Reviewer's report

Title: Factors associated with the impact of quality improvement collaboratives in mental healthcare: an observational study

Version: 3 Date: 29 December 2010

Reviewer: Lisa Rubenstein

Reviewer's report:

This paper has a number of strengths. The study investigated a large naturalistic implementation of quality improvement collaboratives spanning 21 organizations and 29 QI teams. The study also addresses an important area—mental health—and included both mental health specialist and primary care participants. In addition, unlike many papers on the subject of QIC's, the authors defined the key elements of the QIC and determined adherence by QI teams to key elements of their model. The authors also carefully investigated a large number of factors often thought to influence the success of QIC's.

My concerns about the paper are as follows:

1) Methods for dependent variables: There is virtually no methodological data on the outcome variables. Without a believable outcome variable or variables, it is much harder to be interested in the factors predicting that outcome. For example, although we know that only QI teams with at least two months of registration data were included, we do not know how many of the indicated spreadsheets were returned. We do not know how many patients initiated, completed some portions, or completed the entire data collection; we do not actually know what the protocol for data collection was—what data points were the participants to record? We know the “number of patients reached” but we have no information on what that statement means. We know many goals were not reached (Table 4) but it isn’t clear, for example, that the total number of patients in the practice was the right denominator (for anxiety, if I understood it right), or how the authors knew what percent of all dual diagnosis patients were screened.

a. I suspect that the authors know and could summarize enough of this information to give a sense of data quality and composition. As a naturalistic study, it would be unusual if the information were complete/completely specified; I would just want to know more about how complete (or incomplete) it was and the degree to which outcome data quality varied by team. It might be the case that enthusiastic teams were better data collectors, without necessarily being more clinically effective.

b. Being a team that collected data better could in itself be a part of a useful outcome (e.g., the outcome of adherence to the QIC model), however it wouldn’t translate necessarily into better clinical outcomes.

c. (minor) In Table 1, which describes the indicators and standards for the outcome measures, it is not initially clear due to the wording whether the table
was also reporting results (“…50% of the patients with severe anxiety disorders had a score on the CLGI-S lower than three.”)

2) Methods for Independent Variables: Methods related to the data on the teams are not clearly presented. The title calls the study observational, which is technically correct, but which I initially took to mean that the study involved direct observation and qualitative analysis. The abstract mentions nothing about a survey. In the methods, however, (p. 8) we learn that a survey was used to measure factors related to the QIC teams. On p. 9 we learn that team members were questioned on their characteristics; I assume this was from the same survey, but it is written as if it might be an interview. While the survey overall is referenced to several prior publications (p. 8), only one of the actual scales is clearly and specifically referenced to its origins (the Attitude Social Influence Efficacy model, p. 10). However, particularly in the absence of information on the prior validation of the scales used, it would be helpful to include some information on scale performance such as Cronbach’s alpha in the paper; there appear to be plenty of subjects upon which to calculate this statistic. It would also be nice to know descriptively how the teams performed on the key scales; for example, did they perform similarly to previously studied teams? How much did results vary within and between teams? The response rate (reported as 67% on p. 15 though it looks higher from Table 3) is decent for a study of this kind.

3) Data Analysis: The data analysis for this study is qualitative; it most resembles cross-case analysis, which could be referenced but isn't. The method for deciding whether a difference exists (p. 12) is not something I recognize as standard. I would be willing to accept it on an intuitive basis, though a qualitative analyst might (or might not) have other solutions to suggest. There are a large number of hypotheses tested for the number of subjects (teams). Table 5 provides a lot of information but perhaps should be an appendix; a table that honed in on the key hypotheses and findings would be more helpful. To get to this, a clearer view of the conceptual framework or logic model for the analyses might help; it feels as if the authors felt compelled to put every possible variable into their model, rather than either pre-specifying a model or undertaking some sort of variable reduction process to winnow down what they found.

4) The points I’ve listed above are not referenced in the study limitations.

5) The current supplemental material, while nice to have for reference, is too detailed to meet the needs suggested above.

In summary, this is an ambitious and interesting paper. However, as it currently stands, it is not distilled well enough for the reader to easily understand or evaluate its findings. In part this is due to the presentation, which is not as clear as it could be. It might also be that more emphasis on the descriptive aspects of the data versus on the paper as showing factors in relationship to outcome would be useful, if the data based aspects of the latter are too difficult to clarify.

**Level of interest:** An article of importance in its field

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

**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.