Methodologic Issues in the Clinical Trial Study of Depression

Doug Berger, M.D., Ph.D.
Japan Pharmaceutical Consultant
www.japanpsychiatrist.com/DrugDev.html

A recent study by Kirsch et al. entitled, *Initial Severity and Antidepressant Benefits: A Meta-Analysis of Data Submitted to the Food and Drug Administration* published by the Public Library of Science (PLoS Med February 26, 2008) caused a stir in the mass media who chimed in with sound bites like, "Antidepressants don't work much better than placebos for many depressed patients." This study obtained data on all clinical trials submitted to the FDA for the licensing of four new-generation antidepressants for which full datasets were available (fluoxetine, venlafaxine, nefazodone, and paroxetine) and used meta-analytic techniques to assess linear and quadratic effects of initial severity on improvement scores for drug and placebo groups and on drug-placebo difference scores.

While the study is a poignant reminder of the difficulty of proving statistical efficacy for antidepressants, its message that antidepressants do not work much better than placebo except for the most severely depressed patients will leave many patients with significant depressive symptoms without proper medical care.

This is a summary of the major findings of this study:

1. Efficacy reaches clinical significance for only the most extremely depressed patients, and this is due to a decrease in the response to placebo rather than an increase in the response to medication.
2. A substantial response to placebo was seen in moderately depressed groups and in groups with very severe levels of depression. It decreased somewhat, but was still substantial, in groups with the most-severe levels of depression.
3. Given these data, there seems little evidence to support the prescription of antidepressant medication to any but the most severely depressed patients.

How can we interpret these findings in light of our believing in the incredible power of these drugs to help depression? The reason is that clinical practice, especially for depression, is not practiced like a clinical trial, and that for a number of methodologic reasons, the pharmacodynamics of depression are not easy to study in a clinical trial. The result is that only modest improvement may be seen in the active treatment, and the results of trials of some antidepressants that later obtain market approval are negative.

First, let me enumerate some of the differences between a clinical trial and clinical practice:

1. There are few exclusion criteria, inclusion criteria are wider, co-morbid and complex patients are treated.
2. Flexible dosing and multiple medications are employed.
3. In clinical practice, one treats an individual patient, treatments are not based solely on the results of a clinical trial. Many depressed patients clearly remit when a drug is given and relapse when a drug is removed. Also, it is common for patients to respond only to certain types of drugs, e.g., serotonin promoting drugs but not norepinephrine promoting drugs or vice-versa.

Next, we need to review why the study of efficacy of depression is so difficult:

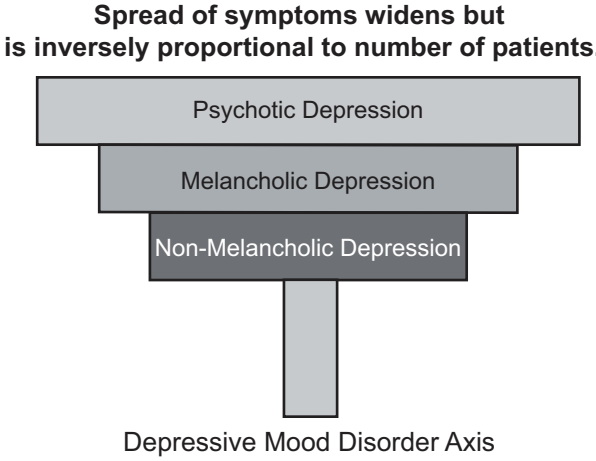
1. Symptoms of depression are not an easily measurable and clearly reactive pharmacodynamic parameter (The action of antidepressants on depression is not as easily measurable as the action of reserpine on blood pressure).
2. Depression is difficult to treat. Close to 50% of subjects, whether on active or placebo will not have much response in a clinical trial. This shows that placebo itself is not exactly a favorable treatment for depression.
3. Depression assessment is subjective, and while there has been much criticism of the use of the Hamilton Depression Rating Scale (HDRS) for a clinical trial, this scale is still seen by the regulatory agencies to be the sine qua non for depression assessment. The Montgomery-Asberg Depression Rating Scale (MADRS) is thought by many experts be an improvement over the HDRS, however, the Kirsch et at. study did not analyze MADRS data, and in general depression rating scales may not be as insightful as a good clinical interview.
4. Many patients will receive sleep medications even on placebo, and this may improve some of the symptoms that are assessed making it seem like the depression has also lifted.

5. Many symptoms rated as part of depression may improve in a therapeutic environment with no drug (hopelessness, low self-esteem), and this may cloud the issue of whether the depression itself had responded as many patients with mild depression respond to a drug or not.

6. The sponsors of the study need to have some response endpoint that they can measure within a 6-8 week time-frame of the study so that "responder" is usually defined as a 50% reduction on the HDRS. So even if you are a responder on placebo, you may still have significant depressive symptoms left over. In other words, many subjects on both sides are "responders". If they have mild non-melancholic depression their depression is not really being treated but they feel better. Even if they have major depression with melancholia, the symptoms that can get better do, but the depression is not better; e.g., 50% reduction should not be considered response.

7. Entry of subjects with only mild symptoms of depression is one of the biggest problems in a clinical trial (Fig 1.). Melancholic depression, or a morbid-type of depression is the most likely type of depression to respond only to active medication. The study sites are motivated to rapidly enter as many subjects as possible, thus many mild (non-melancholic) cases are entered. Non-melancholic patients have enough symptoms of depression to score above the cut-off on the HDRS, but they do not have the loss of appetite, worthlessness, and loss of ability to enjoy life like the melancholic cases, and are more likely to have a number of symptoms that improve in a therapeutic environment under placebo treatment and be coded as a "responder". Psychotic depression should also not be entered in the study as they require an antipsychotic in addition to an antidepressant –this is usually not a problem in a clinical trial.

Fig 1.



8. Drug companies want to do the studies efficiently and at low cost and will set the number of subjects at the lowest number needed based on previous data in order to obtain the estimated statistical power. Because of the methodologic difficulties inherent in the nature of depression studies as described here, this number may not be enough to get wide significance vs placebo.

9. The meta-analysis assumes that all studies were of the same quality, that all the protocols and subject numbers were adequate, that proper and consistent orientation of the sites was comparable, etc. It is well known that inter-site variability in the discrimination of subjects and in the quality of the study is often large in depression studies.

While the conclusion of the Kirsch study may not be valid for the real world of depression treatment, what can we take home from this study to help the medical care and clinical trial construct for depression treatment?

First, we need to improve the methods for clinical trials of depression and the knowledge of these difficulties at the regulatory agency level. We need to show the agencies what makes sense, and work with the experts in the fields, and large pharma needs to take the lead in creating sounder-trials.

Next, I think the response of the industry must be more than a sound bite that sounds like a "sorry you are wrong" message. One large pharma company, for example, was reported to say that the authors of the study had "failed to acknowledge" the very positive benefits of SSRIs and their conclusions were "at odds with the very positive benefits seen in actual clinical practice". I think we must use these opportunities to educate the media and the consumer with a message that describes the complexities involved. For example, "Antidepressants have shown efficacy in depression in spite of the hurdles of studying this complex illness in a clinical trial –and we can provide you with the details, suicide rates decrease in societies in proportion to use of antidepressants, and the medical community has studied and proven the ability of antidepressants to greatly improve the quality of life of millions of persons with this debilitating illness", is perhaps a more well-rounded initial comment.

Finally, both the industry and the medical profession need to make clear and concise responses to these kinds of reports. There should be a response that is written in clear wordage for the lay person, and submitted to the major media organs. A technical response for the professional should also be prepared which should be submitted to a major medical journal. An initial summary response should be placed on the internet on a new website dedicated to this response so that is comes up on a search term (ie., "antidepressants don't work"), and the lay and professional responses placed there later after publication. We should never just say, "of course antidepressants work", we must explain why. The patients who need the care, and the medical professionals and industry teams who put in the sweat to deliver this care deserve this much.