Reviewer's report

Title: Should policy-makers and managers trust PSI? An empirical validation study of five Patient Safety Indicators in a National Health Service

Version: 1 Date: 15 November 2011

Reviewer: Vinita Bahl

Reviewer's report:

This manuscript provides an interesting and useful analysis of the empirical properties of the AHRQ Patient Safety Indicators (PSI) and sheds light on the utility of these indicators as a measure of differences in safety performance between hospitals. The results of the analysis describe real and practical issues with using PSIs to evaluate hospital performance.

In general, I would ask the authors to be more explicit about the details of the AHRQ PSI software they used and whether some of the adverse events that were identified were present on admission. In addition, I recommend that the authors, given the results of their analysis, provide more specific guidance about the use, and particularly the limitations, of the PSIs.

Major Compulsory Revisions

1. The authors must document whether the PSIs they studied identified only those adverse events that were acquired during hospitalization or whether they included adverse events that were present on admission. (From the comment in discussion section about decubitus ulcers, it appears that the adverse events that were present on admission were included)

If the PSIs include adverse events that were present on admission, then these events would have contributed to systematic differences in rates between hospitals. The authors should evaluate the impact of the events that were present on admission on these differences. I don't think the tests of the influence of patient comorbid conditions and hospital case mix index would have uncovered the impact.

2. In the discussion section, the authors were careful to point out the impact of incomplete reporting of secondary diagnoses and the need to compare "like" hospitals. They also commented on the validity of the PSIs, which has been tested to some degree by the Spanish NHS. However, in their closing remarks, the authors recommend that the PSI can be used as a screening tool to "identify those centres from which best practice lessons can be drawn out and those where intervention is clearly needed."

Given the study results, I recommend that the authors be specific about how results that show some centres with low O:E ratios and others with high O:E ratios should be interpreted and also provide more specific guidance to centres
that have these different results. It is possible that centres with high O:E ratios may not be adequately reporting secondary diagnoses and so they should improve the documentation and coding of patient comorbidities. In this scenario, the intervention is about coding, not improvements in patient safety. Yet, the authors use of the term "intervention" seems misleading because it suggests that centres with high O:E ratios have an unfavorable patient safety performance.

3. In the methods section, the authors should explain why they chose to analyze the 5 selected PSIs and why they chose not to analyze other PSIs or even all PSIs.

4. In the discussion section, the authors apply the lessons learned from this analysis of 5 selected PSIs to all PSIs. It would be helpful if they described how and why the analysis of 5 selected PSIs can be generalized to all PSIs.

5. In the results and discussion sections, the authors should provide a specific description of why the analysis of the PSI for death in low mortality DRGs produced results that were different from those of the other 4 PSIs and the implications of these differences.

6. In the discussion section, the authors should document the limitations of their study and of their results.

Minor Essential Revisions

The authors should document the version of AHRQ PSI they are using.

Discretionary Revisions

Modify first three sentences of opening paragraph to improve grammar and quality of the written English. Here is a suggested revision:

"The Spanish National Health Service, like others, has become influenced by the patient safety movement. Evidence from two reports on Spanish hospitals that followed other international works on adverse events, inspired the debate. The first one showed an inpatient incidence of adverse events ranging from 5.6% to 16.1%, where between 17% and 41% were deemed avoidable. The second one, found an incidence of adverse events requiring intervention of up to 10.1%.

The quality of the written English improved after this first section.

Level of interest: An article of importance in its field

Quality of written English: Needs some language corrections before being published

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.

Declaration of competing interests:
I declare that I have no competing interests