Author's response to reviews

Title: A critical review of the research literature on Six Sigma, Lean & StuderGroup's Hardwiring Excellence: the need to demonstrate and communicate the effectiveness of transformation strategies in U.S. healthcare

Authors:

Joshua R Vest (jvest@srph.tamhsc.edu)
Larry D Gamm (gamm@srph.tamhsc.edu)

Version: 2 Date: 21 April 2009

Author's response to reviews: see over
Drs. Eccles and Mittman,

Thank you very much for sharing the very thoughtful comments from the reviewers. We have been responsive to the extent possible given the original focus of our paper. We have, in fact, added some additional research articles from three other citation indexes at the suggestion of the reviewers to the original set that the paper analyzed. Although this addition did not change the tenor of the conclusions, it served to offer the broader representation of sources proposed by one of the reviewers.

Below we present the comments of the reviewers followed by our responses to each and noting any resultant revisions to the paper. Again, we have found the reviews to be fair and very positive regarding the paper. We feel the reviewers have enabled us to strengthen our paper.

Responses to Reviewer #1

1 What do the findings mean, given the methodological limitations in the studies reviewed that might undermine the validity of the results? Management research is tremendously underfunded. The purpose of this research is not to prove hypotheses but rather to improve the quality of managerial decision-making. To what extent can any of these studies be useful in what ways to give managers better evidence for managerial interventions.

The utility of the studies reviewed is limited by the lack of rigorous design that enables one to understand that whatever degree of success reported was directly associated with the intervention/innovation or with particular elements of the intervention. Also, the assumption that results of negative findings are less likely to be reported or published, makes it difficult to assess how typical reported positive results are of other efforts to implement the interventions. Nonetheless, depending upon the quality of the documentation, each of these studies reviewed may be of practice value or heuristic value to managers in terms of guiding their thinking about steps necessary to implement particular interventions.

2 By the way, the cost of the transformational strategies is ignored in the research reviewed and in the analysis.

Addressed as part of comment #3.

3 I judge most of the authors' recommendations (e.g. academic partnerships, research administrators) not easily feasible. Health care organizations and systems generally don't wish to fund management research (although they would be the chief beneficiaries of management research). Government doesn't place a high priority on this kind of research either.

This comment can be contrasted with the fact the cost of research, knowledge development, and innovation is seldom considered in clinical transformation, because it is often externally funded and “not adopted” until it has been tested and produced for adoption. Questions of efficacy, effectiveness, and
efficiency—compared to other options often continues for years after a treatment innovation is adopted. One could argue that management innovation, in contrast, is introduced real-time in a real-life laboratory. Inasmuch as managers are participants in the process being researched, immediate incentives for such research are likely to be positive for management committed to continued improvement and organizational learning, but possibly negative for those who are more committed to personal judgment and quick action. In either case, increasing emphasis on evidence-based practice in medicine, management, and other practices suggests that managers and maybe their boards will be more interested in the ability of managers to capture and structure information to report on the impact of their managerial investments. As such we have added the following paragraph on page 17:

While suggesting this avenue to improvement, we are aware that funding for evaluation and management research is not a priority for many health organizations. Again, however, this re-emphasizes the point for improved research studies in order to demonstrate the value of these strategies. The obvious potential for cost-savings or reductions were implied by the improvements in almost all of the reported studies, however, only a couple specifically indicated how much money was either saved [33] or how revenues were increased [23]. Justification for evaluation and research is made easier when expected savings are available to offset those costs and those savings are expected to be ongoing. Still other opportunities exist for improved partnering between health services researchers and practicing organizations. Academic medical centers represent innovative institutions with a history and expectation of research and appear to be natural settings for these types of investigations. Evaluation of these and other transformative strategies may be slightly different than historical interest in clinical applications, but industrial / system engineers are more accessible and the culture is still one of research. Additionally, those seeking executive health management degrees, student interns, or even professionals returning to school for advanced degrees while still employed all provide opportunities and interested individuals for collaboration.

Responses to Reviewer #2

1. Title and abstract – neither make it clear that the review was limited to US studies only, and this needs to be corrected.

U.S. healthcare organizations and system have been specified in both locations in the text as suggested.

2. The results and conclusions sections from the abstract could be more representative of their findings. The results as presented here are too even-handed and, to be frank, I didn’t understand the conclusions!

Both the methods, results and conclusions sections of the abstract have been changed in response to this comment and comment #4. The new text reads:

Background: U.S. healthcare organizations are confronted with numerous and varied transformational strategies promising improvements along all dimensions of quality and performance. This article examines the peer-reviewed literature for evidence of effectiveness
among three current popular transformational strategies: Six Sigma, Lean / Toyota Production System, and Studer’s Hardwiring Excellence.

Methods: The English language health, healthcare management and organizational science literature (up to December 2007) indexed in Medline, Web of Science, ABI/Inform, Cochrane Library, CINAHL, and ERIC was reviewed for studies on the aforementioned transformation strategies in healthcare settings. Articles were included if they: appeared in a peer-reviewed journal; described a specific intervention; were not classified as a pilot study; provided quantitative data; and were not review articles. Nine references on Six Sigma, 9 on Lean / Toyota Production System, and 1 on StuderGroup meet the study’s eligibility criteria.

Results: The reviewed studies universally concluded the implementations of these transformation strategies were successful in improving a variety of healthcare related processes and outcomes. Additionally, the existing literature reflects a wide application of these transformation strategies in terms of both settings and problems. However, despite these positive features, the vast majority had methodological limitations that might undermine the validity of the results. Common features included: weak study designs, inappropriate analyses, and failures to rule out alternative hypotheses. Furthermore, frequently absent was any attention to changes in organizational culture or substantial evidence of lasting effects from these efforts.

Conclusions: Despite the current popularity of these strategies, few studies meet the inclusion criteria for this review. Furthermore, each could have been improved substantially in order to ensure the validity of the conclusions, demonstrate sustainability, investigate changes in organizational culture, or even how one strategy interfaced with other concurrent and subsequent transformation efforts. While informative results can be gleaned from less rigorous studies, improved design and analysis can more effectively guide healthcare leaders who are motivated to transform their organizations and convince others of the need to employ such strategies. Demanding more exacting evaluation of projects consultants, or partnerships with health management researchers in academic settings can support such efforts.

3. Research question – the authors do not explicitly address a key question about all of these interventions in this review, which is the ‘sustainability of effect’ on practice or culture. For example, a process improvement workshop may well produce a short-term change in practice, but does it last? While the decay of effect is mentioned in one study in the results, explicit identification of this issue upfront is warranted.

While in agreement that we did not explicitly identify a research question around sustainability of effect we contend that our concern was neither hidden nor left unaddressed in the introduction. The cited definitions of transformation imply a concern for sustainability and we cited prior research indicating a primary problem was a view of such strategies as management fads (page 5). Again we would agree our treatment of sustainability from a research perspective is implicit as these transformation efforts
assume or plan for continued effort and success over time. Making the question explicit, however, would not change any results, findings or conclusion given that the articles reported on very recent initiatives. The research, therefore, could not provide much evidence of sustainability or even prolonged evaluation beyond the noted study. We also added the following sentence to page 16 (last full paragraph) to reiterate this conclusion: “Also, the longer time period can offer additional evidence of sustainability.”

4. Search strategy – the databases searched were a little limited, even for the US only, as was the basic key word search. US studies could be published in non-US databases, so their inclusion would have been advisable. An example of relevant databases to search for this topic would be: Cochrane Effective Practice and Organisation of Care Group (EPOC) specialised register, Cochrane Controlled Trials Register, MEDLINE, EMBASE, PsycLit, Cinahl, ABI Inform, the Science Citation Index, ERIC, the System for Information on the Grey Literature and the Health Management Information Consortium database. A check on the missing databases to see if they have missed anything obvious would reassure this reviewer. Or, at the very least, acknowledgement of the limitations.

In order to address the above concerns we expanded our search (using the same timeframe and inclusion criteria) to include the databases suggested above to which our institution has access. The number of new articles are summarized below and the methods and results sections updated to reflect the change.

<table>
<thead>
<tr>
<th>Source</th>
<th>Six Sigma</th>
<th>Lean /TPS</th>
<th>Studer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cochrane library</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CINAHL</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>ERIC</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Totals</td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

The discussion of these articles is primarily incorporated into Pages 10&11 and are summarized appropriately in Table 2. The suggestion from the reviewer was welcomed, as none of the additional articles contradicted our previous findings, but simply strengthened them by contributing additional evidence of less than ideal methods, and doubtful conclusions.

The discussion also needs to acknowledge that this was not a systematic search: some interventions which would have met the criteria to be lean etc, may not have been described as such and therefore missed on a simple keyword search. Further, the authors did not hand search relevant journals to look for studies.
This point is well taken and we have avoided use of the “systematic”. The acknowledged limitations of our strategy were somewhat lacking, so we have added the following text on Page 18-19:

We did not undertake a hand search of any journal titles, because all the healthcare management related journals we would consider are indexed in our original search set.

5. Data extraction – acknowledge that the level of detail that would be extracted in a Cochrane EPOC review was not achieved, in particular on the interventions (“what types of training, education?” etc), and therefore the ability of the reader to transfer to their own setting would be limited.

The point is well taken as our data abstraction was not to this level, but our object was not to critically examine the intervention itself. We added the following text to page 7:

“The goal of this article was not to critique the interventions themselves, so the level of information extracted was not to the depth of very rigorous systematic comparative reviews such as a Cochrane EPOC review. Readers wishing to critically examine the intervention in greater detail are referred to the original publications.”

6. Results – there is some confusion in the description of biases that were not avoided by the constituent studies. For example, confounding biases cannot be eliminated by statistical analyses: this is a design issue, only fully satisfactorily dealt with by randomisation.

In order to be consistent on the use of statistical procedures to adjust for confounding factors, Page 10, 1st paragraph – Changed sentence:

(FROM) “These studies are also similar in that they did not engage any statistical tests, so no inferences can be made and no confounding bias was controlled.”

(TO) “These studies are also similar in that they did not engage any statistical tests, so no inferences can be made. Nor was there adjustment for potential confounding bias.”

7. Results – the critique of the Interrupted Time Series (ITS) studies could be more in-depth. For example, a single group ITS is not as robust as a controlled ITS. I refer the authors to ‘Craig R. Ramsay, Lloyd Matowe, Roberto Grilli, Jeremy M. Grimshaw, Ruth E. Thomas. Interrupted Time Series designs in Health Technology Assessment: lessons from two systematic reviews of behaviour change strategies. International Journal of Technology Assessment in Health Care, 19:4 (2003), 613–623’ for further guidance.

On page 14 we noted the weakness of one study’s interrupted time series by noting it was essentially relying on graphical evidence and on page 17 we mention “well executed time series designs” would be an improvement and within that paragraph we highlight how outside expertise would help. On the above suggestion we include the Ramsay citation in the latter section to point readers to the resource, but disagree an expansion of the discussion on ITS is necessary or fits within the context of the article. In
terms of the studies with ITS, we point out the analysis and methodology were insufficient, as we did with other approaches like those ignoring correlated data. In both these cases, we acknowledge the shortcoming and suggest the means to redress the problem (apply the correct statistics) or for others to employ a better design (ITS vs pre – post). Going into more depth about ITS would be out of place in the article, because for no other shortcoming do we give detail about how to select a better comparisons, or choose the right statistical test.