Reviewer's report

**Title:** Semi-Automated Analysis of Standard Uptake Values in Serial PET/CT Studies in Patients with Lung Cancer and Lymphoma

**Version:** 1  **Date:** 1 December 2011

**Reviewer:** Eric Laffon

**Reviewer's report:**

# Report:

This paper aims to assess a semi-automatic method for calculation of SUVmax in serial PET/CT studies in comparison to a conventional manual method. This reviewer agrees that "the demand for quantitative monitoring of SUV has grown as the indications for performing follow-up 18FDG PET/CT are expanding". Therefore, for saving time in current clinical practice, the aim of this study is valuable. However, this paper suffers from deficiencies that are mainly related to inappropriate statistical analysis.

- **Major Compulsory Revisions:**

**Abstract:**

- methods: indicate the number of readers in this section.
- results: statistical analysis should be totally reconsidered (see below).

**Introduction:**

- The second paragraph could be developed.

**Materials and Methods**

- Manual method: when comparing two PET examinations, was SUVmax always found in the same slice? Indeed, scanning the whole lesion in each examination often shows that the slice involving SUVmax can be different (even for 2 scans successively performed).

- Semi-automatic method: the way this method is implemented introduces a bias in the study, for the above mentioned reason. In other words, the automatic alignment does not take into consideration that SUVmax could be found in a different (contiguous) slice than that marked in the first examination. (Note that a semi-automatic volumic assessment could avoid such a bias.)

- Statistical analysis: should be totally reconsidered by using the method of Bland and Altman, because linear regression is not sufficient to assess the agreement between 2 methods. (Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1:307-310.)
Results
- Linear regression analysis: this reviewer is not sure that there is no confusion between "intercept" and "slope" (please give additional details, such as equations). However, as mentioned above, the analysis is not appropriate.
- Range of SUVmax values: the lowest SUVmax value (2.4) is too low. A cut-off of 2.5 is usually used. Means should also be given.
- Table 1:
  . This reviewer does not understand why the column "manual" shows different figures between readers. Indeed, manual measurements should mandatorily provide similar results! (as much as "One of the readers marked the 26 lesions in screenshots …to secure that all three readers were measuring the same lesions")
  . Last column: what is "p" related to ? Moreover, the second reader found a significant difference between the two kinds of assessment that is not clearly emphasized in the results.
  . "The differences in manual measurements are systematically lower than the semi-automated ones". Why don’t you use a simple sign test to ascertain it?
  . Give unit of SUVmax.

Discussion and Conclusion
- The discussion is not structured because the results are unclear (or clearly understood).
- The first sentence of the conclusion summarizes the deficiencies of the paper: " … but there was a significant difference in the mean difference in SUVmax between studies when comparing both methods". This point is neither clearly presented in the results section nor discussed, whereas it is the main interest of the paper.

Minor Essential Revisions:

Title and Manuscript
- Please choose between "semi-automated" and "semi-automatic" and keep on your choice throughout the manuscript.

Materials and Methods
- An estimate of the effective dose (mSv) might be given for both PET and CT.

Results
- I suppose that Figure 1 is flowchart and Figure 2 is erroneously labelled "Figure 1".
- Second paragraph: "almost-perfect reproducibility …" perfect is awkward.