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

effects of sertraline were "trivial," not clinically significant; 3) we used—he says—"newspeak" and p values "disembodied" from the underlying statistics in order to confuse readers; and 4) we acted unethically in concealing data and in reporting results. These assertions are ill-informed and without foundation, and we reject them.

This trial makes important contributions to clinical pharmacological research in late-life depression, it provides relevant information about the likely effects of selective serotonin reuptake inhibitors (SSRIs) that clinicians can evaluate, and there are no particular controversies to it.

The trial was not "oversized" and not planned to reveal "trivial" differences. As we described, the determination of group size was based on the results of a previous large placebo-controlled SSRI trial in late-life depression (1), the expectation that outcomes in clinically heterogeneous elderly depressed populations with extensive medical comorbidity would be themselves heterogeneous and modest on average, and the ability to assess potential moderators such as melancholia or anxiety. With that exception, other placebo-controlled trials in late-life depression have been underpowered and undersized. The consequences of underpowered trials are that they tend to yield noninformative results and type II errors. Conversely, when results are statistically significant, it is because the effect sizes are implausibly large. In some instances, results from smaller trials have not been published simply because they are negative. Most experts would consider an adequately powered trial of typical clinical patients and outcomes generalizable enough to inform clinical practice as a distinct strength and not a "scientific failure."

Contrary to his assertions, the statistics and outcomes in this report are clearly described and understandable. No "disembodied p values" were reported; every p value was explicitly connected to an outcome parameter and a statistical test. Any reader could assess the baseline characteristics of the population and the magnitudes of differences and calculate effect sizes of outcomes—just as Dr. Carroll did himself. Moreover, if the trial had been underpowered and undersized, he would not have been able to calculate an interpretable number-needed-to-treat statistic because the confidence interval (CI) would have been so broad as to be uninformative.

It is inappropriate and misleading for Dr. Carroll to compare this geriatric depression outpatient trial to the earliest imipramine trials performed around 1960 in younger adults (Klerman and Cole, 1965) in order to support his assertion that sertraline has a "trivial" effect. These trials, landmarks as they were half a century ago, were seriously deficient in nearly all areas. They used inexplicit diagnostic and inclusion criteria (e.g., mixing inpatients and outpatients, psychotic and neurotic depression, schizophrenia and mania) and methods for dosing and maintaining the blind or placebo control (e.g., many used atropine and thiopental as "placebos"). Outcomes assessments were idiosyncratic, and dropouts were not accounted for; most were so small, averaging about 60 to 70 patients, that they were not statistically significant individually.

Subsequent antidepressant trials, those from the 1980s and 1990s that used modern diagnostic criteria, rigorous methods, and specified outcomes, and modern evidence-based reviews based on these trials (2) demonstrated a relative benefit of antidepressant response over placebo of 1.6 (95% CI=1.5–1.7) in primarily young and middle-age adults. By comparison, we found a relative benefit for sertraline of 1.4 (95% CI=1.1–1.7). This effect is hardly trivial. Similarly, although number-needed-to-treat statistics from these studies are larger than what Dr. Carroll calculated, they are not statistically significantly so. The relative benefit (or relative risk) is an effect size measure that accounts for placebo response, something that a number-needed-to-treat statistic cannot (Laupacis et al., 1988).

The relevant comparison to make, however, is to the few other placebo-controlled antidepressant trials in late-life depression. Here, the relative benefit is 1.4 (95% CI=1.2–1.6) (2), nearly identical to our finding. We discussed that the effects of sertraline were modest, nearly identical to a similarly sized trial of fluoxetine (1) and suggested that the two trials probably represent best estimates of the treatment effects of SSRIs in outpatients with late-life depression. We submit that this trial is informative of what likely treatment effects are in elderly outpatients over the short term and, unlike some research, will be more enduring and of practical clinical consequence.

Dr. Carroll goes on to fault us for not providing—or worse—"withholding" or "concealing" what he calls "remission" data, presumably based on cutoff scores on outcome instruments, in order to best "spin" the results. The use of such cutoff scores on continuous or ordinal data is clearly unsatisfactory, especially in elderly groups, where there are substantial somatic and residual depression-like symptoms among both depressed and nondepressed individuals (3, 4). In fact, we used standard definitions of a clinically meaningful response, a 50% reduction in baseline Hamilton depression scale scores and, separately, a Clinical Global Impression Scale (CGI) improvement score of 1 or 2 (i.e., markedly or moderately improved). Moreover, we reported that the CGI response of 1 or 2 had to be sustained throughout the remainder of the trial.

Nevertheless, at his request, we calculated "remission" rates, defined as an endpoint Hamilton depression rating scale score ≤10 and a CGI severity score of 1 or 2 (borderline ill or not ill at all). Remission rates on the Hamilton depression rating scale were 34.6% versus 26.6% (Cochran-Mantel-Haenszel χ^2=5.61, df=1, p<0.02), and remission rates on the CGI severity scale were 32.8% versus 22.8% (Cochran-Mantel-Haenszel χ^2=9.43, df=1, p=0.002), respectively, for sertraline versus placebo. The risk difference, or number-needed-to-treat statistic, and the relative benefit of 1.44 are virtually identical, and the absolute rates are similar to the categorical responses we reported for the Hamilton depression rating scale score (35% versus 26%) and the CGI scale score improvement (45% versus 35%).

Abstracts do not substitute for complete reports and do not contain all results. Dr. Carroll would have put quality-of-life scores in the abstract, arguing that most readers would read only the abstract, and says that we should highlight here that patients could not appreciate any effect. He does not similarly fault us for omitting from the abstract the patients' self-assessed global impression of improvement, which strongly favored sertraline. Contrary to his assertion, the patients, in fact, endorsed their own improvements and with an effect size that was larger than the clinicians' assessments.

In sum, we reject Dr. Carroll's assertion that we put aside scientific and public health considerations to write an article under corporate influence to gain a marketing niche. Con-

trary to his assertion, we presented the whole Cheshire cat: face, ears, and tail. We regret that Dr. Carroll cannot offer his points more collegially or professionally.

References

1. Tollefson GD, Bosomworth JC, Heiligenstein JH, Potvin JH, Holman S (Fluoxetine Collaborative Study Group): A double-blind, placebo-controlled clinical trial of fluoxetine in geriatric patients with major depression. Int Psychogeriatr 1995; 7:89–104
2. Williams JW, Mulrow CD, Chiquette E, Noel PH, Aguilar C, Cornell J: A systematic review of newer pharmacotherapies for depression in adults: evidence report summary. Ann Intern Med 2000; 132:743–756
3. Thomas L, Mulsant BH, Solano FX, Black AM, Bensasi S, Flynn T, Harman JS, Rollman BL, Post EP, Pollock BG, Reynolds CF: Response speed and rate of remission in primary and specialty care of elderly patients with depression. Am J Geriatr Psychiatry 2002; 10:583–591
4. Roose SP, Sackeim HA, Krishnaan KRR, Pollock BG, Alexopoulos G, Lavretsky H, Katz IR, Hakkarainen H, the Old-Old Depression Study Group: Antidepressant pharmacotherapy in the treatment of depression in the very old: a randomized, placebo-controlled trial. Am J Psychiatry (in press).

LON S. SCHNEIDER, M.D.
J. CRAIG NELSON, M.D.
CATHRYN M. CLARY, M.D.
PAUL NEWHOUSE, M.D.
K. RANGA RAMA KRISHNAN, M.D.
TOM SHIOVITZ, M.D.
KAREN WEIHS, M.D.
Los Angeles, Calif.

Polypharmacy in Psychiatric Inpatient Treatment

TO THE EDITOR: Franca Centorrino, M.D., and associates (1) compared the use of antipsychotics in psychiatric inpatients using data from 1989, 1993, and 1998. They found that the proportion of days the patients had received more than one antipsychotic during inpatient treatment had increased from 1.7% in 1989 to 20% in 1998. The most common combinations were typical antipsychotics added to atypical primary agents, and the authors suggested that this might reflect incomplete confidence in the effectiveness of monotherapy with atypical agents. Unfortunately, the authors did not report on other medications besides antipsychotics, and I wonder if a substantial increase had also occurred for these.

As the authors stated, polypharmacotherapy is a growing international phenomenon, and incomplete trust in the effectiveness of atypical antipsychotics obviously is just one of many reasons fostering polypharmacy. I recently reviewed the available literature on the number of psychotropic drugs administered during inpatient treatment (2) and found that the proportion of patients (including all diagnoses) being treated with monotherapy has declined significantly during the last few decades. Studies originating in 1980 or before reported monotherapy in 48%, studies between 1981 and 1990 in 31%, and studies between 1991 and 2000 in 20%. Despite all caveats concerning the small database of available studies, there is little doubt that a powerful trend toward polypharmacy is operating. The reasons for this are certainly quite complex, as follows:

1. A more sophisticated diagnostic process leading to diagnoses of multiple comorbid conditions makes more treatments necessary.

2. There are far more drugs available, both new and old, in new indications, and all are intensely promoted by the pharmaceutical industry.

3. Inpatient treatment has to deal with the most severe and often therapy-resistant cases, for which an increasing number of combination and augmentation therapies have been recommended and are widely used in spite of little empirical evidence.

4. A decreasing number of psychiatric beds and decreasing lengths of stay of inpatient treatment add even more pressure to strive for the most effective treatment.

Psychiatrists have to be aware that their clinical practice is far from evidence based. Two conclusions are important. First, clinicians should monitor the trend toward polypharmacy in their treatment regimens extremely critically. Second, we need studies investigating at least those combinations of drugs that are most widely used, e.g., the combination of atypical and typical antipsychotics.

References

1. Centorrino F, Eakin M, Bahk W-M, Kelleher JP, Goren J, Salvatore P, Egli S, Baldessarini RJ: Inpatient antipsychotic drug use in 1998, 1993, and 1989. Am J Psychiatry 2002; 159:1932–1935
2. Rittmannsberger H: The use of drug monotherapy in psychiatric inpatient treatment. Prog Neuropsychopharmacol 2002; 26:547–551

H. RITTMANNSBERGER, M.D.
Linz, Austria

Dr. Centorrino Replies

TO THE EDITOR: We agree that polypharmacy, including treatment with multiple psychotropic medications not limited to antipsychotics, is a major concern in the field of psychiatry today and one that warrants both consideration and further study. While our focus in our article was primarily on the use of antipsychotic medication, we are writing a second report using the same subject group that examines combination therapy in particular and includes information on combination psychotropic medication in general. The results presented in this report will highlight the increasing prevalence of combination therapy and compare the possible effects of combination versus monotherapy in factors such as length of inpatient stay, clinical status, and side effects. We maintain, however, that further study into the use and outcome of polypharmacy is necessary.

FRANCA CENTORRINO, M.D.
Belmont, Mass.

Fertility and Schizophrenia

TO THE EDITOR: Schizophrenia, a disease with a strong genetic component, has not disappeared, despite the fact that affected patients have lower fertility than the general population. Jari Haukka, Ph.D., et al. (1) tried to explain this apparent paradox by testing the hypothesis that the relatives of schizophrenia patients have higher fertility than the general population. Not surprisingly, the study did not confirm this hypothesis.