Reviewer's report

Title: Improved ability of biological and previous caries multimarkers to explain and predict caries disease as revealed by multivariate PLS modeling

Version: 1 Date: 5 June 2009

Reviewer: Loc Do

Reviewer's report:

The manuscript titled ‘Improved ability of biological and previous caries multimarkers to explain and predict caries disease as revealed by multivariate PLS modeling’ attempted to use partial least square (PLS) modelling in predicting caries. In my opinion, the following issues need to be addressed.

Major Compulsory Revisions

1. The PLS is a method developed for use in econometrics and chemometric. It has not been popular in medical research. The vast majority of readers of medical and health literature are not familiar with this method. It is assumed that even fewer among readers of dental literature. Therefore, it is important for the authors to provide detailed description of the method and to describe the results in an understandable fashion. The manuscript in the current form is very difficult to appraise.

2. The PLS is designed for predictive models with high number of predictors where multicollinearity is high. Its main purpose is for prediction but it cannot offer explanatory modelling. The authors should take that in mind in the text. For example, the title and the pages 11-12 mention ‘explain’. It should be avoided in order not to confuse readers.

3. In the same fashion, values of R-squares need to be discussed. According to Chin (1998), the highest R-square values reported here are of moderate strength.

4. The VIP (variable importance/influence in projection) needs to be described in full before first use.

5. Although the PLS is capable of handling large number of predictors, appropriate sample size is also an important issue. Have the authors calculated the sample size required for this analysis?

6. The last paragraph on page 4, the last two sentences do not belong here. It must be in Method and Result.

7. Was the data collected in 1985 (page 6)?

8. The authors measured caries incidence and progression at tooth surface level. There can be an issue with clustering effect within individual on those dependent variables.

9. The authors should, in my opinion, report the distribution of common
individual-level factors such as oral hygiene practices and use of dental preventive services. If those factors were not equally distributed among the 30 children then it should be controlled for.

10. Any limitations of the method need to be properly discussed.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Needs some language corrections before being published

**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.

**Declaration of competing interests:**

'I declare that I have no competing interests'