I would also question the authors’ conclusion that reliance on self-reports of abuse would more likely take the form of underreporting than overreporting. Although there is evidence that this may occur in a non-clinical sample (Femina et al., 1990), it is conceivable that an opposite pattern might emerge in a clinical sample, especially with the current emphasis by therapists and the media on the damaging effects of childhood sexual abuse. One must also consider the growing number of individuals who claim to have recently uncovered previously forgotten memories of abuse. Insofar as the reliability of these “recovered memories” is now being seriously questioned (e.g. Loftus, 1993), future research on this topic should regularly include memory continuity as an additional variable. Even if one assumes that recovered memories are reliable, they may be associated with a distinctly different pattern of symptoms. Moreover, as the controversy over recovered memories plays itself out, there is a need for an ongoing assessment of the frequency with which patients are reporting such memories.


RUSSELL A. POWELL

Department of Social Sciences
Grant MacEwan College
P.O. Box 1796
Edmonton, Alberta
Canada T5J 4S2

**Community psychiatric nurse teams**

Sir: Mujen et al (*BJP*, August 1994, 165, 211–217), in discussing the results of their innovative and important clinical trial of standard v. intensive care by community psychiatric nurses (CPNs), concluded that “despite the low number of contacts in the generic [standard] group and high number in the CST [intensive] group, no differences in outcome were found”. Although they interpreted these findings cautiously, some methodological considerations were omitted in their discussion of the possible explanations.

**Randomisation, confounding, and type II error rate.** Simple randomisation is a powerful method, because it controls for both known and unknown confounding factors. This procedure, however, requires a large sample size, especially if one is trying to detect relatively small effects against potentially confounding background noise. The sample size used by Mujen et al (82 at baseline) is rather small for this purpose, and consequently there remain some baseline differences between the two groups after randomisation. Therefore, it becomes vital to control adequately for confounding. The presence of some possible confounders, such as the higher proportion of CST patients living alone and being of Afro-Caribbean origin, are acknowledged. However, in spite of the already small sample size, the authors examine for possible confounding using the method of *post hoc* stratification, which results in very small subgroups. Such analyses are unlikely to be informative, break the simple randomisation, and should be avoided if a trial has limited statistical power (Pocock et al, 1987a).

The lack of power in this study is compounded by the relatively high attrition rate (nearly 25%). Absence of “statistically significant differences” should, therefore, not be equated with truly negative results. Instead, the possibility of a type II error (i.e. a false negative result) should be considered. The issue of low power can be clarified by listing the appropriate confidence intervals: a statistically insignificant mean difference with a narrow confidence interval is an informative null finding, but not a statistically insignificant mean difference with a very wide one. Such data would enable the reader to evaluate whether the negative results are genuine or due to small sample size.

Analyses. The authors specify seven end points (considered of equal importance from the statistical point of view), and conduct separate *t*-tests of mean changes in each endpoint, and further separate analyses at each of the three measurement points. Such overuse of significance testing may make any possible results difficult to interpret. Furthermore, because of their larger standard error, analysis involving the mean change between baseline and post-treatment measurements are inferior to methods using the mean baseline measurement as a covariate in a linear model for treatment comparisons of post-treatment means (Frison & Pocock, 1992). The method used will also fail to pick up the fact that, for example, there were small but consistent differences between the two groups on patient satisfaction scores (for a discussion of suitable methods see Pocock et al, 1987b; Frison & Pocock, 1992). It needs to be pointed out that each test is conducted on slightly different samples, because of
Correspondence

withdrawals and loss to follow-up. This is important because loss to follow-up in a small trial may lead to bias. Although the authors examine the issue of bias due to loss to follow-up, by providing comparison material between the group of all patients and the group with complete ratings, the reader also needs detailed information on those lost to follow-up. For example, in Table 1, the difference between the CST and generic groups in mean number of admissions and mean number of pre-hospital days increases considerably from the group of “all patients” to the group of “rated patients” (from 0.6 to 2.2, and from 0 to 23 respectively). This could indicate that loss to follow-up had differential effects on the distribution of important predictors of outcome in the two groups. Another bias might have been introduced by the fact that baseline assessment apparently took place after randomisation, with the possibility of influencing initial expectations of both patients and raters. For example, at baseline there are significant differences between the two groups on the GAS (mean difference 4.9; 95% CI 4.0–5.9) and the related PSE measure (mean difference 4.7; 95% CI 3.4–6). This demonstrates better functioning in CST patients at baseline, a difference that disappeared at subsequent measure points.

Time in hospital. Time spent in hospital, arguably the most important outcome measure of the trial, is usually skewed. In reporting the results on total time spent in hospital, the authors state that the mean time in hospital was 46 days for the CST and 44 days for the generic group. While this would indicate a similar length of stay, they also report that median length in the CST group was 21 days, but 12 days in the generic group. This would suggest that the CST group had longer admissions, which may be related to the fact that patients in the generic group were more often living in hostel care. This point needs to be clarified; if time in hospital is evidently skewed, then the comparison of mean values as presented in Table 4 is invalidated, and the authors should use the appropriate non-parametric test. The fact that for both groups the median values are much lower than the mean values does indicate that the data are positively skewed. It is therefore important that the reader be provided with the appropriate statistical information as to whether the intervention groups differed beyond chance expectation with respect to this outcome measure.

We agree with Muijen et al that the issue examined is an important one, and that one should interpret the findings cautiously. However, we feel that their assumption that the negative findings are informative is not entirely justified. Apart from the possibility of a high type II error rate, the results are difficult to interpret because of unsatisfactory adjustment for confounding and the possibility of bias.


Jim van Os
Kwame McKenzie
Karyna Gilvary
Tom Fahy

Institute of Psychiatry
De Crespigny Park
London SE5 8AF

Impact of new long-stay patients

Sir: Lelliott (BJP, August 1994, 165, 160–169 and 170–178) raises the issue of the impact of new long-stay patients (NLS) on acute psychiatric services, who occupied an average of 7% of acute beds nationally and over 10% in some services. Our own audit of prolonged admissions to acute wards in Nottingham with onset in 1990 revealed 33 patients with stays of 180 days or more, occupying 8720 bed days. Nottingham has 145 acute beds (including 12 forensic/TU) for a population of 620 000 (23 per 100 000). There are integrated community mental health services, described in part by Tyrer (1989), and a well-developed rehabilitation and community care service.

The patients with prolonged stays accounted for 17% of available occupancy, and as overall occupancy was then 76%, 22% of actual occupancy. Only two of the NLS patients remained on acute wards for more than one year (maximum stay 421 days). If patients with acute admissions of 90 days or more (Lelliott argues stays of greater than three months are “inhumane”) are included, the bed use by the long-stay group rises to 26% of available, or 33% of occupied bed days in 1990. The psychiatric and sociodemographic characteristics of the group appeared similar to that of Lelliott’s cohort.

One conclusion that could be drawn from these data is that as services become more community