controlled studies of the treatment of delirium, and this study may be misinterpreted by those who read only the abstract. Delirium needs to be detected early and treated vigorously, despite the conclusions of this study.

Stephen D. Anderson, MD, FRCPC
Clinical assistant professor
Department of Psychiatry
Faculty of Medicine
University of British Columbia
Vancouver, BC

Reference

1. Cassem NH, Hackett TP: *Massachusetts General Hospital Handbook of General Psychiatry*, 2nd ed, PSG Publishing, Littleton, Mass, 1987: 368–371

Cole and associates are to be commended on the first systematic study of interventions designed to reverse or resolve delirium in elderly inpatients. The methods used were reasonable in evaluating the effectiveness of treating delirium.

I am not a statistician and so I wish to ask the authors about the size of their sample.[1,2] In addition, I challenge the authors' choice of $p < 0.05$ as the level of statistical significance. A more conservative approach, as suggested by Hassard,[3] may have been preferable, especially in light of the multiple comparisons undertaken. The level of statistical significance could have been more appropriately revised downward by dividing the original Type I error of 0.05 by the number of statistical comparisons made.[4] For example, if five comparisons were made, the level of statistical significance would drop to 0.05/5 or 0.01.

Table 1: Sample size required to show a significant difference in score of 5 on the Crichton Geriatric Behavioural Rating Scale with a standard deviation of 9.4

Type II error	$p = 0.05$	$p = 0.01$
$\beta = 0.2$	79	111
$\beta = 0.1$	92	125

Table 2: Sample size required to show a significant difference of 1.0 in the Short Portable Mental Status Questionnaire with a standard deviation of 2.6

Type II error	$p = 0.05$	$p = 0.01$
$\beta = 0.2$	152	213
$\beta = 0.1$	176	242

For a standard deviation of 9.4 in scores of the Crichton Geriatric Behavioural Rating Scale (CGBRS) for the control group, a Type I error of 0.05 and a Type II error of 0.20 in a two-tailed test, and a clinically significant difference in score of 5, the sample should be approximately 80 patients per group (Table 1). Similarly, for a standard deviation of 2.6 in scores of the Short Portable Mental Status Questionnaire (SPMSQ), the sample needed for a clinically significant difference in score of 1 would be 150 patients per group (Table 2). Tables 1 and 2 outline the size of samples needed for different levels of Type I and Type II error. In no case was the sample used by Cole and associates sufficient to achieve clinical significance.

These methodologic issues preclude my integrating the authors' results into my practice. Their results need to be replicated several times with larger samples. None the less, I would commend the authors on a difficult task well done and would heartily recommend that they extend their study to involve more patients.

Albert J. Kirshen, MD, FRCPC
Assistant professor of internal medicine
Division of Geriatric Medicine
University of Toronto
Toronto, Ont.
Program director
Continuing Care
Baycrest Centre for Geriatric Care
North York, Ont.

References

1. Donner A: Approaches to sample size estimate in the design of clinical trials — a review. *Stats Med* 1984; 3: 199–214
2. Hassard TH: What sample size will I need? In *Understanding Biostatistics*, Mosby Year Book Publishers, St Louis, 1991: 167–182
3. Hassard TH: Analysis of variance. In *Understanding Biostatistics*, Mosby Year Book Publishers, St Louis, 1991: 75–97
4. Hassard TH: Analysis of variance. In *Understanding Biostatistics*, Mosby Year Book Publishers, St Louis, 1991: 85

[One of the authors responds:]

Dr. Kirshen's comments on the levels of statistical significance are noteworthy. We had anticipated differences between groups greater than those suggested by Kirshen (i.e., 8 to 10 points for the CGBRS and 2 to 3 points for the SPMSQ); consequently, the sample size we calculated was lower.

Dr. Anderson's six major criticisms of our study are unjustified.

First, he asserts that elderly patients with delirium should receive neuroleptic medication. Although this may be true for the agitated, delirious patients usually referred to a consultation-liaison psychiatrist, it was not true for all patients enrolled in this study, many of whom were hypoactive and in many of whom delirium would have gone undetected but for the study.

Second, Anderson states that the intervention was poorly defined but none the less unlikely to be effective in patients with moderate to severe delirium. The intervention, well described in the article, comprised two parts: a consultation by a geriatric internist or psychiatrist and follow-up by a liaison nurse who, among other activities, used a protocol in discussing management with the nurse responsible for each patient. Unfortunately, the beneficial effects of these procedures proved to be small.

Third, Anderson claims that we did not follow the recommendations of an earlier review article, namely, elimination of the confounding effects of dementia and relation of detection and management procedures to course and outcome. In fact, these recommendations applied to studies of prognosis, not intervention.

Fourth, Anderson insists that delirious patients with dementia should have been excluded from the study. We included such patients because there was no evidence they would not respond to our intervention. As it turned out, they improved less than