

## Reviewer's report

**Title:** Using a summary measure for multiple quality indicators in primary care: the Summary Quality Index (SQUID)

**Version:** 1 **Date:** 20 December 2006

**Reviewer:** Martin Lee

### Reviewer's report:

General

-----  
Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

The fundamental concept here is sound. It is reasonable to establish a single pooled quality estimate from multiple quality indicators and to use this to evaluate the performance of a healthcare unit or system. However, the attempt here is incomplete and, in my opinion, overly simplistic, or, at least, the evidence provided is not convincing that this approach is sufficient.

The authors use a basic proportion of indicators met to define their index (and spend far too much space in the paper describing this extremely basic statistical concept). As a basic premise for this paper, I find this not properly developed. Much more needs to be done in order to provide support to publish this simplistic concept. For instance, could this index be used in a prospective evaluation of quality improvement? Why is an unweighted index as good as one that weights according to difficulty of achievement or the severity of the disease involved? (Although the authors do mention this in their discussion, they need to pay much more attention to this as most composite health indices use some sort of weighting in their development.)

Their attempt to provide validation information is commendable, but again I think this is underdeveloped. One key statistic that is missing is the distribution of E among the patients. Clearly, if E tends to be very small for a reasonable percentage of the subjects, then the possible values for SQUID are, of course, very limited. Under these circumstances, the responsiveness of this index is going to be quite different than when E is closer to the maximum. They report the responsiveness in terms of the change in percentage over a 15 month period. The difference seems to be small in absolute terms and is strongly statistically significant only because of the huge number of subjects involved. It would be useful to also know what this numerical change is for all subjects and not for just those who were active during the entire period (i.e. overall value for the 3rd quarter, 2004 versus the overall value for the 3rd quarter, 2005) as I would suspect that those who may not be doing as well might not continue to appear in the system for evaluation.

As an overall evaluation of this paper, I think that conceptually the authors have a interesting idea, but the approach taken here is convincing enough yet. Either they consider a more in depth quantitative evaluation of their index or they use this forum to describe the qualitative issues and difficulties encountered in the actual use and development of such a concept in a large healthcare system.

-----  
Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

-----  
Discretionary Revisions (which the author can choose to ignore)

**What next?:** Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.