Failed Studies Should Not Be Used to Malign Good Treatments

To the Editor: The recent double-blind, randomized, placebo-controlled trial of S-adenosylmethionine (SAMe) versus escitalopram in major depressive disorder (MDD) was reported as a failed trial by Mischoulon and colleagues. In intent-to-treat samples, no significant differences were found in response rates: 36% for SAMe, 34% for escitalopram, and 30% for placebo. The remission rates of 28% for SAMe, 28% for escitalopram, and 17% for placebo suggested that both active treatments produced “a more robust ‘true’ effect” compared to placebo.

Mischoulon and colleagues discuss possible reasons for the failure of antidepressants with established records of efficacy (SAMe and escitalopram) to exceed placebo. They note that Iovieno and Papakostas observed that placebo response rates ≥ 30% correlated with lower risk ratio of response to antidepressants versus placebo. Rutherford and Roose also reported worse performance of drug versus placebo in studies with placebo response rates ≥ 30%, as seen in this Mischoulon study. Nearly 60% of subjects had low levels of depression severity (pretreatment Hamilton Depression Rating Scale score ≥ 19), increasing the difficulty of demonstrating a significant drug versus placebo effect.

Despite the fact that the SAMe response rate was equal to that of escitalopram, discussion of the results gave unequal treatment to the 2 antidepressants. Mischoulon and colleagues state that the sample size was only two-thirds the number for which the study was powered, a “major limitation”; however, they write that the sample “is large enough to provide a conclusive statement about the efficacy of SAMe as a monotherapy for MDD” [our emphasis]. How can one make a conclusive statement from a failed trial? No such statement was made about escitalopram, though it performed no better than SAMe. They speculate, “SAMe may be better suited as an augmentation therapy than as a monotherapy.” Bias is introduced and then is driven home by suggestions that further trials will be needed to clarify SAMe’s potential antidepressant effect and its place in treating depression. According to the US Department of Health and Human Services Agency for Healthcare Research and Quality 2002 report, SAMe has already proven efficacy as a monotherapy in depression, based on 13 randomized, placebo-controlled trials and 19 randomized trials comparing it with standard antidepressants (imipramine, amitriptyline, clomipramine, nomifensine, minaprine, and desipramine). The manner in which researchers discuss study results impacts clinical practice and can be particularly damaging in the case of less well-known or less conventional treatments. Publications must be impartial in order to avoid introducing errors of bias into clinical decision making.

This study highlights the growing threats to validity in large, expensive antidepressant trials that appear to have excellent methodology (eg, high Jadad scores), and yet produce no useful information. The answers would more likely be found in problems with the patient selection process and the heterogeneity within patient groups with respect to genomic and metabolic biomarkers. At the very least, until the true causes of invalidity are identified, such studies should not be used to create unwarranted doubt about the efficacy of highly beneficial treatments such as SAMe.

References

Dr Mischoulon and Colleagues Reply

To the Editor: We thank Drs Gerbarg and colleagues for their interest in our recent publication, “A double-blind, randomized, placebo-controlled clinical trial of S-adenosyl-l-methionine (SAMe) versus escitalopram in major depressive disorder.” Dr Gerbarg et al expressed concern about the statement that our study sample, while smaller than expected, was "large enough to provide a conclusive statement" about the efficacy of SAMe monotherapy in major depressive disorder. We take this opportunity to address these concerns and further expand on and clarify our statement.

The study, as we pointed out, was originally powered for a sample of 300 patients. Statistical projections at the time of study closure led us to conclude that even if we had been able to recruit the full complement of patients per the original study design, the findings would not have been significantly altered, i.e., neither of the active treatments would have separated from placebo. In essence, this was a futility analysis (such analyses are usually conducted as interim analyses to determine whether a trial should be continued). Inadequate power may explain a negative study, but less so a “failed” study, particularly one such as ours that had relatively low response rates to the active treatments, coupled with a robust placebo response rate. Because many trials of complementary and alternative medicine (CAM) therapies have been limited by small samples, we wanted to emphasize that our findings should not be attributed to the sample size. Furthermore, the study was carried out by 2 teams of seasoned investigators with lengthy track records of successful execution of clinical trials of various therapies for major depression, including CAM. There was therefore no reason to think that the findings could be explained by inexperience in
the clinicians who assessed the patients and executed the study procedures.

Our conclusions reflected what we observed: patients in our sample who experienced significant alleviation of depression did so independently of which treatment they were assigned. This in itself must suggest the question of whether the apparent beneficial effects of SAMe, as well as of escitalopram, were due to a placebo effect, though we acknowledge that the latter has considerably more established efficacy based on large datasets. In our Discussion section, we put forth a number of possible explanations for the unexpected findings, but did not claim to have shown that SAMe monotherapy was ineffective for depression. Indeed, we acknowledged the stronger remission rates for SAMe and escitalopram as supportive of their efficacy, and also cited our recent encouraging findings for SAMe augmentation therapy.2

A single study is never definitive and must always be carefully interpreted in the context of the literature as a whole. In this regard, our colleagues cite a 2002 report from the Agency for Healthcare Research and Quality (AHRQ) of the US Department of Health and Human Services,3 stating that “SAMe has already proven efficacy comparable to prescription antidepressants as well as in many placebo controlled trials.” This statement, while a reasonably accurate interpretation of the results of the AHRQ meta-analysis, does not preclude the fact that, relative to most antidepressant medications that comprise current standard of care, the evidence base for efficacy of SAMe has been limited by studies with small samples and with other methodological issues.4 The AHRQ report, in fact, concluded that the statistically significant effect of SAMe versus placebo was “equivalent to a partial response to treatment”5(p2) and that “too few studies were available”5(p2) to calculate a risk ratio for categorical response (50% decrease on HDRS). The report recommended that “once efficacy of the most effective oral dose of SAMe has been demonstrated, larger clinical trials are indicated for the use of SAMe for depression, osteoarthritis, and cholestasis. Such trials would need to enroll large numbers of patients with homogeneous diagnoses, and focus on significant clinical outcomes. Ideally, they would compare SAMe to both placebo and standard care.”5(p3) Our study, originally funded in 2004 by the National Institutes of Health and the National Center for Complementary and Alternative Medicine, was designed to advance the field of scientific inquiry into antidepressant effects of SAMe as one of the largest investigations of SAMe monotherapy, and the first comparison with an SSRI.

Members of our research team, like some of the authors of the critique, have received different forms of support from companies that manufacture and market SAMe. For this reason, we felt it was especially important that we present our findings as objectively as possible, without appearing to “spin” the results in a direction that would seem to favor SAMe. The cautionary tone of our article reflected this desire. We continue to prescribe SAMe in our practices and have seen many patients benefit. Our report was not intended to discourage clinicians from recommending SAMe to appropriate patients, but simply to emphasize that more work needs to be done before any firm recommendations can be made.

Finally, we strongly agree with our colleagues’ closing statement that potential explanations for our findings may lie in heterogeneity within patient groups. Various subinvestigations into biological, site-related, and gender-related factors and their potential impact on the findings of this study are either in progress or in press.5