Reviewer’s report

Title: An Approach to Measure Compliance to Clinical Guidelines in Psychiatric Care.

Version: 2 Date: 1 February 2008

Reviewer: James Rohrer

Reviewer’s report:

Review: BMC Psych 020108

Overall

In this study, 2 clinics actively implemented guidelines (GLs) for depression, 2 did so for suicide attempters, and 2 received the GLs but were not active implementers. This is interesting because the issue has contemporary importance. It is interesting methodologically because quality improvement studies in psychiatric services would benefit from more frequent of similar quasi-experimental designs. However, the findings must be clearly analyzed and presented, which the investigators have not done.

MAJOR COMPULSORY REVISIONS

Important points:

Did the investigators consider creating an additive process quality index for each GL rather than analyzing each indicator separately? This is not a compulsory change but would greatly simplify the presentation.

The investigators reported selection bias in that more motivated clinics volunteered to be implementation sites. This underscores the importance of controlling for baseline patient differences.

Abstract

The study is described as controlled. That is technically correct but incorrectly infers random assignment. The study is observational in nature. It is pre-test post-test with comparison group. The comparison group in non-equivalent since those sites are described as being less motivated to implement the quality improvement program. In the abstract, the authors can merely describe the study as a quasi-experiment involving observations before and after the intervention and a comparison group.

The abstract should contain numbers for each group and significance levels in the results section.

The results section of the paper reports that there were no baseline differs in gender mix between implementation and control clinics. However, it does not say
that age differences were not present. I'm afraid that suggests that all results should be age adjusted.

Methods

Hypotheses should be stated at the beginning of the methods section. For example, consider these:

H1: the odds that a patient was treated in accordance with GLs are greater after implementation than before.
H2: the odds that a patient was treated in accordance with GLs are greater in the implementation clinics than in the comparison clinics.
H3: the overall depression quality score is higher for patients treated in implementation clinics than in control clinics.
H4: the overall suicidality quality score is higher for patients treated in implementation clinics than in control clinics.

The paper would be a lot simpler if you dispensed with H1 and H2.

Inter-rater reliability was assessed but not reported in the methods section. I prefer to see the results of that assessment in the methods rather than the results because it attests to the accuracy of the instrumentation and is not part of the hypothesis testing.

Results

Assuming the above hypotheses are being tested, Table 1 should show the percentages of patients being treated in accordance with each depression indicator at baseline with columns for implementation and control groups. One row can show the mean quality scores for the groups. Significance tests should be reported to indicate how much the groups differed at baseline.

Table 2 should show the percentages of patients being treated in accordance with each suicide indicator at baseline with columns for implementation and control groups, along with one row for the mean quality scores for the groups. Significance tests should be reported to indicate how much the groups differed at baseline on each row.

Since significant baseline differences were present, Table 3 should report how the odds of compliance for each depression indicator in Time 2 differed in the implementation group, after adjusting for baseline differences. The same table can show how the mean quality differs by time period. Table 4 does the same for the suicide indicators. A graphic showing the trend lines would be very enlightening.

Discussion

The cost of this intervention is not trivial. Did other aspects of clinical care suffer because staff members spent time improving depression treatment?
Whether the intervention has a long-term impact on GL adherence is unknown. Presumably some return to baseline will occur.

A substantial amount of random fluctuation in process quality scores occurs. Therefore, the study should be replicated so as to confirm that random change and selection bias have not combined to produce a spurious result.

MINOR COMPULSORY REVISIONS

Clarifications needed:

How were the six psychiatric clinics different from each other?

Were patient records randomly selected or were they a convenience sample, perhaps of consecutive patients?

DISCRETIONARY REVISIONS

Shifting the paper toward a focus on additive process quality measures.

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

'I declare that I have no competing interests'