Reviewer's report

Title: Measuring team factors thought to influence the success of Quality Improvement in primary care: a systematic review of instruments

Version: 2 Date: 23 August 2012

Reviewer: Miranda Laurant

Reviewer's report:

Title manuscript:
Measuring team factors thought to influence the success of Quality Improvement in primary care: a systematic review of instruments

Comments
A well written manuscript, which is a relevant, but only first, step for our understanding of factors that influence the effectiveness of team work.

The review focus on one specific aspect, namely “the team”, whereas in an companion manuscript other aspects at organizational and individual level are reported. I haven’t read this companion manuscript, so it is difficult to understand why these domains of the framework are not reported in one manuscript. By presenting this in two different papers, it seems that the interaction between organizational level factors and team-level factors and individual level factors is ignored. But perhaps this comes all together in a third paper, where the development and content of the final InQuIRe framework will be reported.

Finally, although I appreciate this paper, I miss at the end some specific guidance. As many instruments seem appropriate to some extent, as only few address the whole domain, most only address certain constructs. And as researcher, self-reported surveys have to be short to avoid non-respond and missing values. You didn’t report anything on these issues, which to my opinion would strengthen the manuscript. In other words, you forgot to guide me how to choose my instruments to measure team-level factors that might influence the outcomes of CQI.

Minor Essential Revision

1) Page 6: There is no reference to the initial framework. It would be very helpful to understand on which work this initial framework is based.

2) On page 6 the term “measurement review” is used, but it is not clear what is meant. Perhaps you used this term as you were looking at measurement instruments which have been used to measure influencing team-level factors. The term ‘measurement’ is redundant, you are performing a review, without meta-analysis.

3) Page 7: although figure 4 reports the inclusion criteria for stage 2, it would also be very helpful if criteria for full text screening (stage 1) is mentioned. The
inclusion criteria stage 1 and inclusion criteria stage 2 are preferable not only reported in figure 4, but also in main text as this is important information to understand the results.

3) Page 11: Please included the number of identified papers (n=6296), and refer to figure 4.

4) Page 12: it would be very helpful if the number of papers and instruments is reported. E.g. Teamwork context was measured in xx papers, including xx unique instruments. Similar for other domains, and perhaps also for the different constructs, eg. Comprehensive measures of enabling conditions was measured in xx papers.

5) Page 28/29: Limitation of the study: Only 1 person included the papers and extracted the data. Only 10% of all papers included in stage 4 were extracted by a second person (page 8). This is a severe limitation of study design. In the discussion (page 29) you concluded that “this is first application of our taxonomy, refinement is likely. …. Alternative categorizations are possible”. Given this limitation (e.g. why not extraction by at least two persons, consensus in broader group) how should I value the outcome of this study. Is it very likely that constructs are different categorized or that categories within constructs should be further refined? Perhaps you can strengthen your conclusion on the value of the taxonomy, refer to 10% double extracted papers, was there good inter-rater score, or did both vary within extraction, and what were (and perhaps still are) the main ‘discussions’ regarding which domain, construct and/or categories.

6) Page 25: Limitation of the study: Only self-report instruments are included in review, while at this page you report that actual behavior can best be measured by observation. To my opinion this is something you could have figured out at the start of the study. I don’t read anything about reasons to limit the review to self-reported instruments, in other words why are other methods, such as observations and interviews, excluded from this review. Furthermore, this restriction should be reported as limitation of this study. Observations, interviews might also be valid to measure team level factors thought to influence CQI-outcomes and could contribute to the taxonomy and guide researchers in their choice of appropriate instruments.

Lay-out/textual issues

1) Page 6: Abbreviation of InQuIRe misses a underline “Quality”

2) Page 8: The InQuIRe framework (figure 1) provided….. . In previous sentences figure 1 was referred to as initial framework. Please be clear is figure 1 the final framework, or initial framework.

3) Page 11: Observational designs were most commonly reported in other papers (n=214). Delete ‘other’.

4) Page 24, line 6: Please rephrase “In combination with the already limited evidence for the measurement properties of some instruments, this means that additional testing of instruments in relevant context is required”.

And please also report what kind of testing you mean, do you mean to test the instruments validity and realiability or do you mean that instruments should be
tested whether or not these constructs the instruments represent/test do influence CQI outcomes.

5) Page 24, last sentence: this sentence seems out of place here.

6) Table 1, page 44: Example content validity seems missing

7) Table 3: for readability, consider # or other symbol instead of X (bold X) and X*. Reader a much better and quicker overview of the table. Similar for tables 4 and 5

8) Table 6: for readability there should be a clear distinctive line between ‘other healthcare’ and ‘non-healthcare’

9) Figure 2: is this figure correct? “Team composition & structure” is a category with different elements like Size, Tenure, etc. but Competencies of team members, again with different elements as sublevel, seems to be an element (font size), but is also called category. Please check.

10) Additional file 2: only report search PsychInfo, please also include search terms used Medline and HaPI.

11) Additional file 4 and 6: add (Stage 4)

**Level of interest:** An article of outstanding merit and interest in its field

**Quality of written English:** Acceptable

**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