Reviewer’s report

Title: Dutch healthcare reform: did it result in performance improvement of health plans? A comparison of consumer experiences over time

Version: 1 Date: 16 February 2009

Reviewer: Hans Maarse

Reviewer’s report:

This article deals with the question whether the introduction of managed competition along with reporting of quality information in Dutch health care is associated with quality improvement in health plans. To answer this question, the study uses seven quality aspects of health plan services. The number of items to measure the quality aspect ranges from 1 to 5.

The authors hypothesize that changes over the years (2005, 2006 and 2007) would be more profound for quality aspects that were identified as aspects that needed improvement most and for health plans with an inferior performance at first measurement. This hypothesis seems to imply that there was no need to improve the performance of ‘above-average’ health plans. I consider this questionable (‘good is the enemy of better’).

My greatest problem with the article is that the authors compare 2005 with 2006 and 2006 with 2007. I strongly doubt whether the comparison 2005-2006 makes any sense, because that year was very hectic due the introduction of the new health insurance scheme. The authors seem to realise this when they question whether a period of six months is long enough to initiate improvement projects in health insurance service. So, I suggest to restrict the analysis to a comparison of 2005 with 2006. This measure would also strongly contribute to the overall readability of the article, in particular the section ‘comparison of changes in performance between health plans’. Furthermore, it is no longer necessary to find an explanation for the fluctuations in the results (note: a convincing explanation is nowhere given).

I consider the discussion as disappointing for at least three reasons:
- the earlier distinction between professionalism, market share and reputation does not appear very helpful. The authors don’t do much with it, because data are lacking. Furthermore, I do not believe that a clear distinction can be made between reputation protection and fear for loosing market share.
- It would have been highly interesting to find out whether the under-average performers have indeed lost market share whereas the above-average performers have indeed won market share.
- The authors seem to forget that their analysis concerns about 40% of the market. They do not spend a single word on the role of collectivities in health insurance (presently, about 60% of the market).
- The discussion on the relevance of public reporting at the end seems to me irrelevant. This is just another topic.

Further questions:
- there is much variation in the number of observations per insurer. Consequences?
- Why no non-response analysis?
- At several places the authors speak about to anticipate. I wonder whether this word is correct. I prefer words like ‘to expect’ or ‘to assume’.
- The comparison interval is defined as 1,36 x standard error. Why 1,36?
- Does statistical significance also imply policy relevance (remember that statistical significance is not always clinical relevant).
- Why do the authors suggest that relatively bad-performing health plans had more possibilities for improving their service? Is that always true?

Conclusion: major revision required