled clinical trial intended to demonstrate those effects. I am aware that clinical studies often fail to document the efficacy of effective drugs, but I doubt the public, or even the majority of medical community, are aware of this fact. I am persuaded they not only have a right to know, but should know. Moreover, I believe that labeling that selectively describes positive studies and excludes mention of negative ones can be viewed as potentially 'false and misleading'" (Leber, 1998, p. 11).

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 two 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 two 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.

A typical example of the implementation of this criterion is provided by the FDA file on citalopram. Seven controlled efficacy trials were conducted. Two showed small but significant drug-placebo differences. Two were deemed too small to count. Three failing to show any significant benefit for the drug were deemed "adequate: and "well controlled," but were "not counted against citalopram" because there was a "substantial placebo response" (Internal memo by T. P. Laughren, FDA Team Leader for Psychiatric Drug Products). Thus, citalopram was approved on the basis of two clinical trials despite negative results in five other trials.

CLINICAL CONCLUSIONS

To summarize, there is a strong therapeutic response to antidepressant medication. But 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 (Hollon, *et al.*, 2002; Moerman, 2002). But the risks of side effects, suicide, withdrawal symptoms and drug interactions render this alternative problematic (Kirsch, in press).

A third possibility is the use of alternative treatments. Physical exercise, for example, has been shown to produce clinical benefit in moderately depressed people (reviewed in Kirsch, 2009). This might also be a placebo effect, but the difference in the side effect profile can be considered. Side effects of antidepressants include sexual dysfunction, insomnia, diarrhea, nausea, anorexia, bleeding, forgetfulness, seizures, and increased suicide risk. Side effects of physical exercise include enhanced libido, better sleep, decreased body fat, improved muscle tone, longer life, increased strength and endurance, and improved cholesterol levels. Finally, psychotherapy might be considered as a first choice treatment for depression. The data indicate that psychotherapy is as effective as medication in the short run and considerably more effective when relapse rates are assessed, even for severely and chronically depressed patients (reviewed in Kirsch, 2009).

REFERENCES

Abramson L.Y., Seligman M.E.P. & Teasdale J.D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology* 87, 49-74.

Barbui C., Furukawa T.A. & Cipriani A. (2008). Effectiveness of paroxetine in the treatment of acute major depression in adults: a systematic re-examination of published and unpublished data from randomized trials. *Canadian Medical Association Journal* 178, 296-305.

Hollon S.D., DeRubeis R.J., Shelton R.C. & Weiss B. (2002). The Emperor's New Drugs: Effect Size and Moderation Effects. *Prevention and Treatment* 5, Jul 2002, ArtID 28.

Kirsch I. (1985). Response rxpectancy as a determinant of experience and behavior. *American Psychologist* 40, 1189-1202.

Kirsch I. (Ed.) (1999). *How Expectancies Shape Experience*. American Psychological Association: Washington DC.

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

Kirsch I., & Sapirstein G. (1998). Listening to Prozac but hearing placebo: A meta-analysis of antidepressant medication. *Prevention and Treatment* 1, Jun 1998, ArtID 2a.

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 and Treatment* 5, Jul 2002, ArtID 23.

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 Food and Drug Administration. *PLoS Medicine 5*. Retrieved August 21, 2009, from http://medicine.plosjournals.org/perlserv/?request=get-document&doi=10.1371/journal.pmed.0050045

Moerman D.E. (2002). "The loaves and the fishes": A comment on "The emperor's new drugs: An analysis of antidepressant medication data submitted to the U.S. Food and Drug Administration". *Prevention and Treatment* 5, Jul 2002, ArtID 29.

National Institute for Clinical Excellence (2004). *Depression: Management of Depression in Primary and Secondary Care*. Clinical practice guideline No. 23 Retrieved August 21, 2009, from http://www.nice.org.uk/guidance/CG23

Philipp M., Kohnen R. & Hiller K.O. (1999). Hypericum extract versus imipramine or placebo in patients with moderate depression: randomised multicentre study of treatment for eight weeks *British Medical Journal* 319, 1534-1539.

Rabkin J.G., Markowitz J.S., Stewart J.W., McGrath P.J., Harrison W., Quitkin F.M. & Klein D.F. (1986). How blind is blind? Assessment of patient and doctor medication guesses in a placebo-controlled trial of imipramine and phenelzine. *Psychiatry Research* 19, 75-86.

Spitzer E. (2004). Major Pharmaceutical Firm Concealed Drug Information: GlaxoSmithKline Misled Doctors About the Safety of Drug Used to Treat Depression in Children. Press Release: Office of the New York State Attorney General, June 2, 2004. Retrieved August 21, 2009, from http://www.oag.state.ny.us/media_center/2004/jun/jun2b_04.html

Teasdale J.D. (1985). Psychological treatments for depression: How do they work? *Behaviour Research and Therapy* 23, 157-165.