Reviewer's report

**Title:** Determining the interviewer effect on CQ Index outcomes: a multilevel approach

**Version:** 2  **Date:** 12 February 2010

**Reviewer:** Dario Gregori

**Reviewer's report:**

The work is aimed at describing whether experienced interviewers can influence the perceived quality of healthcare services reported by residents of nursing homes and homes for the elderly. In order to achieve the goal, a multilevel analysis was used to estimate the proportion of variance in quality-of-care satisfaction (CQ index sub-domains scores) determined by the interviewers.

The statistical analysis lacks in many points.

First of all, it is not so obvious the interpretation of the Intra-class correlation coefficient (ICC). You should discuss it at the light of multilevel models.

The ICC can be interpreted in several ways. First, it represents the degree of common environments that observations share. It would increase if observations in the same cluster were under more similar environments and, as a result, if the responses of observations became more alike. Another interpretation of the ICC is the proportion of total variance that is attributed to the cluster level. Therefore, as the relative variance of the clusters increases, the less likely you are to assume that the groups are similar. A third interpretation is that the ICC represents the degree of homogeneity among observations at the cluster level. In this sense if observations were not correlated, they would not affect one another nor would they be similar. Finally, which is your interpretation as I intended, the ICC is the correlation between two observations that are randomly chosen from the same cluster (e.g., correlation of respondents who are interviewed by the same persons).

In this sense, it was interesting to compare the raw ICC (ICC from an empty model with random intercept) along with the ICC adjusted for respondents and interviewers' characteristics. It is reasonable to think that adjusted ICCs are lower than raw ICCs, probably because of differences in case mix. It is likely that respondents within the same health-care institution agree more on satisfaction. Why you did not consider adding to the multilevel model the health-care institution level?

A chi square test for the statistical significance of the ICC should be performed and ICC should be discussed at the light of their statistical significance.

Instead of reporting variances explained by the models, it was more interesting a table with all coefficients and standard error of the covariates in the models.

Applying a multilevel model require some assumption are satisfied. Residual
analysis usually is carried on to check that residuals are normally distributed. If this not the case, because scores are skewed, you have to transform your data. This point was not covered in your analysis.

Minor comments

For clarity, patients' socio-demographic information along with their length of stay, health status, and type of care should be included all in table 2. Be careful that categorical variables such as sex (women/male) do not require standard deviation as reported in the first line of the result section. In these table (and the analogous table 3 with interviewers' characteristics) consider add absolute frequency also.

Be more precise about model fitting: statements as "saw a substantial decrease in model fit -2 log likelihood" (pg17, row 17) is too much general. Consider to provide exact data in tables and then discuss it

**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.