Reliability of SUV estimates in FDG PET as a function of acquisition and processing protocols

J. Feuardent[1], M. Soret[1,2], O. de Dreuille[3], H. Foehrenbach[2], I. Buvat[1]

[1]U494 INSERM, CHU Pitié-Salpêtrière, Paris, [2]HIA Val de Grâce, Paris, [3]Siemens, Saint Denis, France

***Abstract*-- Standardized uptake values (SUV) are commonly used in FDG PET to characterize suspicious high uptakes. To better understand the reliability and the limits of SUV, we studied the accuracy of SUV estimates using phantom data.**

Methods: Using the Data Spectrum thorax phantom in which spheres were inserted, we studied the effect of the sphere sizes, of out-of-the-field-of-view-activity, of the emission scan duration and of attenuation and partial volume effect (PVE) corrections upon SUV biases. Considering a specific acquisition and processing protocol as a reference, we determined the changes in SUV when modifying this protocol.

Results: For a true SUV of 8, estimated SUV were strongly dependent on lesion size, but also on the acquisition and processing protocol. Depending on the method used to derive the attenuation map, estimated SUV could change by more than 50% for lung spheres. Using PVE correction, SUV increased by a factor greater than 2 for spheres less than 2 cm in diameter. The very definition of SUV could also change the SUV by more than 50%. Even with CT-based attenuation correction and PVE correction, SUV estimates remained underestimated by more than 20% in lesions less than 2 cm in diameter.

Conclusion: Biases in SUV estimates strongly depend on the acquisition and processing protocols. This suggests that comparing SUV between studies make sense only if the acquisition and processing protocols are strictly identical.

I. Introduction

Standardized uptake values (SUV) are commonly used in FDG PET to characterize suspicious high uptakes. However, SUV estimates are biased by a number of phenomena, making their interpretation potentially misleading. We started a detailed investigation of the biases affecting SUV estimates using numerical simulations [1], by considering the effects of lesion size, uptake heterogeneity within the lesion, tumor-to-background activity ratio, attenuation, spatial resolution, tumor location and image sampling. The purpose of this work was to further investigate the reliability of SUV estimates using real phantom experiments (and not only simulations) for a better understanding of the limits and reproducibility of SUV. We thus investigated the effect of two phantom parameters (lesion sizes and out-of-the-field-of-view activity), of an acquisition parameter (emission scan duration) and of two processing parameters (attenuation and partial volume effect corrections). We also compared two methods of SUV measurements. Acquisition and processing protocols that minimize the errors in SUV estimates are also suggested.

II. Material and Methods

Phantom. A thorax phantom (Data Spectrum Corporation, Hillsborough, NC) including lungs and soft tissues was considered. Two identical sets of 4 spheres, with diameters of 10.5, 16, 22 and 33 mm, were introduced in the lung and mediastinum compartments. FDG activity concentrations were set to 30.4 kBq/mL in the spheres and 3.8 kBq/mL in the mediastinum while no activity was introduced in the lungs (Cf. Table 1). Knowing the total weight of the phantom, the activity distribution in the phantom could be described in terms of Standardized Uptake Values (SUV) with the following formula:

$$SUV = \frac{\text{Activity concentration in volume of interest (kBq/mL)}}{\text{Injected activity (kBq) / Total weight (g)}}$$

For the considered phantom, theoretical SUV were 8 in the spheres, 1 in the mediastinum and 0 in the lungs (Table 1).

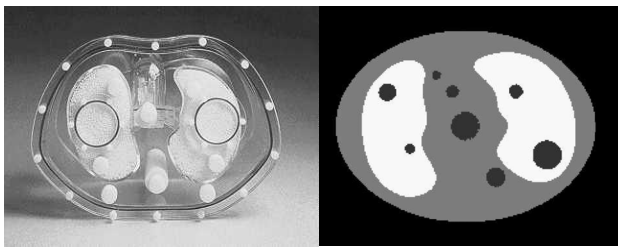

Fig. 1. Data Spectrum thorax phantom and location of the spheres inserted in the phantom.

TABLE I. Activity distribution in the thorax phantom and associated SUV.

	Activity concentration (kBq/mL)	*Theoretical SUV*
Soft tissues	30.4	8
Spheres	3.8	1
Lungs	0.0	0

The impact of out-of-the-field-of-view activity (OFOVA) was studied by adding two activity sources mimicking the brain

0-7803-8257-9/04/$20.00 © 2004 IEEE.

(7.4 kBq/mL in a water cylinder) and the bladder (81.4 kBq/mL in a perfusion bag) (Fig. 2).

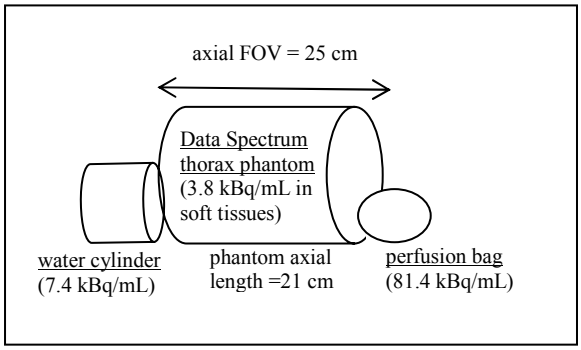

Fig. 2. Activity concentrations in the Data Spectrum thorax phantom and in the out-of-the-field-of-view sources.

PET acquisitions. 3D mode emission acquisitions were performed using the CPET tomograph (ADAC-UGM/Philips, Philadelphia, PA). A 2.25 x 2.25 x 4 mm spatial sampling was used when acquiring the sinograms. The single count rate was around 4.10^6 cps, similar to the count rates observed in clinical acquisitions with the CPET. The impact of the emission acquisition duration on SUV estimates was studied by considering 6 min and 18 min emission scan, respectively. The 18 min emission scans were obtained by summing 3 consecutive 6 min emission scans. All emission data were corrected for "background", supposedly including randoms and scatter, by extrapolating background activity within the phantom from the tails of activity detected outside the borders of the phantom in the sinograms (method implemented on the CPET).

To perform attenuation correction, 2 min transmission scans were also acquired after the emission scan, with the ^{137}Cs transmission source of the CPET system.

CT acquisition. A CT acquisition of the phantom 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 corresponded to the axial sampling.

Data processing.

1) RECONSTRUCTION. Sinograms were reconstructed using Ordered Subset Expectation Maximization (OSEM) with 8 subsets and 6 iterations after Fourier rebinning. The spatial resolution in the reconstructed images was visually assessed by comparing the reconstructed images with numerically simulated images of the phantom filtered with a 3D Gaussian of FWHM varying between 8 and 14 mm. The estimated spatial resolution was 9 mm.

2) ATTENUATION CORRECTION. All images were corrected for attenuation using two different attenuation maps. A so-called Cs map was obtained after "remapping" the ^{137}Cs transmission scan. This "remapping" consisted in assigning theoretical 511 keV attenuation coefficients to the soft tissues compartment ($\mu = 0.096$ cm^{-1}), while the attenuation coefficients measured in the lungs are kept unchanged. The μ values at the border between the soft tissues and lung compartments are interpolated from the neighboring μ values [2]. A so-called CT map was obtained by assigning theoretical 511 keV attenuation coefficients to the lung compartment ($\mu = 0.035$ cm^{-1}) and to the soft tissues and sphere compartments ($\mu = 0.096$ cm^{-1}) as identified from the manual segmentation of the CT into 11 compartments: 1 for each of the 8 spheres, 1 for the lungs, 1 for the soft tissues and 1 for the region outside the phantom. To perfom the CT-based attenuation correction, the CT images were first registered with the PET images at the CT image sampling (1.56 x 1.56 x 1 mm), using an algorithm of a mutual information maximization.

3) PARTIAL VOLUME EFFECT (PVE) CORRECTION. Two PVE corrections were implemented. The first method consisted in multiplying the activity measured in a volume of interest (VOI) drawn around each sphere (see below) by an appropriate recovery coefficient (RC). The RCs were estimated using the size of the spheres as estimated from the manual segmentation of the CT, and considering a 9 mm spatial resolution in the reconstructed images and the true contrast between the sphere and the surrounding tissues (8:1 for the mediastinal spheres and 8:0 for the lung spheres). The second method [3] consisted in inverting a 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 as segmented from the CT data. The CC matrix coefficients were obtained by calculating the attenuated projections of each compartment as identified on the CT map, by blurring the projections with a Gaussian filter of 9 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.

SUV measurements. We systematically measured two SUV indices: SUVavg and SUVmax were estimated by considering the mean and maximum values in VOIs manually defined on the CT data after registration of the CT with the PET data.

Data analysis. We first considered the SUV as estimated using a "reference" configuration, defined as the configuration without out-of-the-field-of-view activity, with a 6 min emission scan, with the Cs map for attenuation correction, without any PVE correction, and when considering SUVavg. We then studied how the SUV estimates changed when increasing the emission scan duration, when out-of-the-field-of-view activity was present, when a CT map was used instead of the Cs map for attenuation correction, when PVE corrections were performed, and when SUVmax was considered instead of SUVavg. Changes in SUV estimates as a function of the imaging and processing protocol and absolute biases in SUV were systematically investigated.

III. RESULTS AND DISCUSSION

A. "Reference configuration"

SUV as estimated in the reference configuration are given in Tables II and III for the lung spheres and the mediastinal spheres respectively. Results were consistent with previous results obtained using numerical simulations [1]: the smaller the sphere, the greater the SUV underestimation (35% and 91% respectively for the largest and the smallest sphere in the lung). SUV underestimations were systematically less severe for mediastinal spheres than for lung spheres. This is because for the lung spheres, PVE affected the Cs transmission measurement as the spheres had a density different from that of the surrounding lung tissue. As a result, attenuation in the lung spheres was underestimated and attenuation correction did not restore enough counts. Because the mediastinal spheres had a density similar to that of the surrounding soft tissues compartment, PVE did not affect the Cs transmission measurements much, and attenuation was more accurately compensated for than for the lung spheres.

TABLE II. SUV AND SUV UNDERESTIMATION IN PERCENTAGE (IN PARENTHESIS) AS ESTIMATED IN THE "REFERENCE" CONFIGURATION FOR THE LUNG SPHERES.

	Spheres in the lungs (Ø in mm)			
	10.5	16	22	33
SUVavg	0.7	2.1	3.6	5.2
Underestimation (%)	91	74	55	35

TABLE III. SUV AND SUV UNDERESTIMATION IN PERCENTAGE (IN PARENTHESIS) AS ESTIMATED IN THE "REFERENCE" CONFIGURATION FOR THE MEDIASTINAL SPHERES.

	Spheres in the mediastinum (Ø in mm)			
	10.5	16	22	33
SUVavg	1.7	3.0	4.9	6.5
Underestimation (%)	79	63	39	19

B. Changes in SUV estimates as a function of the imaging and processing protocol

Out-of-the-field-of-view activity. SUV tended to be greater with out-of-the-field-of-view activity than without (see Tables IV and V, for the spheres located in the lungs and in the mediastinum, respectively). The changes in SUV were never greater than 11% however. The small impact of OFOVA can be explained by the systematic "background" subtraction applied to all acquired data, which subtract both randoms and scatter on the CPET. Our observations suggest that this correction is not ideal, as there is still an impact of OFOVA, but that it is effective at removing part of the signal coming from activity out of the field of view, as the impact of OFOVA, although present, was small.

Duration of the emission acquisition. Lengthening the emission acquisition from 6 min to 18 min did not affect the SUV much: changes in SUV were never greater than 5% (Tables IV and V). These results are consistent with previously reported results [4] in which the bias on the SUV decreased by less than 15% when the emission scan duration changed from 15 min to 5 min. Indeed, the scan duration mostly changes the noise level in the data, while SUV_{avg} is not very sensitive to noise – unlike SUV_{max} – as it is calculated from the average of signal in the VOIs.

Attenuation map. The type of attenuation map affected the SUV only for the lung spheres, with SUV systematically higher (hence SUV biases systematically smaller) when the CT map was used instead of the Cs map (Tables IV and V). The smaller the sphere diameter, the greater the impact of the attenuation map. Because of the limited spatial resolution of the Cs map, PVE affected the Cs transmission measurements. PVE in transmission yielded an underestimation of the attenuation coefficients in the spheres located in the lungs, as counts transmitted through the lung tissues were detected at the sphere locations. Therefore, attenuation was under-corrected in the lung spheres. This effect was not so severe when using the segmented CT map for attenuation correction, since this map had a better spatial resolution. The SUV were therefore less underestimated when using the CT map for attenuation correction than when using the Cs map. The changes in SUV observed in the lung spheres tended to be greater (between 15 and 120%) than those previously reported in another phantom study (between 3 and 15% [5]) and in a patient study (between 4.3 and 15.2% [6]). This might be because the difference in spatial resolution of the two attenuation maps we considered (around 1 mm for the CT map and more than 16 mm for the Cs map) was actually greater than the difference in spatial resolution of the two attenuation maps considered in these other studies in which both maps were segmented.

The nature of the attenuation map did not affect SUV estimates in the mediastinal spheres. Indeed, these spheres had an attenuation coefficient similar to that of the surrounding water compartment so that the spheres and surrounding medium were not differentiated in the transmission scan, had it been performed with the Cs source or with the CT. As the Cs and the CT transmission maps were almost identical at the mediastinal sphere location, the resulting SUV estimates were also very close (differences less than 5%).

Partial volume effect correction. Without PVE correction, the SUV underestimation was between -19% (largest mediastinal sphere) and -91% (smallest lung sphere) (Tables II and III). PVE correction strongly increased the SUV, by a factor greater than 2 (percent changes > 100% in Tables IV and V) and up to 6 for spheres less than 2 cm in diameter. This is consistent with the fact that PVE strongly affect measurements in structures less than 2 to 3 times the spatial resolution in the reconstructed images, i.e. less than 18 to 27

mm in diameter in our case. The CC correction method was more efficient at recovering SUV than the RC method (Tables IV and V), resulting in biases between +11% for the largest mediastinal sphere and –57% for the smallest lung sphere. This is because the matrix coefficients involved in the CC method account for the interference between PVE, attenuation, and attenuation correction as they are calculated by reproducing closely the attenuation and attenuation correction processes. This is not the case for the RC method, which does not interfere at all with attenuation correction.

Even if the CC correction greatly reduced the bias in SUV estimates, SUV were estimated with a bias less than 15% only for spheres greater than 22 mm in diameter. For spheres < 22 mm in diameter, SUV remained largely biased even with the CC correction (Table VI). The high biases that still affect SUV estimates for the smallest spheres might well be due to errors in estimating the sphere volumes. Indeed, sphere volumes as estimated from the manual segmentation of the CT were systematically underestimated (Table VII). These errors in volume estimates from the CT are typical of segmentation errors reported in the literature [7,8]. The accuracy of the segmentation step, hence of sphere volume estimates, affect attenuation correction, PVE correction and SUV estimates. When the volume of the structure of interest is underestimated, PVE will be corrected too strongly, which should lead to SUV overestimation. On the other hand, for the spheres located in the lungs, an underestimation of the sphere volume yields an underestimation of attenuation, as part of the sphere tissue is seen as lung tissue. Insufficient attenuation correction at the level of the lung spheres yields SUV underestimation. Finally, the size of the VOI considered to calculate SUV_{avg} also affects the SUV estimates. If the VOI is too small, SUV tend to be less biased by PVE than if the VOI is too large. The precise way these SUV overestimation and underestimation combine to yield the final bias in SUV still need further investigation. Our results suggest that despite overcorrection of PVE resulting from the underestimation of the sphere volume, SUV remain underestimated overall. This still needs clarification.

TABLE IV. CHANGES IN SUV WITH RESPECT TO SUVREF, DEFINED BY 100(SUV–SUVREF)/SUVREF FOR THE LUNG SPHERES.

Changes with respect to the reference values (%)	*Spheres in the lungs (Ø in mm)*			
	10.5	16	22	33
With OFOVA	11	2	2	6
18 min emission scan	-4	4	0	4
CT attenuation correction	120	78	31	15
RC PVE correction	492	162	77	41
CC PVE correction	531	233	97	47
SUVmax	71	60	65	74

TABLE V. CHANGES IN SUV WITH RESPECT TO SUVREF, DEFINED BY 100(SUV–SUVREF)/SUVREF FOR THE MEDIASTINAL SPHERES.

Changes with respect to the reference values (%)	*Spheres in the mediastinum (Ø in mm)*			
	10.5	16	22	33
With OFOVA	-8	8	1	6
18 min emission scan	5	-2	-2	1
CT attenuation correction	3	3	-3	4
RC PVE correction	197	107	61	34
CC PVE correction	222	109	55	36
SUVmax	26	51	45	59

SUV index. SUVmax were systematically higher and less biased than SUVavg. This is because SUVmax is less affected by PVE. Although this result could suggest that SUVmax should be preferred to SUVavg for accurate SUV estimates, using SUVmax yields biases that strongly depend on the noise level in the images. For instance, SUVmax was 9±6% higher on average on the 18 min scan compared to the 6 min scan, because of the difference in noise in these two scans, while SUVavg differed only by 1±3% on average between the 18 min scan and the 6 min scan. In [3], a large variability of SUVmax was also demonstrated when comparing 15 min and 1 min transmission scan durations.

TABLE VI. RESIDUAL BIASES ON SUV ESTIMATED ON DATA PROCESSED WITH AN OPTIMAL PROTOCOLE INCLUDING CT-BASED ATTENUATION AND PVE CORRECTION.

SUV underestimations in %	*Ø in mm*			
	10.5	16	22	33
Lung spheres	45	13	11	5
Mediastinal spheres	33	22	5	-11

TABLE VII. PERCENT ERRORS IN VOLUME ESTIMATES DEFINED BY 100(ESTIMATED VOLUME–TRUE VOLUME)/TRUE VOLUME) FOR THE LUNG AND MEDIASTINAL SPHERES.

Error in %	*Ø in mm*			
	10.5	16	22	33
Lung spheres	-57	-44	-20	-11
Mediastinal spheres	-57	-37	-18	-12

Overall, differences in SUV estimates greater than 100% (i.e., by a factor greater than 2) were observed only because of differences in the acquisition and processing protocols. This suggests that comparing SUV between studies makes sense only if the acquisition and processing protocols are identical, and that meta-analysis of SUV reported in different papers is almost impossible given the variability of the biases affecting SUV as currently assessed in different PET centers. This is all the more true given that many other factors that might also affect SUV have not been considered in our study, such as blood glucose level, time between injection and imaging session, or reconstruction algorithm.

IV. CONCLUSION

Biases in SUV estimates only slightly depend on the emission scan duration and on the presence of out-of-the-field-of-view activity, but strongly depend on the μ map used for attenuation correction, on whether PVE is corrected for and on whether SUV is calculated using the average count value within the tumor VOI or using the maximum value at the tumor location. Differences in SUV estimates greater than 100% (i.e., by a factor greater than 2) can be caused only by differences in the way data are acquired and processed, which suggest that comparison of SUV between PET centers using different scanning and processing protocols is almost impossible. Even when using CT-based attenuation correction and PVE correction, SUV estimates remain largely biased in lesions less than 2 cm in diameter, due to the difficulty in precisely assessing tumor volume from the CT.

V. REFERENCES

[1] Soret M, Riddell 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.

[2] Benard F, Smith RJ, Hustinx R, Karp JS, Alavi A. Clinical evaluation of processing techniques for attenuation correction with 137Cesium in whole-body PET images. *J Nucl Med 1999*; 40:1257-1263.

[3] Soret M, Koulibaly PM, Darcourt J, Hapdey S, Buvat I. Quantitative accuracy of dopaminergic neurotransmission imaging using 123I SPECT. *J Nucl Med* 2003; 44:in press.

[4] Visvikis D, Cheze-Le-Rest C, Costa DC, Bomanji J, Gacinovic S, Ell PJ. Influence of OSEM and segmented attenuation correction in the calculation of standardised uptake values for [^{18}F]FDG PET. *Eur J Nucl Med* 2001; 28: 1326-1335.

[5] Visvikis D, Costa DC, Croasdale I, Lonn AHR, Bomanji J, Gacinovic S, Ell PJ. CT-based attenuation correction in the calculation of semi-quantitative indices of [^{18}F]FDG PET uptake in PET. *Eur J Nucl Med* 2003; 30: 344-353.

[6] Nakamoto Y, Osman M, Cohade C, Marshall LT, Links JM, Kohlmyer S, Wahl R. PET/CT: comparison of quantitative tracer uptake between germanium and CT transmission attenuation-corrected images. *J Nucl Med* 2002; 43:1137-1143.

[7] Sohaib SA, Turner B, Hanson JA, Farquarson M, Olivier RTD, Reznek RH. CT assessment of tumour response to treatment: comparison of linear, cross-sectional and volumetric measures of tumour size. *Br J Radiol* 2000;73:1178-1184.

[8] Schiepers C, Brown M, Rogers S, McNitt-Gray M, Phelps ME, Dahlbom M. Lesion size and volume in combining PET and CT: phantom experiments. *J Nucl Med* 2002;43:14P.