Reviewer's report

Title: Factors associated with data quality of the Routine Health Information System in Benin

Version: 2 Date: 21 April 2014

Reviewer: Francoise Renard

Reviewer's report:

The authors have included the suggested corrections/precisions; the addition of table 2 greatly improves the readability.

Minor revisions:

1. It seems that there is a contradiction between the description of the results of the table 1 and the table itself with regards to supervision of data quality, for which a p=0.651 is seen in the table vs 0.009 in the text; probably due to a shift in the lines?

2. Please refer to table 2 at the beginning of the section Univariate Results, rather than after the second variable.

3. The table 2 could include the results for the 2 variables that you show in the graphs, it would be clearer for the reader.

4. It would also be clearer if you could present in a table the 2 variables that you analyse with a score, either in a separate small table, either in a subsection of table 2.

5. Is it so that the probability of batch rejection was lower for people with a lower perceived self-efficacy (HR=0.7, p=0.052). With a p so close to the significance level, it is difficult just to affirm that those results are not significant (even if stricto sensu they are not). This result is contra-intuitive (we would have expected a better result for the people with a higher perceived self-efficacy); moreover when you analyse the same variable on a continuous scale (score), you get to the inverse conclusion (perceived self-efficacy is associated with better results). I was wondering if there was not an inversion in the table, otherwise, do you have an idea why, when treating the variable in 2 different ways, you find results going in opposite directions.

6. For work engagement, I suppose you find consistent results between the analyse of the dichotomized variable (HR=1.56, p<0.01), and the analyse of the continuous variable - even if the significance level is borderline; please could you show those results, cf remark 3. I would suggest to express the last sentence of &2 p11 as: the average score for work engagement was xxxx within the group of health workers whose batches were not rejected, versus yyyyy in the group whose batches were rejected (p=0.06).

Level of interest: An article of importance in its field
Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

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