Reviewer's report

Title: Intracluster correlation coefficients for the Brazilian Multicenter Study on Preterm Birth (EMIP): methodological and practical implications

Version: 1 Date: 14 January 2014

Reviewer: Mohamed Shoukri

Reviewer's report:

Background: Cluster-based studies in health research are increasing. An important characteristic of such studies is the presence of intra-cluster correlation, typically quantified by the intra-cluster correlation coefficient (ICC), that indicate the proportion of data variability that is explained by the way of clustering. The purpose of this manuscript was to evaluate ICC of variables studied in the Brazilian Multicenter Study on Preterm Birth.

Major Compulsory revisions:

1- Confidence intervals on the ICC depend on the standard errors, which in turn depend on the type of variable being measured. Therefore, for normally distributed data we have an expression for the SE that is very different from the SE of ICC when the data are binary. The SE expression for both cases should be shown.

2- There are no methodology presented in the paper

3- I took the liberty of calculating the DFF for some of the variable, and to my disappointment, many were wrong.

4- If you plan to conduct a cluster-sampling based study, which includes many covariates, it was reported that the estimated ICC is affected by the number of covariates and the way such covariates are measured (on the cluster level, or individual within cluster level).

Discretionary revisions:

1- It might be useful to see how the cost of obtaining clinical or laboratory data affect the design effect.

2- The authors may want to look at the paper:

Contemporary Clinical Trials 36 (2013) 244–253

Minor revision: None
Level of interest: An article of limited interest

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

Declaration of competing interests:

'I declare that I have no competing interests'