Reproducibility of Uptake Estimates in FDG PET: a Monte Carlo study

Juliette Feuardent, Marine Soret, Irène Buvat[1]

Abstract—**Tumor glucose metabolism measurements from Fluoro-deoxyglucose (FDG) Positron Emission Tomography (PET) could be helpful for therapeutic follow-up, assuming the index used is highly reproducible in the absence of change in tumor metabolism. We compared the variability of uptake values as estimated by Tumor-to-Normal tissue Ratio (TNR) depending on whether they are calculated from maximum value (TNRmax) or from the mean value (TNRavg) in a volume of interest (VOI).**
Methods: Monte Carlo simulations of CPET acquisitions using a Data Spectrum thorax phantom including eight spheres mimicking tumors with various sizes and locations were performed with SimSET. The phantom and acquisition protocols were designed to study the differences in TNRavg and TNRmax estimates when there were changes in tumor size, tumor-to-background activity contrast, tumor uptake, noise level in the image, or VOI location with respect to the tumor. Twenty realizations of each configuration were simulated.
Results: Both TNRavg and TNRmax were highly sensitive to a 50% change in tumor uptake. However, uptake estimates also differed by more than 25% when the tumor uptake did not change but the tumor size decreased by 30, 50 or 70%. With TNRavg, this dependency upon tumor size could be almost fully removed using partial volume effect correction. When only the non-specific activity changed, TNRmax could vary by as much as 50% for small spheres, while TNRavg variations remained less than 25%. TNRmax did not depend on the VOI location with respect to the tumor, while the reproducibility of TNRavg was highly dependent on the correct positioning of the VOI with respect to the tumor. TNRavg and TNRmax were both sensitive to the noise level in the images, suggesting that repeated scans should be performed with similar counting rate and duration.
Conclusion: Overall, TNRavg appears less sensitive to changes different from tumor uptake variations and should be preferred to TNRmax in the context of therapeutic follow-up, provided a correction for partial volume effect is used and the positioning of the VOI with respect to the tumor is precisely controlled.

I. Introduction

Quantitative measurements of tumor glucose metabolism from FDG PET should be helpful for therapeutic follow-up. In that context, the index used for characterizing the tumor metabolism should be highly reproducible in the absence of change in tumor metabolism, while being highly sensitive to a change when it actually occurs. Several indices for characterizing glucose metabolism in tumor have been proposed [1-3]. An ongoing question related to the definition of an index is whether tracer uptake should be measured from the averaged activity concentration in a volume of interest (VOI) or from the maximum activity concentration in that volume. The purpose of this study was thus to compare the variability of uptake values as estimated by Tumor-to-Normal tissue Ratio (TNR), depending on whether they are calculated from the maximum pixel value (TNRmax) or from the mean value (TNRavg) in a VOI. We looked at which differences in TNRmax or TNRavg could be observed when one of the following parameters changed between two scans: the tumor uptake, the tumor size, the non-specific activity in normal tissue, the VOI location with respect to the tumor and the noise level in the image. The consequences of our findings for the use of TNR in the clinical context of therapeutic follow-up in oncology are discussed.

II. Material and Methods

A. Phantom.

A thorax phantom (Data Spectrum Corporation, Hillsborough, NC) including lungs, soft tissues and eight spheres was considered. The 8 spheres consisted in two identical sets of 4 spheres (10.5, 16, 22 and 33 mm in diameter). One set was placed in the lung while the other was placed in the mediastinum (Fig. 1).

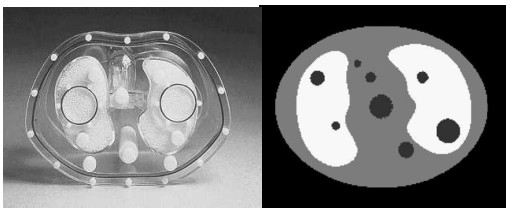

Fig.1. Data Spectrum thorax phantom and location of the spheres inserted in the phantom.

B. CT acquisition.

A CT acquisition of the phantom with the sphere in place was performed using the Lightspeed scanner (General Electric Medical Systems, Milwaukee, WIS). The spatial sampling of the CT data was 1.56 mm x 1.56 mm x 1 mm, where 1 mm

[1] J. Feuardent and I. Buvat are with INSERM U494, CHU Pitié-Salpêtrière, 91 bd de l'Hôpital, 75634 Paris cedex 13, France. M. Soret is with the Nuclear Medicine Service of H.I.A. Val-de-Grâce, bd de Port-Royal, 75005 Paris, France, and with INSERM U494, CHU Pitié-Salpêtrière, 91 bd de l'Hôpital, 75634 Paris cedex 13, France.

corresponded to the axial sampling. This high resolution CT scan was used to define the activity map and the density map needed as input for the SimSET Monte Carlo simulation code. Only the lung and soft tissues compartments were segmented from the CT, while the tumors were added a posteriori as spheres with perfectly known inner diameters.

C. Simulations.

Simulations were performed with the Monte Carlo simulation code SimSET [4], previously validated for 3D PET acquisitions on the CPET (ADAC-UGM/Philips, Philadelphia, PA) [5]. A 2.25 mm x 3 mm x 4 mm spatial sampling was used when simulating the sinograms. The size of the sinograms was 256x192x64 pixels. Two series of simulations were performed, with two different uptake levels in tumors: the tumor-to-soft tissues uptake ratio was 8 in the first series, and 4 in the second series. In both series, no activity was included in the lungs. For each level of tumor uptake, 20 realizations of the phantom acquisition were simulated, changing only the seed of the random generator in SimSET for each simulation. A total of 4.10^9 events were generated per simulation, yielding sinograms including 22.6 million detected counts.

D. Data processing.

1) RECONSTRUCTION. Sinograms were reconstructed using Ordered Subset Expectation Maximization (OSEM) with 8 subsets and 6 iterations after Single Slice Rebinning, using a home-made reconstruction software. Because spatial resolution in the reconstructed images is an input parameter for partial volume effect correction (see subsection 3 below), spatial resolution was visually assessed by comparing the reconstructed images with the activity map corresponding to the simulation input convolved with a 3D Gaussian functions of FWHM varying between 8 and 14 mm. The estimated spatial resolution was found to be 12 mm. The reconstructed images were then interpolated to the CT spatial sampling, namely 1.56mm x 1.56mm x 1mm, yielding images with a size of 256x256x78 pixels.

2) ATTENUATION CORRECTION. All images were corrected for attenuation using the attenuation map obtained as follows: theoretical 511 keV attenuation coefficients were assigned to the lung compartment (μ = 0.035 cm^{-1}) and to the soft tissue and sphere compartments (μ = 0.096 cm^{-1}); the resulting attenuation maps were blurred with a 3D Gaussian filter to match the spatial resolution of the reconstructed emission data, namely 12 mm.

3) PARTIAL VOLUME EFFECT (PVE) CORRECTION. A PVE correction was implemented, requiring the definition of VOIs corresponding to the 11 compartments describing the phantom: 1 for each of the 8 spheres, 1 for the lungs, 1 for the soft tissues and 1 for the region outside the phantom. The method [9] consisted in inverting an 11 x 11 cross-contamination (CC) matrix describing the spill-in and spill-out between the 8 spheres, the lung and the soft tissues compartments. The CC matrix coefficients were obtained by calculating the attenuated projections of each compartment as identified on the CT map, by blurring the resulting projections with a Gaussian filter of 11 mm FWHM mimicking the detector response and by reconstructing the resulting projections with attenuation correction using the same reconstruction scheme as that used for the PET images. It was checked that the spatial resolution in the image of each compartment reconstructed after an 11 mm blurring of the projections was similar to that of the reconstructed phantom images (i.e., about 12 mm).

E. TNR measurements. Two TNR indices, TNRavg and TNRmax, were calculated by considering the mean and maximum values in the 'tumor' VOIs. The tumor VOIs were taken as the sphere of exact volume inserted in the CT input image of the simulation. The VOI used to estimate the normal value was defined in a soft tissue region chosen to be large enough to have sufficient statistics and far enough from other compartment boundaries to avoid introducing any PVE. For both TNRavg or TNRmax, the 'normal' uptake value was the average value in the VOI drawn over the normal soft tissue.

F. Data analysis. Five situations of change were considered, corresponding to situations that could possibly be encountered during therapeutic follow-up.

1) A change in tumor uptake during therapy was modeled by comparing the dataset corresponding to a TNR of 8 in the spheres with the dataset corresponding to a TNR of 4 in the spheres. The "starting point" (TNR1 in equation (1) below) was TNR = 8, so that the modeled change mimicked an actual response to therapy (decrease of tumor uptake by 50%).

2) A change in tumor size in the absence of change in tumor uptake was modeled by comparing the uptake values measured in the spheres of different sizes. We studied three cases: sphere diameter decreasing from 33 mm to 22 mm (33% decrease), from 33 mm to 16 mm (52% decrease) or from 33 mm to 10.5 mm (68% decrease). The "starting point" was always the TNR in the 33 mm sphere, so that the changes modeled a shrinking of the tumor without change in uptake of the remaining tumor.

3) A change in non-specific activity was modeled by comparing the TNR of the spheres located in the lungs (no surrounding activity) with those of the spheres located in the mediastinum (relative surrounding activity of 1). The "starting point" was taken as the spheres located in the lungs, mimicking an increase in non-specifc activity over the course of the therapy.

4) A change in the VOI location with respect to the tumor location. The sphere VOI was shifted by 3 pixels in all directions (corresponding to 4.7, 4.7 and 3 mm shifts in the transaxial and axial directions respectively). This mimicked a misregistration of the VOI drawn on a first study when applied on the images from a second study.

5) A change in the total counts in the images. We compared the results obtained using 22.6 million detected counts with those obtained when only 5.6 million counts were detected (corresponding to 10^9 simulated events). The noise increase in

the "second" scan compared to the "first" scan, with noise computed as the ratio between standard deviation and mean in a VOI drawn over soft tissues, was of 30%.
For each simulated change (in tumor uptake, in tumor size, in non-specific activity, in VOI location or in noise level), the TNR variation between the two states 2 and 1 (1 corresponding to the starting point) was calculated as (in %):

$$100 \times (TNR_2 - TNR_1) / (TNR_1). \quad (1)$$

The mean and standard deviation of the TNR variation over the 20 realizations were computed.

III. Results and Discussion

To interpret the results in a practical context, we referred to the EORTC recommendations [7]. These recommendations state that a progressive metabolic disease would be characterized by an SUV increase of more than 25%, that a stable metabolic disease would correspond to an SUV increase less than 25% or to an SUV decrease of less than 15%, and that a partial metabolic response would be characterized by an SUV decrease between 15% and 25% after one therapy cycle, or by an SUV decrease of more than 25% after more than one therapy cycle.

A. Changes in tumor uptake. When the TNR changed from 8 to 4, mimicking a response to therapy for instance, with a 50% decrease in uptake, TNRavg changes were between 26.4% and 54.9%, while TNRmax changes were between 31.8% and 63.2% (Fig. 2). Both indices were thus sensitive to large changes in tumor uptake, and the change of TNRavg was slightly less dependent on tumor size and tumor location than the change of TNRmax.

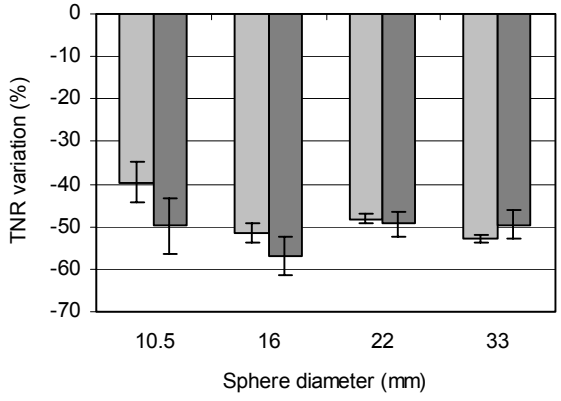

Fig.2. TNR variation (in %) for a change in tumor uptake from 8 to 4. Results are shown for TNRavg (in light grey) and for TNRmax (in dark grey). Bars and error bars represent the average and standard deviation over the 20 realizations respectively.

B. Changes in tumor size. When the tumor decreased in size (Fig. 3) without any change in tumor uptake, TNRavg decreased. The decrease was from 20% to 27%, from 32% to 38% and from 52% to 65% when the 33 mm tumor had its diameter go down to 22 mm, 16 mm and 10.5 mm respectively. TNRmax also decreased, from 12% to 30%, from 20% to 36% and from 42% to 63% when the 33 mm diameter tumor had its diameter decrease to 22, 16 and 10.5 mm respectively. This suggests that variations greater than the 15% and 25% thresholds can be observed, not as a result of an actual change in tumor metabolic activity, but just because the tumor size decreases in time. This is because of PVE, which makes the apparent tumor uptake dependent on the tumor size. This dependency was actually removed when applying PVE correction (Fig. 3). After PVE correction, changes in TNRavg were less than 10% when the tumor size decreased from 33 mm to 10.5 mm while activity concentration remained constant in the tumor.

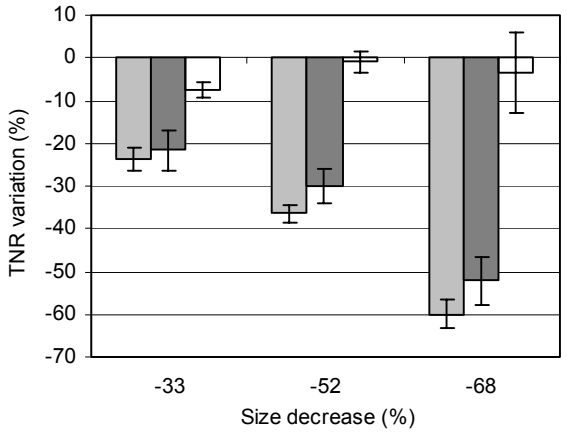

Fig.3. TNR variation (in %) for a change in tumor size from 33 mm to 22 mm, 16 mm and 10.5 mm. Results are shown for TNRavg (in light grey), for TNRmax (in dark grey), and for TNRavg after PVE correction (in white). Bars and error bars represent the average and standard deviation over the 20 realizations.

C. Changes in non-specific activity. When normal tissue activity, i.e. non-specific activity, increased from 0 to 1, TNRavg changes were between 0% and 24%, while TNRmax changes were between 0% and 52% (Fig. 4). TNRavg was thus much less sensitive to the level of non-specific activity than TNRmax. Also, the smaller the tumor size, the greater the dependency to non-specific activity.

D. Changes in VOI positioning with respect to the tumor location. When the VOI was shifted with respect to the actual location of the tumor, TNRmax remained unchanged as long as the pixel with the maximum value in the tumor was still included in the VOI (which is basically always the case in practice). Unlike TNRmax, TNRavg was sensitive to the proper positioning of the VOI with respect to the tumor, and dropped by up to 36% for the smallest tumor size (Fig. 5).

Careful re-positioning of a VOI drawn on a scan upon a subsequent scan might be feasible however, for instance using a simple criterion such as maximizing the number of counts within the VOI.

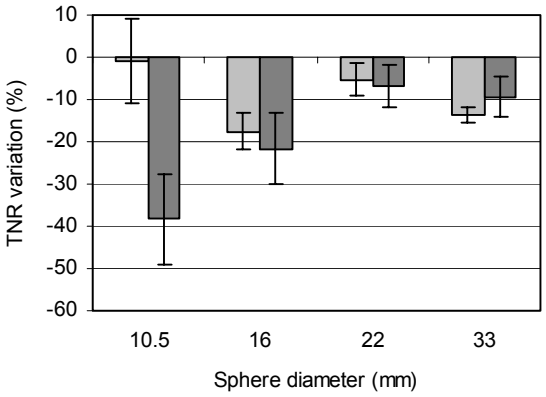

Fig.4. TNR variation (in %) for a change in non-specific activity from 0 to 1. Results are shown for TNRavg (in light grey) and TNRmax (in dark grey). Bars and error bars represent the average and standard deviation over the 20 realizations.

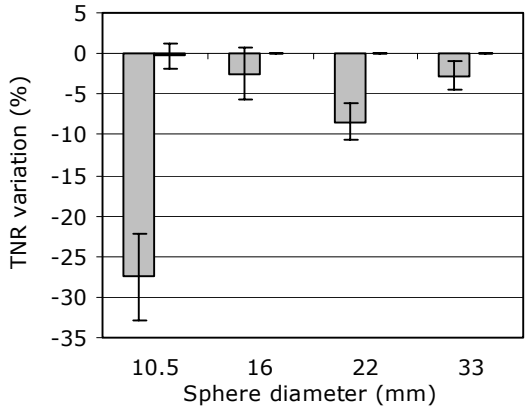

Fig.5. TNR variation (in %) for a change in VOI location with respect to tumor location. Results are shown for TNRavg (in light grey) and for TNRmax (in dark grey). Bars and error bars represent the average and standard deviation over the 20 realizations.

E. Changes in noise level. When comparing TNR measured from images corresponding to two different acquisition statistics, it appeared that TNRavg could changed by as much as 24% (sphere with a 10.5 mm diameter) and TNRmax could change by as much as 52% (Fig. 6).

In summary, TNRmax could change by more than the 15% or even 25% threshold just because of a change in image noise level, tumor size, or non-specific activity. On the other hand, TNRavg could change by more than 15% or 25% because of a variation in image noise level, tumor size, non-specific activity or VOI positioning. These misleading changes can be dealt with more or less easily. For both TNRavg and TNRmax, acquiring similar numbers of events for all scans to be compared should allow one to mostly avoid the apparent changes in tumor uptake that are in fact due to changes in noise level. When considering TNRavg, the dependency between apparent tumor uptake and tumor size can be efficiently removed using a PVE correction (Fig. 3). VOI positioning can also be optimized to reduce this source of variability. When using TNRmax (to avoid the use of PVE correction), there is not much one can do regarding the dependency of TNRmax with non-specific activity and tumor size. TNRavg thus appears more suitable for therapeutic follow-up than TNRmax, as it can be made less sensitive to changes different from changes in activity concentration than TNRmax.

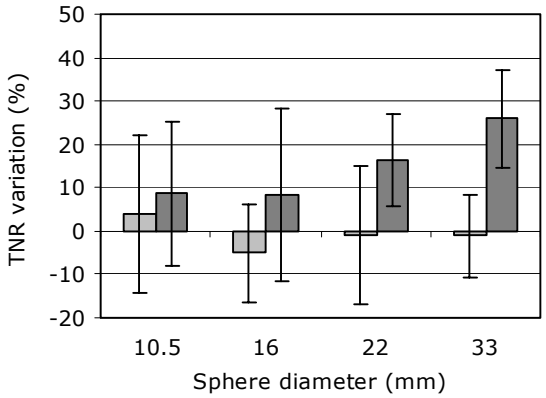

Fig.6. TNR variation (in %) regarding a change in noise level. Results are shown for TNRavg (in light grey) and TNRmax (in dark grey). Bars and error bars represent the average and standard deviation over the 20 realizations.

As TNR is a ratio between two measured values, we also studied the variability of uptake in a 'normal' region. This variability (measured as the ratio between the standard deviation and the mean value within the 'Normal' VOI) was found around 0.4%. As a result, the variability of TNR is mostly caused by the variability of tumor uptake. Percent differences in TNR are then almost identical to percent differences in SUV, and our results would have been almost exactly the same had we considered SUV instead on TNR.

IV. Conclusion

Substantial changes in estimated TNR (TNRmax or TNRavg) can be due to many effects different from actual change in tumor uptake. To reduce the risk of mis-interpreting an observed change, scans performed over time should be acquired at similar count levels, the tumor VOI used for the measurement should be the same and carefully located around the tumor, TNRavg (or SUVavg) should be used instead of TNRmax (or SUVmax), and finally, partial volume effect correction should be performed.

V. REFERENCES

[1] Weber WA, Ziegler SI, Thödtmann R, Hanauske AR, Schwaiger M. Reproducibility of metabolic measurements in malignant tumors using FDG PET. *J Nucl Med* 1999; 40: 1771-1777.

[2] Nakamoto Y, Zasdany KR, Minn H, Wahl RL. Reproductibility of common semi-quantitative parameters for evaluating lung cancer glucose metabolism with positron emission tomography using 2-deoxy-2-[^{18}F]Fluoro-D-glucose. *Molecular Imaging and Biology* 2002; 4:171-178.

[3] Paquet N, Albert A, Foidart, Hustinx R. Within-patient variability of ^{18}F-FDG : standardized uptake values in normal tissues. *J Nucl Med* 2004 ; 45 : 784-788.

[4] Harrison RL, Vannoy SD, Haynor DR, Gillispie SB, Kaplan MS, Lewellen TK. Preliminary experience with the photon history generatot module of a public-domain simulation system for emission tomography. *Conf Rec IEEE Nuclear science symp and Medical Imaging Conf* 1993; 2:1154-1158.

[5] Buvat I, Castiglioni I, Feuardent J, Gilardi MC. Unified description and validation of Monte Carlo simulators in PET. *Conf Rec IEEE Nuclear science symp and Medical Imaging Conf* 2003; M10-184.

[6] Soret M, Ridell C, Hapdey S, Buvat I. Biases affecting the measurements of tumor-to-background activity ratio in PET. *IEEE Trans Nucl Sci* 2002; 25: 2112-2118.

[7] Young H, Baum R, Cremerius U, Herholz K, Hoekstra O, Lammertsma AA, Pruim J, Price P. Measurement of clinical and subclinical tumour response using [^{18}F]-fluorodeoxyglucose and positron emission tomography: review and 1999 EORTC recommandations. *Eur J Cancer* 1999; 35: 1773-1782.