Author's response to reviews

Title: The structure of quality systems is important to the process and outcome. An empirical study of 386 hospital departments in Sweden.

Authors:

Stefan T Kunkel (stefan.kunkel@pubcare.uu.se)
Urban Rosenqvist (urban.rosenqvist@pubcare.uu.se)
Ragnar Westerling (ragnar.westerling@pubcare.uu.se)

Version: 2 Date: 11 May 2007

Author's response to reviews: see over
Revision Letter

1 Referee 1: Dr Shekelle

1.1 Minor Essential Revisions

The typology presented in the background is not something that I think is universally accepted and I would suggest simply saying that QI efforts can go under a lot of different names like CQI and TQM and balanced scorecard etc, and that for this paper we've defined quality systems as...

We have changed the recommended paragraph in Background (third paragraph) according to the referee's suggestion. It now indicates that quality systems "have many names" rather than "can be defined as".

I think the authors need to acknowledge that the head of department may not be the best person to answer questions like "Do the clinic's employees and managers have time to work with quality improvement?" and "In general, is it easy to get support from the clinic's managers when trying to implement new organizational improvements?" I'm pretty sure that the head of my department would answer these things somewhat to a lot differently than those of us actually working in the clinic!

Head of departments and quality coordinators were chosen since they were thought to possess a more detailed knowledge of quality systems. We have added an explanation of this to Methods (The data obtained from the questionnaire, first paragraph).

We have also added a discussion of the limitations of this choice to Methodological Considerations (eighth paragraph). There we discuss the possibility of different answers among managers and non-managers, and we also mention that there were no significant differences between how heads of departments and quality coordinators responded. Furthermore, an appendix B has been added with the frequency distribution of all variables and described under Methods (The analyses of variables).

I would suggest moving the paragraph in the methods about the possibility of bias resulting from the non-responders from the methods section to the discussion section. This is normally where I see this presented.
We agree and have moved the paragraph as requested to Methodological Considerations (fifth paragraph).

*I think the authors need to state in the discussion that there is no demonstrated link between their "outcome" and the outcome that usually matters most to patients - better health outcomes.*

We have added a paragraph under Implications (last paragraph) reminding the reader that since no health outcomes were measured, no links between quality systems and health outcomes could be demonstrated.

*The authors may want to consider pushing the conclusions a bit further - would it be incorrect to conclude that organizations that say they have a quality system but are not devoting adequate resources to it are fooling themselves? Or that if organizations want to improve the quality of their care then they need to devote sufficient resources to their quality systems in order to achieve substantial results?*

We appreciate the suggestions and have added a variation of the second suggestion to Conclusions (last sentence): "This would indicate that, for instance, adequate resources and administration may play an important role in systematic quality work."

### 2 Referee 2: Dr Boelle

#### 2.1 Major Compulsory Revisions

- *In the questionnaire sent to the hospitals, was the subdivision (Structure/Process/Outcome) made apparent? In this case, it must be evaluated how this presentation could have led to some of the reported associations.*

  No, the subdivision was not made apparent in the questionnaire. Structure, process, and outcome were not mentioned. The subdivision in the questionnaire and the numbering of questions in Appendix A have only been done to improve readability. We have added an explanation to Methods of this fact (The operationalisation of the model, first paragraph).
I find it very troubling that out of 6 carefully crafted questions, 1 or 2 may not be associated with the corresponding factor (once for Outcome, once for Process). Although the authors claim (p.7) that “the questions were able to represent the different factors with adequate precision”, I find otherwise that this very much questions the sensitivity and specificity of the questionnaire; it also emphasizes the fuzziness in the distinction between the three item types.

We have added explanations and clarifications to Methodological considerations on the matters. We have clarified, in the first paragraph, the steps that we have taken to assure that the questions are good reflections of their factors (sensitivity, significant factor loadings), that the factors are separated from each other (specificity, factor correlations <1), and that the measurement model as a whole reasonably represents reality (model fit indices, p>0.05).

The structural equation model was not based on the inexact questions since they were removed. The model is only based on those questions that were found to be good reflections of their factors. The statistics also indicate that they are reasonable representations of their factors. This has also been clarified in the second paragraph.

This study addresses “Attitude” rather than “Practice”, furthermore it may be based on a very partial view of the quality system. For example, take question “B3: Are the clinic’s employees positive to reporting incidents?” How may the answer be evaluated? Is this based on the “feeling” of the person who fills the questionnaire? In this case, it will likely depend on its role in the clinic. This issue could have been addressed by having multiple answers from the same department. More generally, the questions are very general (as the authors acknowledge) and therefore may be very open to “subjective” judgement. It is not clear how this may depict the reality of quality systems implementation.

We agree that some of the questions may be sensitive to attitudes of what role a responder had in a department. More clearly stated in Methodological considerations (paragraph 8), there were no significant differences between head of departments and quality coordinators for any of the questions. This would indicate that at least between these two different roles, there were no differences.

Question B3, which was given as an example by Dr Boelle, was not included in the structural equation model since the factor analysis indicated that it did not adequately reflect its factor. It is quite probable, as the referee suggests, that the answers on this particular question
would have been different if, instead, physicians or nurses without manager position had been asked. We have clarified that since this question was removed, along with a few other potentially sensitive questions, it would not affect the presented structural equation model (Methodological Considerations, second paragraph).

- It is not clear how much support for Donabedian’s description lies in these numbers. It is obvious that provided a large number of subjects, one may eventually find significant associations everywhere. The magnitude is much more at stake here. The association 0.2 between process and outcome is rather weak, and gives little support to the “Process -> Outcome” link.

We agree that the association between process and outcome is rather weak. We have improved our explanation that this should be seen in the light that this relationship represents the correlation between these factors when the effects of structure have been removed. Thus, another way to describe this is that even though structure seems very important to outcome, there still remains an effect from process. Discussion (fifth paragraph) has been rewritten to increase clarity on this matter.

Dr Boelle is correct that with a sample large enough, all correlations become significant. However, with structural equation modelling this effect is countered by the use of model fit indices. The larger the sample, the smaller differences between the model and the data set (="reality") become significant, and the larger the chance of rejecting a model (p<0.05). We have showed that the correlations are significant and at the same time demonstrated that the model as a whole is a good representation of reality. We have added an improved explanation of the advantages of using structural equation modelling in this respect to Methodological Considerations (ninth paragraph).

- It is essential to analyze quantitatively, for example by sensitivity analysis, how the non responding hospitals may affect these conclusions.

We agree that it is important to analyse how non-responding departments might have affected the conclusions and have added a paragraph to Methodological Considerations (sixth paragraph). There we acknowledge the potential problem but point at circumstances that hopefully lessens the impact of the problem. For instance, we have conducted new LISREL analyses of the early responder group and have compared the SEM estimates and model fit
indices to those of the total sample. The differences were very small and did not change the conclusions. Moreover, all statistical tests were passed with good margins.

We have also rearranged the paragraphs so that it should be clearer what type of potential differences that we have analysed: between responders and non-responders (sixth paragraph), between different methods of estimation (seventh paragraph), and between heads of departments and quality coordinators (eighth paragraph). The conclusions of these analyses were that we could not find any statistically significant differences between any of the groups.

2.2 Minor Essential Revisions

- It would be nice to display the result of the 18 questions in terms of percentages.

The table has been included as Appendix B and an explanatory text has been added under Methods: Analyses of variables.

3 Referee 3: Dr Staccini

3.1 Discretionary Revisions:

- What was the history of quality management in each department? Do they have the same quality management structure? Is there a central department which specifically handles all quality programmes in the hospital? Have all departments been involved in quality management for a same period of time?

We have added a paragraph to the Background stating that departments are required by law to have a quality system, but that it differs how departments choose to organise their quality work (third paragraph). We have also added a paragraph to Implications (sixth paragraph), where we discuss areas for further research: "In further studies, it would be interesting to investigate the implementation of quality systems and relate these processes to departments' history of quality management and overall organisational structure of hospitals."

- What about risk as an outcome? Has it been admitted that risk reducing can be a good incentive to motivate health care staff?
Reducing risk or penalties for employees that report mistakes is often mentioned as a way to increase reporting. It is an interesting suggestion to include risk (or risk reduction) as an outcome. Risk would thus be an excellent candidate aspect to be included in a future more detailed and developed model. We have also added a paragraph to Implications (paragraph 6), where we discuss areas for further research: "It could also be interesting to evaluate the achievement of specific quality goals because some goals are probably more easily achieved than others."

- *Is there an explanation for the inconsistency of B2 and B3 items? Could it be linked to the heterogeneity of the organization of quality management in the departments?*

That is a possible explanation. Another explanation could be that the questions are about sensitive matters. We have added a paragraph to Methodological Considerations (second paragraph) where we discuss the inconsistency and our decision to remove these variables from the analyses.

- *Will it be possible to use this method to test the consistence of the questions of a quality referential?*

The methods CFA and SEM can be used to analyse consistency of the questions as long as the variables are measured on an ordinal or metric scale and sample size is large (n>100, minimum, higher depending on model complexity and method of estimation). It is also necessary to have an idea of what relationships to expect (otherwise it is an exploratory factor analysis, not a confirmatory). We have added to Implications (fifth paragraph): "For instance, it could be used to analyse the consistency of questions in quality referentials."

- *How do you explain the poor link between process and outcome? Is the inconsistence of B2 and B3 a cause? B3 depicts a great challenge for all risk/quality management systems to motivate adverse event reporting with a positive feeling. Outcomes include adverse events statistics. So, could it be an explanation?*

Maybe, but we think that the main reason for the low correlation between process and outcome is due to the fact that all effects of structure have been removed. Thus, it is a conditional
(or partial) correlation. A Discussion paragraph has been modified to further clarify this (fifth paragraph).

4 Referee 4: Dr Hutchinson

4.1 Minor Essential Revisions

Page 4, analysis of non-responding depts, sentence 2 'this may suggest etc'. There is no evidence for this statement and it should be removed.

The statement has been removed as requested.

Page 6, discussion section, paragraph 5, sentence 'for instance support from etc.' It is not clear what is meant by this sentence, particularly the section that says 'to which a quality system is evaluated'. Requires clarification for the reader.

We have rewritten and clarified the requested paragraph. It now states that: "For instance, work to increase support from colleagues could increase the probability that quality efforts get systematically evaluated."

Page 7, second paragraph, last line. 'would probably have' goes beyond the evidence. The evidence for quality improvement to support guidelines is much weaker than for clinical practice guidelines (where the evidence is also weak). I would accept 'might' instead of 'would probably'.

We agree and have exchanged 'would probably' for 'might'.

The next 2 paragraphs are examples of what the authors refer to in paragraph 1 and therefore should be more closely linked together, with a 'for instance' or a 'for example'.

We have added the linking words "For instance", "Moreover", and "Last" to improve readability as suggested.