Reviewer's report

Title: Patient safety culture in a Large Teaching Hospital in Riyadh: Baseline Assessment, Benchmarking and Opportunities for Improvement

Version: 3 Date: 24 June 2013

Reviewer: Christine Hartmann

Reviewer's report:

This is a well-written description of a one-time survey of safety culture in one large hospital (with two sites) in Riyadh in the Kingdom of Saudi Arabia. Analyses were also performed to compare the survey results with data from studies in Lebanon, Turkey, and the U.S. that used the same survey instrument.

Major Compulsory Revisions:

• Abstract: Words such as "determinants" should be avoided here. Cross-sectional studies can assess associations and relationships but not determinants or causes.

• I suggest that the use of the word “benchmark” throughout may be problematic. AHRQ has not traditionally referred to its database as a "benchmarking" database but rather as a "comparative database report." AHRQ has done this for a reason--the data they collect cannot truly be considered to be representative of a standard of excellence. The sample reporting is a completely voluntary sample. It may well be that "comparing hospital safety climate among different populations of hospitals" or something similar may be more accurate to use in this manuscript as well. However, first a better description of the results from the other countries is warranted. What were the samples? How large? What types of respondents? How many hospitals? How were data collected? What were the response rates? For the US sample, were data received from the AHRQ HSOPSC database? From what year(s)? When the above information is included, it will be easier to determine whether benchmarking is the best term to use.

• Design, Setting, and Sampling: How was the HSOPSC used in this study “adapted” from the version used in the study in Lebanon?

• Data management and analysis: Regarding the cross-country comparisons: The wording used is unclear. Was the comparison made for each composite in pairs, or was there an aggregate score created across all the composites and the comparison only made for this aggregate score? This also pertains to the description of the results on p. 11. On a related note, it would be valuable to have the significances indicated in Figure 1.

• Discussion, p. 18: The impact of hospital size on results was, indeed, interesting. What might help account for the differences in this case? What does the literature on this subject indicate?
• Limitations: Other limitations worth noting include the conduct of the study in only one hospital, the cross-sectional nature of data collection and the lack of ability to conclude anything about causation, the lack of information about other factors affecting safety climate.

Minor Essential Revisions:
- Introduction, p. 4: Does a higher score reflect stronger or weaker safety climate? It would be good to clarify for the HSOPSC naïve reader.
- Design, Setting, and Sampling: It’s unclear whether the electronic version of the survey was also returned as a hard copy or by some other means.
- Benchmarking against regional and international findings: The first sentence here omits Turkey.
- Discussion, first sentence: The sample was large in absolute numbers but comprised only 2 settings, both of which were part of the same hospital. This study may have been performed with the largest sample to date, but the wording of "large scale" and "wide scale" is somewhat misleading.

Discretionary Revisions:
- Results: It would have made the results more interesting to see data on differences among employees of different job types or management levels, as has been done in prior work on safety climate/culture. Safety culture data have also been analyzed at the unit level, which would have also been interesting to see.
- Discussion: More discussion about how these results might be used in practice would be welcome. The foundations for eventual bridges between safety climate assessment and safety climate improvement need to be constructed, even when the study is cross-sectional in nature. Ideally, in the future, longitudinal work would be performed, as well as assessments of the relationship between safety climate and clinical and workforce outcomes.

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 declare that I have no competing interests.