Studies Weigh Antidepressants Against Herbs and Placebos

January 01, 2003 | Major Depressive Disorder [1]

Two new studies have raised questions about the placebo effect in clinical trials. Issues of methodology must be examined and the changing trends of placebo responses need to be evaluated.

In a study sponsored by the National Institute of Mental Health and published in *JAMA*, St. John's wort (*Hypericum perforatum*) was found to be no more effective than placebo for treating depression but to be no worse than an antidepressant comparator (Hypericum Depression Trial Study Group [HDTSG], 2002). Critics of the study attributed these results to insufficient assay sensitivity; while those critical of antidepressant pharmacotherapy found the results supported their view that medication contributes little more than side effects to depression treatment.

The study, co-sponsored by the National Center for Complementary and Alternative Medicine, followed another U.S. trial that had equated the efficacy of hypericum with that of placebo (Shelton et al., 2001). That head-to-head comparison of herb and placebo was funded in part by Pfizer Inc., manufacturer of the selective serotonin reuptake inhibitor sertraline (Zoloft) that was also used as the antidepressant comparator in the NIMH study.

It had been anticipated that the HDTSG study would resolve questions about the efficacy of hypericum shown in smaller, less well-controlled European studies and in a meta-analysis of those trials (Linde et al., 1996). It was not anticipated that the study would heighten debate about the measure and efficacy of antidepressants approved by the U.S. Food and Drug Administration, but the equivocal results emerging from this growing debate prompted the editors of *JAMA* to juxtapose the study with a literature review of placebo response in depression treatment trials (Walsh et al., 2002) and an editorial on the necessity of placebo control (Kupfer and Frank, 2002a).

In the eight-week, double-blind, randomized comparison of hypericum to placebo and sertraline, 31.9% of patients responded fully to placebo, compared to 24.8% responding to sertraline and 23.9% to hypericum (HDTSG, 2002). Full response was defined as achieving a Clinical Global Impression Scale-Improvement (CGI-I) score of 1 (very much improved) or 2 (much improved) and a Hamilton Rating Scale for Depression (HAM-D-17) score of ≤8. Sertraline was ranked better than placebo on the CGI-I, but neither the SSRI nor hypericum statistically separated from placebo in the HAM-D measure.

Critics argued that the study methodology was faulty and that selecting patients with long-standing depression and a score of at least 20 (moderate to severe) on the HAM-D-17 could obfuscate beneficial effects of the herb for patients with less severe illness (Cott and Wisner, 2002; Jonas, 2002; Linde et al., 2002; Wheatley, 2002). While the acuity of illness in this studied population may leave opportunities to evaluate hypericum in patients with less severe symptoms, the similarity of placebo and drug effect could foster additional questions about antidepressant effectiveness. In particular, this resonates with questions about SSRI effectiveness for severe depression posed over a decade earlier by the Danish University Antidepressant Group (1990). In addition, these results heightened consternation with the apparent potency of placebo for depression study populations. Accompanying the HDTSG hypericum study was a literature review that found different levels of placebo response in depression studies over a period of two decades (Walsh et al., 2002). The review evaluated 75 placebo-controlled trials of treatment for adult outpatients with major depression that were conducted between 1981 and 2000. Most of the studies involved head-to-head comparison of active agents. Taking the active medication group with the greatest response, they reported that the mean proportion of patients responding to medication was 50.1% (range=31.6% to 70.4%); compared to 29.7% (range=12.5% to 51.8%) responding to placebo.

Over the two decades, the responsiveness to both active medication and placebo increased statistically significantly, but with a significantly greater increase with placebo than with medication. Walsh and colleagues (2002) indicated that this pattern of response to placebo is "highly variable and often substantial, and has increased in recent years, as has the response to medication."

Although unable to identify causal factors for the phenomenon, they suspected there to be relevant differences between participants in recent studies who may be more likely to be recruited from the
public, and those in the past who were typically referred by clinicians. "It is likely that clinically important characteristics of patients participating in treatment studies have changed as a result of these practices," Walsh and colleagues indicated.

In an accompanying editorial, Kupfer and Frank (2002a) emphasized the value of incorporating placebo control in depression treatment studies. They noted that an absence of placebo in the HDTSG trial might have led to the conclusion that hypericum is comparable to sertraline in the treatment of depression, rather than to issues of study assay sensitivity. They acknowledged, however, that placebo response in clinical trials remains problematic and requires further assessment.

"Taken together, the two reports return full circle to the placebo response and understanding its mechanism of action and highlight the perplexing complexity of the placebo and its ability to cause 'mischief' in scientific inquiry," Kupfer and Frank observed.

Schneider and Small (2002) have found that depression remission rates with active medications and placebo differed significantly between clinical sites within multisite studies, including those reviewed by Walsh et al. (2002). Among 29 studies, Schneider and Small found that remission rates at individual sites ranged from 0% to 71% with active medication and from 0% to 75% with placebo. Private research sites showed greater rates of remission with both medication and placebo than did academic sites, with more than twice the drug-placebo effect sizes. In addition to considering individual patient characteristics to understand changes in placebo response, as Walsh and colleagues (2002) suggested, Schneider and Small (2002) recommended discerning differences between investigators and investigational sites. They noted that most depression studies utilize the interview-based HAM-D and/or unstructured CGI for investigators' assessments of improvement. Both of these can reflect interaction between interviewer and subject.

"Individual investigators, despite participating in the same clinical trial and following the same protocols, might have vastly different clinical experiences," they observed. "Subtly different methods of subject selection, perceived financial incentives, unaddressed bias, and unique characteristics of individual sites may contribute to response variability" (Schneider and Small, 2002).

In their letter on the HDTSG hypericum study, Cott and Wisner (2002) suggested that the study could serve as a role model for the difficulties in carrying out antidepressant trials in which the average failure rate is 50%. In a letter separate from their editorial, Kupfer and Frank (2002b) advocated for trial designs that emphasize clinical rather than statistical significance, albeit with adequate assay sensitivity and effect size.

Kupfer and Frank suggested that clinical trial investigators obtain expert consensus on the level of illness severity and the therapeutic outcomes to assess, in order to provide the study treatment with "a fair opportunity to demonstrate its true therapeutic potential."

In July, a study on placebo trials and related commentaries were published in Prevention & Treatment (Brown, 2002; Kirsch et al., 2002; Thase, 2002). The Kirsch et al. (2002) study, a meta-analysis of 38 placebo-controlled studies of six antidepressants reviewed by the FDA between 1987 and 1999, found that while patients receiving antidepressants achieved a mean 10-point reduction in HAM-D score, patients on placebo did almost as well, achieving an eight-point reduction. Kirsch et al. argued that the antidepressants may only be responsible for the two-point increment over placebo, rather than the 10-point therapeutic response, and so their effect could be characterized as clinically negligible. This view was shared in some of the commentaries but countered in others (Brown, 2002; Thase, 2002).

Brown (2002) commented, "Placebo is better at alleviating depression than it is at preventing relapse." He also noted that antidepressant effects have been fairly consistent across studies, while placebo effects have varied more depending on patient populations and the severity of illness. Thase (2002) observed that the Kirsch et al. findings are not revelatory and that similar assessments were made in early studies of fluoxetine (Prozac) almost a decade earlier (Greenberg et al., 1994). Thase attributed the high placebo response and small relative effects of active antidepressants principally to the study populations with moderate, rather than severe, illness.

"Another problem is that industry has been slow to adapt to the knowledge of relatively small expected effect sizes and continues to conduct conventional underpowered studies," Thase commented.

References: References

Studies Weigh Antidepressants Against Herbs and Placebos
Published on Psychiatric Times
(http://www.psychiatrictimes.com)

Sept. 4.
5. HDTSG (2002), Effects of Hypericum perforatum (St. John's wort) in major depressive disorder: a randomized controlled trial. JAMA 287(14):1807-1814 [see comment].

Source URL:

Links: