parison of lithium and T_3 augmentation following two failed medication treatments for depression: a STAR*D report. Am J Psychiatry 2006; 163:1519–1530

SHELDON HEATH, M.D.
Toronto, Canada

Dr. Heath reports no competing interests.

STAR*D Level IV Methodology

TO THE EDITOR: The article by Patrick J. McGrath, M.D., et al. (1) based on STAR*D trials was an important randomized trial of refractory depression. We would like to address two methodological issues in the analysis, however, which may impact its clinical implications.

First, according to Table 1, 41.4% of the patients randomized to the tranylcypromine group entered this study because of previous medication intolerance with other STAR*D trials, while only 21.6% of the patients randomized to the venlafaxine/mirtazapine group had previous medication intolerance. This imbalance, which was not adjusted in the statistical analysis, might have biased the results against tranylcypromine in the tolerability measure, casting doubts on the finding that participants taking tranylcypromine were more likely to exit the study because of side effects.

Second, the primary outcome of remission, based on the Hamilton Depression Rating Scale, was similar in both groups and thus a negative result. However, the secondary outcome of treatment response, based on the Quick Inventory of Depressive Symptomatology-Clinician-Rated, was 12.1% for tranylcypromine versus 23.5% for venlafaxine/mirtazapine. The article downplayed this difference by stating that "response rates were also low and not significantly different" (1, p. 1535). Hypothesis testing methods are most appropriate when the study is designed and powered to test the hypothesis; otherwise, the most appropriate statistics for secondary outcomes and other post hoc analyses are effect estimates and confidence intervals (2, 3). The p value for assessing significance in this comparison is less relevant because the study was powered to assess the primary outcome measure of remission rate, not response rate (3). Hence, type II error is important: when assessing outcomes that the study was not powered to assess, one cannot equate statistical nonsignificance with mathematical nondifference or clinical nonsignificance (4). We recalculated the relative risk between the two groups, which was 1.95, with a 95% confidence interval of 0.83–4.57. Although the null is included, these results are most consistent with a probability of an effect. The most likely clinical conclusion, pending further studies, could thus be that the venlafaxine/mirtazapine combination may have moderate benefits over tranylcypromine based on response rate.

References

1. McGrath PJ, Stewart JW, Fava M, Trivedi MH, Wisniewski SR, Nierenberg AA, Thase ME, Davis L, Biggs MM, Shores-Wilson K, Luther JF, Niederehe G, Warden D, Rush AJ, STAR*D Study Team: Tranylcypromine versus venlafaxine plus mirtazapine following three failed antidepressant medication trials for depression: a STAR*D report. Am J Psychiatry 2006; 163:1531–1541
2. Rothman KJ: A show of confidence. N Engl J Med 1978; 299: 1362–1363
3. Goodman SN: Toward evidence-based medical statistics, I: the p value fallacy. Ann Intern Med 1999; 130:995–1004
4. Freiman JA, Chalmers TC, Smith H Jr, Kuebler RR: The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial: survey of 71 "negative" trials. N Engl J Med 1978; 299:690–694

ALIZA P. WINGO, M.D.
S. NASSIR GHAEMI, M.D., M.P.H.
Atlanta, Ga.

Dr. Ghaemi currently receives research grants from GlaxoSmithKline and Pfizer, has previously served on the speakers' bureaus of GlaxoSmithKline, AstraZeneca, and Abbott Laboratories, and has previously served on the advisory boards of GlaxoSmithKline, Janssen, Pfizer, and Abbott Laboratories. Dr. Wingo reports no competing interests.

Dr. McGrath Replies on Behalf of the STAR*D Trial Team

TO THE EDITOR: Drs. Wingo and Ghaemi raise two methodological points regarding our article describing the results from Level 4 treatment in the STAR*D. The first point is that adjusting for previous medication intolerance, which differed between the tranylcypromine and venlafaxine/mirtazapine groups, would be important to estimate comparative effects of the treatments. We agree, and in fact did adjust for this in the analyses we presented, as is noted in footnote b to Table 3.

The second point suggests that we downplayed a numerical difference in the response rate between the two treatments that was not statistically significant, essentially endorsing a lack of difference in treatments possibly because of a type II error. Drs. Wingo and Ghaemi are correct that our primary endpoint of remission had limited power to detect small effect sizes. However, we chose to emphasize the primary endpoint of remission because it was specified a priori, which limits the possibility of chance findings to the specified significance rate, which is not the case if multiple outcome measures are analyzed separately. They assert that the nonsignificant difference on the secondary measure of response is "most consistent with a high probability of an effect." We disagree—the probability of an effect in this data is clearly not "high" because the probability is most correctly described by the probability of a type I error, that is, the p value. Whether a larger study might find these numerical differences to be significantly different remains speculative.

We believe that Drs. Wingo and Ghaemi neglect another more likely explanation of any putative difference between treatments, which is that poorer tolerability of tranylcypromine in this study resulted in shorter durations of treatment and less chance to show improvement. However, even accepting their interpretation of the response rates leads to the same conclusion stated in our article, perhaps with the addition noted in brackets: "The lower side effect burden, lack of dietary restrictions, [numerical though not significant advantage in response rate], and ease of use of venlafaxine and mirtazapine suggest that this combination may be preferred over tranylcypromine for patients with highly-resistant depression who have not benefited adequately from several prior treatments."

STAR*D Trial Team

Author disclosures accompany the original article.