Reviewer’s report

**Title:** Development and validation of a novel MR imaging predictor of response to induction chemotherapy in locoregionally advanced nasopharyngeal cancer: a randomized controlled trial substudy (NCT01245959)

**Version:** 0  **Date:** 24 Apr 2019

**Reviewer:** Frederique Frouin

**Reviewer's report:**

General comments

This paper aims at predicting the impact of induction chemotherapy (ICT) in locoregionally advanced nasopharyngeal cancer using a radiomic signature (ICTOS) based on MRI data. This study was motivated by the following previous results: 1) a global improvement of FFS was observed with ICT, when compared to CCRT; 2) a more specific analysis has shown that some patients could benefit from the ICT while some others would not benefit from such a therapy (due to some adverse effects).

Thus the aim of the present study was to find out a predictor, which could help in defining a category of patients, which could benefit from ICT. Results shown in the present paper suggest that ICTOS, a radiomic signature that is newly proposed in this paper, could help in defining such a subgroup of patients.

This paper presents several positive points, which deserve to be underlined. First of all, the total number of cases is quite high (502) and the validation cohort (248 patients) is as large as the training data set, which is noticeable and gives results high significance. Moreover, the radiomic signature, which is defined in this study, is explicitly defined in the paper. This is a very important contribution, which makes the study easily reproducible by other groups. Too many radiomic papers did not publish their predictive model; for that reason the publication of signature should be rewarded. Furthermore as this signature depends on a reduced number of parameters (three parameters derived from contrast enhanced MRI), the proposed model is assumed to be robust.

However there are some issues, which need to be answered before publication. In addition, some missing information needs to be introduced.

First of all, the definition of ICTOS appears to be erroneous. Is the exponential function really used? In those conditions, it is not possible to consider positive and negative values of ICTOS. I can guess that ICTOS refers to the exponent part. In such a case, the exponential function would not be useful. Furthermore, when looking at values that are reported in Figure S5, ICTOS term seems to be simply: 

\[-0.668 \times \text{skewness} - 0.442 \times \text{GLCM variance} + 0.410 \times \text{GLRLM}_{\text{LRHGLE}}\]

Please clarify that point.
The intra- and inter-observer variability of radiomic features according to the segmentation has been studied through repeated segmentations, which is valuable. However, it would be also important to define the reproducibility of the parameters of interest (skewness, GLCM variance, and GLRLM_LRHGLE) according to each MR scanner and MR acquisition protocol. Finally, it would be better to introduce the definition of the ICTOS in the main text, since it is one of the main contributions of the paper.

MR data: Are coronal and sagittal slices necessary? Indeed, they do not seem to be further analyzed. Appendix S2 indicates that several MR scanners were used for the study. Could you give the distribution of the patient acquisitions according to the different scanners, for both the training set and the validation set? Could you also introduce the in-plane voxel size? Indeed, several studies have shown that these spatial resolution parameters have an impact on some radiomic features, especially on texture parameters. It would be interesting to report the variability of the ICTOS according to the different acquisition MR parameters and scanners. This is

How were texture parameters computed: which value is chosen for Ng? How are minimal and maximal values chosen? Which bin size is used for the computation of co-occurrence and other matrices? The standardization of MR intensity that is proposed is interesting in that sense that it provides a solution to compare different exams. However could you precise the rationale for such a procedure? Could you briefly demonstrate the impact of such a procedure?

Constitution of the original training cohort
There are several inconsistencies in the text, which needs to be addressed. Figure 2 indicates that patients in the training cohort have either events or more than 3 years of follow-up. However 5-years of follow-up are indicated in the text. Thus 3 years of follow-up is not adapted to the study. Furthermore it is indicated that last follow-up was on December 30, 2017. Could this be updated? Furthermore plots in Figures 3, 4, S3, and S4 show results with more than 60 months of follow-up. For sake of simplicity, I suggest displaying results from 0 till 60 months.

Minor points
Please introduce the term of external validation cohort (instead of validation). Give the reasons why only a part of the clinical trial was used to define the validation cohort (why not using the whole trial?).

Page 13, lines 2-3: remove the sentence, which is already written in the Procedures subsection

Page 13, lines 15-17: consider the true number of patients (116.2 and 131.6 are irrelevant values). Furthermore which group do you consider here: training or validation? The total number of patients being equal to 247.8, this result generates some confusion.

Figures: detail the legends, which are very brief.

Figure 1: annotations: CT is not suitable (replace it with MR, 4 occurrences) Illustration of "Radiomic features extraction" (Panel c) is not relevant: why are colored maps displayed? Which parameters do they illustrate? Panel d is also difficult to understand without detailed legend, and in its present form does not bring some useful information.

Figure 2
More than 3-years follow-up are indicated in the Figure 2. In the text, 5-years follow-up is considered. Could you explain this inconsistency?

Figures 3 and 4: Please consider the x-axis, representing time in months: I assume that the labels ‘26’ should be ‘36’
Same remark for Figures S3 and S4

Some points are introduced as supplemental data, but they could be reduced, since they are already published, and not new. Consider for instance IPTW.

**Are the methods appropriate and well described?**
If not, please specify what is required in your comments to the authors.

No

**Does the work include the necessary controls?**
If not, please specify which controls are required in your comments to the authors.

Yes

**Are the conclusions drawn adequately supported by the data shown?**
If not, please explain in your comments to the authors.

Yes

**Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?**
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I recommend additional statistical review

**Quality of written English**
Please indicate the quality of language in the manuscript:

Acceptable

**Declaration of competing interests**
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?
2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

Statement on potential review bias
Please complete a statement on potential review bias, considering the following questions:

1. Did you co-author any publication with an author of this manuscript in the last 5 years?

2. Are you currently or recently affiliated at the same institution as an author of this manuscript?

If you can answer no to all of the above, write 'I declare that I did not publish with these authors in the last 5 years and also meet the affiliation criteria". If your reply is yes to any, please give details below.

I declare that I did not publish with these authors in the last 5 years and also meet the affiliation criteria

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal