Reviewer’s report

Title: The relationship between baseline Organizational Readiness to Change Assessment subscale scores and implementation of hepatitis prevention services in substance use disorders treatment clinics.

Version: 4 Date: 27 February 2010

Reviewer: Martin Charns

Reviewer’s report:

Major Compulsory Revisions

The re-written portions of the paper do add to clarity and better describe the importance of assessing the ORCA.

There are several problems with the measures and analysis that together contribute to doubt that a relationship exists between ORCA ratings and amount of implementation. The small sample size is the major weakness of this paper. The categorization of sites into high and low is also a major problem. I have tried to look at these data in as many ways as possible to find a clear relationship between the ORCA scores and level of implementation and I cannot be convinced of the relationship.

1. The paper in several places notes that the ORCA is intended to measure the organizational readiness for implementing practice change. For example, on page 5 the authors state, “…what is sorely needed is a measure that demonstrates a correlation to actual uptake of new practices following … interventions.” That would imply that the ORCA measures should be compared to practices that are implemented in a subsequent time period. The appropriate measure for this seems to me to be the number of practices that are implemented between baseline and the final measure. That would be computed as the final number of LHI practices implemented minus the number at baseline. In the limitation section, the authors note “…[the classification of two clinics highly adherent to LHI practices at baseline make] … it difficult to distinguish whether the ORCA was assessing ability to implement new practices or was simply correlated with baseline clinic functioning.” I agree with this conclusion, but it raises a major concern about the classification of the nine sites as high or low. A better fit between conceptualization of “implementation” and its measurement would be the difference between final and baseline LHI practices. This would by my calculations result in re-classification of site 3 from Low to High and sites 4 and 8 from High to Low, based on a split at the median of the change score. Among the 9 sites that is a substantial reclassification but puts the measure of implementation more in line with the concept of implementation (vs baseline level of practices). The paper does not present the ORCA comparisons based on this classification.

2. The comparisons between groups is also a problem. Do the comparisons assume a normal distribution of the ORCA scores? With an ordinal scale or at
best an interval scale, we do not usually worry about normal distributions when we have large samples, but that is not the case here. The authors may want to compute some non-parametric statistics. I also would be more convinced seeing the actual ORCA scores for each site. The fact that 2 of 9 did not have overlapping confidence intervals is not convincing for 9 measures, especially when 2 of the nine were not in the predicted direction. I am concerned that we are seeing random variation.

3. Most other measurement issues of concern are noted in the paper. I would add one other aspect of the measurement of implementation as a major concern. What is used as the measure of LHI practices is the team leaders’ reports that the practice is implemented. For some LHI practices, implementation is defined as having is routine testing of all new patients or referral of all hepatitis A and B patients. Does the requirement of “all” strictly apply? For practices 5 and 7 in table 2, having the practice implemented requires something to be in place, but this does not necessarily mean that the practice is used by staff. It would seem to me that a readiness to change measure would be more related to those practices where staff had to be involved to implement a practice. Might the ORCA be more highly associated with some LHI practices than others?

We have no data to explain the differences measured by ORCA, so at best the paper can only speculate on why the results are as reported. Having some qualitative data to explain the quantitative findings would be very helpful, especially given the measurement problems.

**Level of interest:** An article of limited interest

**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 am employed by the Department of Veterans Affairs, which has funded the research presented in this paper.