

6. Di Nardo PA, Brown TA, Barlow DH: Anxiety Disorders Interview Schedule for DSM-IV: Lifetime Version (ADIS-IV-L). New York, Graywind Publications, 1995
7. Kraemer HC, Wilson GT, Fairburn CG, Agras WS: Mediators and moderators of treatment effects in randomized clinical trials. *Arch Gen Psychiatry* 2002; 59:877–883

BARBARA MILROD, M.D.
ANDREW C. LEON, Ph.D.
M. KATHERINE SHEAR, M.D.
New York, N.Y.

Sertraline and the Cheshire Cat in Geriatric Depression

TO THE EDITOR: The study by Lon S. Schneider, M.D., and associates (1) on the treatment of geriatric depression with sertraline does not rank among the glories of clinical research. It does raise questions about corporate influence and Orwellian “newspeak” in reporting clinical trials.

The study is remarkable first for its size, determined a priori by a power analysis. The aim was to achieve power sufficient to detect a mean difference of 2 points in change scores on the 17-item Hamilton Depression Rating Scale. With a projected pooled standard deviation of 8 points, this difference would represent an effect size of only 0.25. Based on past trials, a group of 700 patients was deemed necessary. The group finally enrolled numbered 747, a stunning instance of excess to answer the straightforward question of whether sertraline is superior to placebo, especially considering the low bar that the drug was asked to clear. The study has all the hallmarks of an “experimercial,” a cost-is-no-object exercise driven by a corporate sponsor to create positive publicity for its product in a market niche.

The authors concluded that sertraline is superior to placebo. The difference in mean Hamilton depression scale change score in the key intent-to-treat group was 0.8 points, less than half the stated goal. This clinically trivial difference achieved statistical significance by virtue of the gargantuan group size and because the pooled variance was less than the authors had assumed in the preliminary power analysis. “Statistically significant” differences on other dimensional primary outcome measures were likewise clinically trivial. Somewhat more encouraging data were obtained for the “completer” group, but with 131 fewer patients, that group was not representative of the drug’s performance in clinical settings. Completer data are no longer accepted as evidence of efficacy.

In the intent-to-treat group, the authors further reported a “statistically significant” advantage for sertraline in a categorical measure of response, defined as a 50% reduction of Hamilton depression scale score (35% response rate for sertraline and 26% for placebo). This difference is also clinically trivial. It translates to a number needed to treat of 11. This means that clinicians would have to use sertraline 11 times to obtain one response that would not have occurred anyway with placebo (2). In an earlier time, when antidepressant drugs first were developed, the drug-placebo difference in response rates averaged 30%–35% (3, 4), based on a number needed to treat of about three. Clearly, as reflected in this trial and elsewhere, there has been much “dumbing down” of expectations for antidepressant efficacy in recent years.

And where, by the way, are the data on remission? There is currently wide agreement that remission is the optimal indicator of antidepressant efficacy (5). The authors withheld remission data. When challenged, they will doubtless use the procedural rationalization that remission was not specified a priori as an outcome measure. The question must be, why not? By this fig leaf they conceal clinically relevant data that would probably reflect poorly on the putative efficacy of sertraline. This technique allows the authors to present their results with the best “spin.” Thus does the corporate mandate to put lipstick on the pig prevail over the academic duty to communicate independent analyses of the data (6–8). The *Journal* is complicit in this scientific failure.

The authors also failed to emphasize in the abstract (where most readers would notice it) that none of the functional or quality-of-life outcome measures favored sertraline over placebo. Something has changed in our conceptual paradigm when a drug can be described as “effective” for depression, but the patients do not confirm that their lives are any better with respect to vitality, social functioning, emotional role functioning, or mental health. Like the Cheshire cat’s smile, the only evidence that sertraline was there is the disembodied p value, grinning in statistical space, with no connection to clinical reality. That is not quite what Percy Bridgman had in mind when he introduced operationalism in science. Lewis Carroll, on the other hand, would have appreciated the irony.

References

1. Schneider LS, Nelson JC, Clary CM, Newhouse P, Krishnan KRR, Shiovitz T, Weihs K (Sertraline Elderly Depression Study Group): An 8-week multicenter, parallel-group, double-blind, placebo-controlled study of sertraline in elderly outpatients with major depression. *Am J Psychiatry* 2003; 160:1277–1285
2. Laupacis A, Sackett DL, Roberts RS: An assessment of clinically useful measures of the consequences of treatment. *N Engl J Med* 1988; 318:1728–1733
3. Klein DF, Davis JM: Review of mood stabilizing drug literature, in *Diagnosis and Drug Treatment of Psychiatric Disorders*. Edited by Klein DF, Davis JM. Baltimore, Williams & Wilkins, 1969, pp 187–298
4. Klerman GL, Cole JO: Clinical pharmacology of imipramine and related antidepressant compounds. *Pharmacol Rev* 1965; 17: 101–141
5. Keller MB: Past, present and future directions for defining optimal treatment outcome in depression: remission and beyond. *JAMA* 2003; 289:3152–3160
6. Angell M: Is academic medicine for sale? *N Engl J Med* 2000; 342:1516–1518
7. Davidoff F, DeAngelis CD, Drazen JM, Hoey J, Hojgaard L, Horton R, Kotzin S, Nicholls MG, Nylenna M, Overbeke AJPM, Sox HC, Van Der Weyden MB, Wilkes MS: Sponsorship, authorship and accountability. *N Engl J Med* 2001; 345:825–827
8. Greenberg DS: Conference deplores corporate influence on academic science. *Lancet* 2003; 362:302–303

BERNARD J. CARROLL, M.B., B.S., Ph.D., F.R.C.Psych.
Carmel, Calif.

Dr. Schneider and Colleagues Reply

TO THE EDITOR: Dr. Carroll’s essential complaint seems to be that there was no reason to perform this trial but to “create positive publicity” for “niche” marketing. He elaborates with sarcasm and hyperbole that 1) statistical significance was achieved as a product of an excessively large group size; 2) the