Reviewer's report

Title: Can we trust measures of healthcare utilisation from household surveys?

Version: 2 Date: 13 June 2013

Reviewer: Manoj Mohanan

Reviewer's report:

Major compulsory revisions:
g
Minor essential revisions:
c, d, e, f,

Discretionary revisions:
a, b

This manuscript by Ansah and Powell-Jackson poses an important and timely question about whether researchers can rely on commonly reported measures of healthcare utilization from household surveys.

The authors compare findings from a standard household survey, with findings from a pictorial diary to make inferences about problems with recall and reporting in household surveys. The underlying premise is that since the diary method requires the respondents to enter information on a regular basis, the data would be relatively more accurate as compared to surveys that have been known to have recall problems. The authors also use the two sources of data to demonstrate how an analysis of the impact of removal of user fees would arrive at strikingly different conclusions depending on the data source.

I think this is a very interesting investigation and believe that we need more papers that understand the problems with standard survey methods. That said, there are a few areas that need to be addressed in the current version of the paper.

a. A prominent paper that has addressed similar concerns of recall is by Das, Hammer, & Sanchez JDE 2012. The authors mention this paper in the discussion – would be helpful to cover this in the literature summary and also point out what the current paper adds in light of Das et al.

b. Figure 3 reports morbidity / illness episodes across time periods from the two data sources. If I am reading this correctly, the graphs on the right side show % of households that reported illness in past 6 months. The panels on the left, however, report number of episodes over a year (diary) and over a month (survey). It would be helpful to standardize the two for comparison – I realize this is not ideal, but would make the discussion on page 10 on pattern of morbidity more transparent.
c. The authors report that the figure 4 shows evidence of a "modest negative wealth gradient" using the diary, and a small positive relationship with wealth. I would recommend that the authors demonstrate some statistical test of significance before one can make this inference. Looking at numbers in the figure, the % of households that report a visit in 6 months in the diary seems very flat. The number of visits per year appears to vary between 2.6 and 2.4 visits per year. Similarly, for household survey, the visits per year appear to vary from just below to just above 0.2. (I assume this is % of households that report any visits in last year?)

d. On the same point, the authors mention about a reversal in gradient in the discussion section. Without a significant difference (in magnitude and statistical significance) in levels of utilization, it might be premature to draw conclusions about a gradient or its reversal.

e. The authors also mention that the patterns observed raise question about positive income elasticity of demand for healthcare. I am not sure the paper presents adequate analysis to support this conclusion. Looking only at the data from diary, figure 3 shows a small negative gradient for annual morbidity, and figure 4 shows a relatively flat relationship between visits per year and wealth quintiles (between 2.6 and 2.4). Taken together, wouldn't this suggest the opposite: the rich are less sick, but tend to use about the same level of healthcare.

f. Table 2 probably has a typographical error .. the CI for any clinic visit in Panel A appears to include a negative sign before the 0.006

g. Overall, I would suggest focusing this paper on recall issues to keep it tight and avoid unnecessary distraction. The figures on wealth quintile morbidity or utilization patterns does not add to the understanding of recall problems in surveys.

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.

Declaration of competing interests: No competing interests.