Reviewer's report

Title: Supported local implementation of clinical guidelines in psychiatry: A two-year follow-up

Version: 1 Date: 3 October 2008

Reviewer: Ramesh Raghavan

Reviewer's report:

The authors report on a study designed to improve certain administrative characteristics, procedural improvements in care, and consequent improvements in outcomes among recipients of care in 6 mental health facilities in an urban county in Sweden. There are several grounds for potential enthusiasm for this study – the authors conduct a 2-year follow-up of various outcomes, they address care in a wide geographic area with all its attendant challenges, and their results report on changes in the care received by over 2000 patients. However, several concerns exist with the operationalization, conduct, and analytic treatment of this study as currently written; these are presented below for consideration by the authors.

Major Compulsory Revisions

1. In their Background section, the authors might wish to more clearly delineate the difference between training and implementation, since they appear to be using them interchangeably.

2. The authors might also more clearly differentiate what they mean by “process” – it is unclear if they are referring to implementation processes and outcomes, to resultant changes in clinical processes, or to the ultimate outcomes of clinical improvement measured at the client level. Implementation outcomes can be changes in clinical process, and this can be measured at the organizational level, as well as (ultimately) at the client level.

3. The authors might also expand their literature review of implementation technologies, especially those like the learning collaboratives model that most resemble what they appear to have done. They might also briefly summarize the ability of procedural standards as instruments of clinical quality improvement, and the challenges associated with efforts in this direction.

4. A terminological issue – the authors’ “clinical guidelines” contain elements such as use of rating scales for assessment, and presence of treatment plans in charts. These are standards of care rather than practice parameters. Making this explicit up front might serve to better orient readers who might reasonably expect to see clinically focused indices of care.

5. The next section dealing with Stockholm county’s implementation of these guidelines will benefit by better organization. Why not move it to the Methods section? It could serve as a discussion of pre-implementation efforts undertaken by health planning authorities in the region.
6. What is the relationship between departments and clinics? The authors mention that 2 departments declined to participate, but 6 clinics were included. How did the departments that declined participation differ from the ones that accepted participation?

7. Regarding the guidelines that are being implemented, the authors might wish to insert a table here that contains the list in the text on p.9.

8. Better operationalization of these guidelines is required. Could clinics substitute one instrument for another and still meet the criteria for use of a standardized rating scale? What constitutes a treatment plan according to these criteria? In their table, the authors could list all these guidelines in one column, and provide further explanation and details for each indicator in a second column.

9. The implementation process (p. 6) is reasonably well described. The authors state that staff set “local goals” for the implementation – does this mean that the set of guidelines to which adherence was expected varied by clinic?

10. The authors also discuss adaptation of care conducted by implementation teams. This needs more elaboration, given the fidelity issues in implementation research.

11. In their Data Collection section the authors mention fulfillment of inclusion criteria – what were these criteria? Also, the charts do not seem to have been randomly selected as the authors state. The date of selection was random, and the charts seemed to have been selected using some type of consecutive sampling strategy. The description of chart selection at the control clinics should be made clearer.

12. If 120 charts were selected at each of 6 clinics at each of 4 waves of data collection, the authors should have access to 2880 charts. Yet, their final tally is 2165. Are the authors presenting a complete case analyses? How were charts with missing or incomplete data treated?

13. In addition to problematic operationalization and validity of some of these indicators, the authors seemed to have dichotomized their outcomes inappropriately. If the original coding is “1=not clear according to the definition” and “2=a clear occurrence”, then the authors cannot combine 1 and 2 and compare that to “0=absence of quality indicator”. The authors should consider evaluating the 0s against the 2s, i.e. examine the presence or absence of meeting that standard. What is the distribution of these 0s, 1s, and 2s?

14. There are serious specification issues in the statistical analyses (p. 10). To control only for age and gender in an observational study is unacceptable. The authors need covariates at the client level, at the provider level, and at the clinic level (at the very least, clinic level fixed effects), in addition to a time variable.

15. The authors seem to be estimating a series of logit models, comparing binary variables of meeting guidelines (yes/no) between baseline and 6 month, 6 month and 12 month, and 12 month and 24 month. This is unacceptable given the significant autocorrelation issues that exist in this kind of a longitudinal study and the nested nature of these data, and the heightened risk of a Type I error in such
repeated analyses.

a. The unit of analysis is unclear. Are these odds ratios comparing outcomes at the client level? Presumably, the authors are not conducting analyses at the level of 6 clinics. These data need a panel data analysis approach. If their unit of analysis is 6 clinics, then it is unclear how the authors are surmounting the issues with lack of power.

16. Leaving aside questions on the validity of these indicators and problems with analyses, the results do show that implementation clinics have higher odds of adherence to these indicators between the baseline and 6 month follow-up. However, it is not clear why the authors would want to compare ORs in this way in models stratified by condition. Why not just introduce a binary variable of intervention clinic or control clinic on the right hand side, unless they believe that the mechanisms that produce outcomes are fundamentally different across conditions?

a. The authors should tell the whole story – improved rates of adherence among control clinics for some indicators, and persistent declines in rates over the course of follow-up among experimental clinics.

17. In their Discussion section, the authors should justify why they think adaptation of these guidelines is a good thing, as opposed to interpreting it as a fidelity problem.

18. The Discussion could be fleshed out in several more directions – their approach to implementation, and critical success factors within it; challenges encountered, and how they overcame them; the contextual policy and health services environment within which they worked and how that affected their endeavors; any secular effects or confounds that occurred at this time; implications for conducting implementation research in other health systems, if any, among several others.

Minor Essential Revisions
1. In both Figures, please provide all X-axis and box-and-whisker labels in the English language.

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

**Quality of written English:** Needs some language corrections before being published

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare I have no competing interests.