Reviewer's report

Title: Systematic analysis of 18F-FDG PET and metabolism, proliferation and hypoxia markers for distinction of head and neck tumors

Version: 1 Date: 27 January 2014

Reviewer: Eirik EM Malinen

Reviewer's report:

Review of "Systematic analysis of 18F-FDG PET and metabolism, proliferation and hypoxia markers for distinction of head and neck tumors" by Hoeben and Starmans et al.

This is a well written account on FDG-PET and immunohistochemistry (IHC) of 14 different head and neck (HN) cancer xenograft lines in nude mice. The aim of the study was to assess if PET and IHC parameters could be used to identify the different tumor lines. The methodology used, both for the experiments and analysis, are really be state of the art. My minor objections to this study, which should be commented in the revised manuscript, are the following:

- Why is it important to distinguish the cell lines? In cancer therapy, it is more relevant to identify chemo- or radioresistant lines. I guess many of the xenografts lines studied here shows about the same therapy response, and it would be more interesting to cluster them according to such features. Therefore, I didn't find the research issue that interesting, although "treatment allocation purposes" was indeed mentioned.

- As was noted by the authors, the SUV levels in the tumors was very low, and varied between 0.8 and 1.8 in population-based mean SUVmax (table S2). This range of values is far below what is seen clinically in HN cancers. Therefore, with such low absolute values and so little relative variation between the different cell lines, I do not understand why the authors hypothesized (implicitly) that "FDG can be used to distinguish different xenograft lines". With such low inter-line variability, I guess you don't need the "random forest" algorithm to understand that FDG-PET is useless for the purpose addressed. Thus, the starting point of the investigation seems a little artificial.

- What do we learn from the current study? That different tumor cell lines show different levels of biological markers of relevance for cancer phenotyping, and that we need a "full" IHC screening for accurate prognostication? But how dependent are the findings on the classification methodology used? Is the "random tree" method as such something that the authors recommend to be used in future clinical trials, or was this method a part of the experimentation? Such classification methods have their advantages for sure, but the importance and understanding of the underlying biological features somehow drowns in such a "black box" approach. Furthermore, the manuscript title should contain
“classification”, as this is a key methodology.

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