Reviewer's report

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

Version: 3 Date: 13 December 2010

Reviewer: Loes Schouten

Reviewer's report:

This is a good paper that adds to the growing literature about Quality Improvement Collaboratives (QIC) and QI methodologies. I recommend publication of this paper.

I have some minor essential revisions/remarks:

Abstract

Background: I suggest the authors include in their first sentence describing QICs that these groups of health professionals come from different organizations

Results: teams with an inspirational leader. I suggest the authors use the term teamleader here as well (as they also do describing "an active QI team leader") to make a clear distinction between senior leadership support from the base organisation. It would be more informative to describe 'better results' in terms of what was achieved for patients in mental health care (or at least give some examples).

Methods

May be the authors could describe the (some) context of the QICs in mental healthcare in the Netherlands. Were these QICs the first collaborative projects in mental health care? Was it one QIC with three kinds of topics or three different QICs. What was the time frame of the three QICs? When were they performed (year). Where they performed at the same time or using different years, time-frames (12 months?)?

Do the authors mean 29 QI teams participated in the study or in the collaborative. Or were these numbers the same for the collaborative and the study. This could be more clearly described.

I would expect a more extensive description of the intervention. What kind of improvement did the teams work on, what were there goals and what did they change to improve care? What were the evidence based recommendations. The authors should include some examples to give the reader some information to interpret the findings and improvements in the quality of mental health care.

Outcome measures

I suggest another format (a matrix) for table 1 to be more informative and concise to the reader and reduce words and redundancy
Table 1
Indicators Anxiety Dual diagnosis Schizofrenia
Outcomes xxx xxx xxx
Monitoring xxx xxx xxx
Screening/assessmentg xxx xxx xxx.

Other measures
I suggest using the heading 'measures related to determinants of succes' or another (in stead of 'other measures'). When describing ‘self designed and validated questionnaires were used' what do the authors mean. That they all self designed and validated the questionnaires or that they used two kind of questionnaires. This should be stated more clearly and in the text or tables the authors should distinguish between self designed and other instruments.

I suggest another format for table 2 to be more informative and concise to the reader and reduce words and redundancy ('QI team' and 'have better results' can be removed in each sentence when the authors include these once in the heading of the table)

Table 2
Hypotheses
QI teams have better results if they........
  a. Composition
     ..... spend more time on the improvement
     ..... have previous experience in quality improvement
     Etc
     Etc

To make a clear distinction between senior leadership support from the base organisation and team leadership I suggest the authors use consequently the term teamleader in their description in their text as well in the tables.

functioning of the QI team
The authors describe '...the attitudes towards the improvement. May be the authors could explain for the readers what they imply here (e.g. .........)

In general the authors should be more precise with placing the references. References about the team climate for example should not be placed at the end of the paragraph but at the end of the text referring to references about team climate.

Organizational context of the QIC team.
The authors should be more precise with placing the references here also (not
just putting them at the end of the paragraph).

Data analysis

The authors state that 'at the start of the study none of the QI teams worked in a structured way on the implementation of multidisciplinary practice guidelines. Consequently the baseline score of the QI teams was assumed zero.' Although this might be true for scores/indicators regarding monitoring and screening and assessment (using specific instruments), this may not be the case for patient outcome measures. I really do not understand why this would be the case for indicators measuring anxiety, quality of life or social functioning. The authors cannot state that the baseline scores for these indicators are assumed zero at baseline. What does this mean for the interpretation of the results relating determinants of success to patient outcomes in the manuscript. This needs more explication in the manuscript.

Results

Table 3

Table 3 refers to the characteristics of the QI teams participating in the study. It may be informative to include some data about the number of teams participating per QIC as well. I should also prefer to include in table 3 the total number of team members studied per QIC and a response rate in percentages.

Table 3 also shows an enormous range in number of patients reached. This may bias the results (are so little number of patients representative). This may be reflected on in the Discussion Section.

Proces of care outcomes: screening /assessment rates

In the text the authors refer to an average of 0.14 patients. In table for this seems to be 1.4. Is this a slip up or something different?

The percentages described in the patient and process of care outcome rates paragraph raise the question what these figures look like if it were absolute number of patients. I think the authors should indicate this in the text and should highlight this in their discussion section.

Factors associated with impact of the QI team

The authors state that of 41 non respondents the demographic characteristics were known (table 3). Do they mean that these were taken in account in the figures described in table 3? Or do they imply that respondents and non respondents do not differ in demographic characteristics? This could be more clearly described.

The authors mention that they selected twenty two factors. Do they mean that they did take all measured factors in account or did they make a selection. If they made a selection the selection procedure should be described.

Organizational context

Please use team leader in stead of leader to be more concise to the readers.
Discussion
The discussion section seems a little rush up. The authors could reflect in some more detail on the results, taken other studies and literature on determinants of success (not especially for QICs but in general) in account.

In the discussion section and table 5 the authors introduce the terminology transformational QI team leader. In the previous paragraphs the refer to inspirational leader. I think the authors should choose one term or give some more text (in the method section) to describe that these terms are more of the same?

Some of the references and descriptions to affirm or support their findings seems not be chosen very well and are sometimes a bridge too far. Many authors have reflect on team leadership and management support to support teams. Why referring to the study of Ovretveit here? Ovretveit [ref 28] suggests that successful leaders need to create an inspiring vision of what the improved organization should do. Although there may be some parallels. Senior leadership in a whole organization and team leadership and management support providing means and time for QI teams etc are not the same (but quite different).

Also the interpretation of the authors of the backgrounds of the team members is hard to follow. What do they mean with ‘professionals that are able to control the improvement’? This requires a more clear description/interpretation and some references or examples supporting these interpretation.

Also including reference 29 [Vos] to support differences in background seems not be chosen very well or at least not clearly related to their own findings.

In the discussion section some general reflections about the operationalization of the items/concepts.

Even the theory-based determinants did not consistently relate to success. This might be caused by a sub-optimal operationalization (and multi dimensionality) of the concepts. It may, on the other hand, be caused by the fact that the association between determinant and effect may be not lineair and may be much more complex. It can be hypothesized that these kind of relations cannot be meaningfully captured by cross-sectional analyses on the association between a single determinant and success parameter. It may be insufficient to take a snapshot of a single determinant. May be a more dynamic, interconnected web of determining factors should be considered in analyses based on several different theoretical models and using a mixed-method research design.

In addition what was agreement like within teams? How much variance was there within teams in their scoring of the items? Did the authors measure the internal consistency-reliability of the items within teams? Perhaps an important point of differentiation among teams would be the extent to which team members agreed about the presence of the success determinants? Consider the team coefficient of variation on a given item as an independent variable?
Having only one observation on the determinants of success is a limitation of the study and should be mentioned. The Nembhard finding suggests that asking your teams to rate determinants of success at the end could have biased their ratings upward for example (assuming they already knew their results when they filled out the survey).

**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 have nothing to declare