Reviewer's report

Title: Comparison of logistic regression and classification and regression tree model: An application in identifying high-risk populations in alternative tobacco product use

Version: 2  Date: 26 December 2014

Reviewer: Haider Mannan

Reviewer's report:

Review of: 'Comparison of logistic regression and classification and regression tree model: An application in identifying high-risk populations in alternative tobacco product use'

This is an interesting paper examining the predictors of tobacco use and tobacco use and use of alternative tobacco products. The question posed by the authors is well defined. It achieves this objective by applying two statistical methods—one ordinary logistic regression and the other CART. For logistic regression, model building is made using a method suggested by Frank Harrell. In my opinion, this approach is alright although the LASSO would have been a better choice. It is not clear what the statement “logistic regression involves more distributional assumptions” means. To my knowledge, distributional assumptions for logistic regression are quite robust. This sentence should be rephrased or totally removed. The models selected by CART are more parsimonious than those selected by logistic regression and also includes interactions which are clinically relevant. The data used in this paper is sound.

The figures appear to be genuine, i.e. without evidence of manipulation. The manuscript adheres to the relevant standards for reporting and data deposition. The discussion and conclusions are well balanced and adequately supported by the data. However, the limitations of the work are not clearly stated. The authors do not clearly acknowledge any work upon which they are building, both published and unpublished. The title and abstract accurately conveys what has been found. Finally, the writing is acceptable.

In conclusion, I would recommend accepting the paper for publication subject to some minor changes. First, the authors need to clearly acknowledge any work upon which they are building. Secondly, the limitations of the work should be clearly stated in a separate section prior to the discussion.