Reviewer's report

Title: Quality improvement collaboratives and the wisdom of crowds: spread explained by perceived success at group level

Version: 1 Date: 15 October 2013

Reviewer: Mathilde Strating

Reviewer's report:

Although the quality of data is somewhat weak which has several methodological limitations, the topic of this paper is highly relevant for research on QICs and certainly an understudied issue. Given the methodological limitations and the narrow research questions it seems suitable for a short report.

- Major Compulsory Revisions

Theoretical background
1. The authors do not reflect on the concept of spread and related to that how this is operationalized in their study (a rather narrow definition and measurement of spread by the number of units where projects were implemented). Although understand the manuscript is intended to be a short report I would like to see some more theoretical reflection on the concept whether in the background section or in the discussion/conclusion section.

2. What’s new in this report compared with the earlier study of these authors (reference 11 in this paper) is that the analysis is now conducted on QIC level. The relevance of this according to the authors is that QICs may play an important role in the dissemination of projects within hospitals. But why not include characteristics of these QICs in the analysis? And if not available why not reflect more upon what underlying mechanisms at the QIC level can be expected? The added value of conducting the analysis at QIC level is not clear now. We now have to guess why for example the QIC on operation theatre is less successful in effectiveness and spread.

Methodological issues
3. From the results section it is not clear on which N the correlation analysis was conducted. The total number of program coordinators providing dissemination data was 15, but as I understand each program coordinator provided dissemination data for more than one QIC? So in total there were 90 observations (8*6=48 = 7*6=42). These 90 observations are not independent (observations from the same program coordinator may be related), to what extent does this influence the results?

- Minor Essential Revisions

4. It’s not clear from the abstract and main text how many quality improvement collaboratives are included in the study. The abstract says N=12, but concluding
from table 2 there 6 different QICs (with each a different topic) which were conducted in two groups (so: 2*6=12).

5. It is necessary to distinguish the first and second group of hospitals, since the authors do not report on significant differences between these groups and do not seem to have a hypothesis on this.

6. To follow the number of respondents it’s recommended to include some more information on the number of respondents that participated in the text under ‘methods’.

7. I wonder whether differences between QICs could be tested by for example analysis of variance. That would substantiate particular findings such as ‘the operation theatre QICs revealed the lowest average success and dissemination rate’.

8. Since aggregated scores are used it would be helpful to see standard deviations in table 2.

- Discretionary Revisions
It seems to me that table 1 and figure 2 are not that relevant to include. The authors might delete these.

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

**Quality of written English:** Acceptable

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

**Declaration of competing interests:**

no competing interests