J of Nuclear Medicine Technology, first published online March 15, 2021 as doi:10.2967/jnmt.120.259168

Accuracy Assessment of Standardised Uptake Value Measurements in SPECT/CT: A Phantom Study

Fatin Halim[1,2], Hizwan Yahya[1,2], Khairul Nizam Jaafar[3], and Syahir Mansor[1,3*]

[1] Oncological and Radiological Science Cluster, Advanced Medical and Dental Institute, SAINS@BERTAM, Universiti Sains Malaysia, Penang, Malaysia
[2] Department of Nuclear Medicine, Penang Hospital, Penang, Malaysia
[3] Nuclear Medicine Unit, Advanced Medical and Dental Institute, Universiti Sains Malaysia, Penang, Malaysia

*Corresponding author:
Syahir Mansor, Oncological and Radiological Science Cluster, Advanced Medical and Dental Institute, SAINS@BERTAM, Universiti Sains Malaysia, 13200 Kepala Batas, Penang, Malaysia
E-mail: syahir.mansor@usm.my
Tel: +604-5622538
ORCID ID: https://orcid.org/0000-0001-9315-652X

First author:

Fatin Halim, Oncological and Radiological Science Cluster, Advanced Medical and Dental Institute, SAINS@BERTAM, Universiti Sains Malaysia, 13200 Kepala Batas, Penang, Malaysia
E-mail: fatinnad@student.usm.my
Tel: +60172030231

Conflict of Interest: The authors declare no potential conflicts of interest.

Abstract

Introduction: The advanced development of iterative image reconstruction enables the absolute quantification of single-photon emission computerised tomography/computed tomography (SPECT/CT) studies by incorporating the compensation for a collimator–detector response, attenuation, and scatter into the reconstruction process. This study aimed to assess the quantitative accuracy of SPECT/CT based on the different levels of ^{99m}Tc activity (low/high) on the standardised uptake value (SUV) measurement for different SUV metrics (SUV_{mean}, SUV_{max}, $SUV_{0.6\ max}$, and $SUV_{0.75\ max}$). **Methods**: A Jaszczak phantom equipped with six fillable spheres was filled with low- and high activity phantom setup at 1:4 and 1:10 (background:sphere) ratios. The low- and high activity concentration phantom setups consisted of ≈10 and ≈60 kBq/ml background, respectively, at both ratios. The fixed-size volume-of-interest (VOI) based on the diameter of each sphere was drawn on SPECT using various metrics for SUV quantification purposes. **Results**: The convergence of activity concentration was dependent on the iteration number and application of post-filtering. For the background-to-sphere ratio of 1:10 with low activity concentration, the SUV_{mean} metric showed an underestimation of about 38% from the actual SUV, and SUV_{max} exhibited an overestimation of about 24% for the largest sphere diameter. Meanwhile, the bias reductions of as much as −6% and 7% for $SUV_{0.6\ max}$ and $SUV_{0.75\ max}$ were observed, respectively. SUV_{max} indicated the most accurate reading compared with the others, although points that exceeded the actual value were detected. At 1:4 and 1:10 background, the low activity concentration attained a value close to the actual ratio. Using two iterations, 10 subsets (2i10s) without post-filtering had the best parameter in terms of the most accurate values for reconstruction and provided the best image overall.

Conclusion: SUV_{max} is the best metric in high- or low-contrast ratio phantom with adequate iteration number of at least 2i10s without any post-filtering.

Keywords: Quantitative, SPECT/CT, SUV, ratio, phantom

INTRODUCTION

The development of multimodality gamma camera (single-photon emission computed tomography with computed tomography, SPECT/CT) instrumentation, image reconstruction algorithms, and advanced compensation methods to correct photon attenuation, scattering, and resolution have rendered the quantitative SPECT a feasible method that is comparable to the well-established quantitative positron emission tomography (PET) (*1*).

Filtered back projection (FBP) and ordered-subset expectation maximisation (OSEM) algorithms are the two most commonly used algorithms in SPECT reconstruction (*2*). Although the FBP algorithm is simple and fast (*3*), it unnecessarily amplifies the high-frequency noise which in turn will affect the quality of the final reconstructed SPECT image (*4*). Another limitation of FBP is that attenuation cannot be readily integrated and compensated with FBP. Before or after reconstruction, the data should be corrected to compensate for attenuation in the FBP reconstruction, presenting a challenge for scholars (*5*).

Three-dimensional (3D) OSEM is a quantitative image reconstruction algorithm used in the state-of-the-art SPECT/CT system (*6*). OSEM separates the measured datasets into various subsets and uses a single subset for every iteration, which will accelerate the algorithm by a factor equal to the number of subsets (*7*).

The quantitative accuracy of reconstructed SPECT images deteriorates due to several physical factors, namely, photon attenuation, compton scattering, and spatially varying collimator response (*2*). The inclusion of collimator response, i.e. resolution

recovery in the GE system will increase the accuracy of the final reconstructed images (*8,9*). Attenuation causes inconsistent projection information that may increase or decrease counts in the image, particularly in the area close to the detector plane (*10*). Meanwhile, the presence of scattered photons will result in less contrast and loss of quantification accuracy in reconstructed images.

Different activity concentrations affect the quantification of SPECT/CT. Based on the study by Francis *et al*. (*11*), standardised uptake value (SUV) increases with the increased activity concentration for the same sizes of spheres. Their study proved that radionuclide uptake values correspond to activity concentration in organs or tissues.

The use of SUV in the quantification of SPECT/CT is gaining interest as a quantitative capability of SPECT/CT. SUV is defined as the concentration of radioactivity in the tissue normalised to the injected dose and body weight. SUV_{mean} is defined as the average SUV in the VOI. The SUV_{max} is defined as the maximal SUV in the VOI. In this study, we included $SUV_{0.6\ max}$ and $SUV_{0.75\ max}$, which are defined as the average values that include pixels greater than 60% and 75% of the maximal SUV in the VOI, respectively. Finally, we compared the four SUV metrics in terms of the most accurate reading.

Factors that potentially affect SUV measurements include spatial resolution and reconstruction parameters. For small objects, image resolution has a partial volume effect on the measured SUV (*12*). Usually, SPECT collimators are equipped with a maximum permissible resolution to partially offset the limited detection performance (*13*). Any changes in the reconstruction parameters, such as matrix size, filtering, field of view size, and iteration number, will have a significant effect on SUV calculation in clinical cases.

In this study, we aimed to assess the quantitative accuracy when different levels of activity concentration (low/high) with different reconstruction parameters were used for various SUV metrics (SUV_{mean}, SUV_{max}, $SUV_{0.6\ max,}$ and $SUV_{0.75\ max}$) during SUV measurement.

MATERIALS AND METHODS

Phantom studies

Quantitative $^{99\ m}$Tc-SPECT/CT acquisitions of a Jaszczak phantom containing six spheres of various diameters (9.9, 12.4, 15.6, 19.7, 24.8, and 31.2 mm) were performed on GE Discovery NM/CT 670 SPECT/CT device (GE Healthcare, Waukesha, USA) equipped with a low-energy high-resolution collimator. For the first experiment, the background compartment was filled with an activity concentration of ≈60 kBq/ml with a sphere-to-background ratios of 10:1 and 4:1. For the second experiment, the background compartment was filled with an activity concentration of ≈10 kBq/ml and sphere-to-background ratios of 10:1 and 4:1.

Data acquisition and reconstruction

SPECT acquisitions were acquired and reconstructed using the proprietary iterative conjugate gradient algorithm on a GE Xeleris workstation with 128 × 128 × 128 voxel grid. The acquisition voxel size was 4.42 × 4.42 × 4.42 mm^3. A CT scan with an energy of 120 kVp and a tube current of 205 mAs was used for attenuation correction. CT based

attenuation correction and dual-energy window scatter correction were systematically applied in SPECT reconstructions. Both experiments were acquired using 20 s per view for a total of 60 views per camera head with no zoom application (×1 multiplication), and images were reconstructed using 2, 10, and 20 iterations with 10 fixed subsets with and without the Gaussian filter. SPECT/CT data were reconstructed in an isotropic voxel size of 128 × 128 × 128 with a dimension of 4.42 × 4.42 mm^2 and a slice thickness of 4.42 mm. All images were reconstructed using the OSEM algorithm with attenuation correction, scatter correction, and resolution recovery. We utilised the Gaussian post-filter using a 4 mm full width at half maximum.

Image analysis

The reconstructed data were processed using A Medical Image Data Examiner (AMIDE, version 1.0.4) freeware tool (*14*). The VOI was drawn for six spheres based on the CT images and then fused on the SPECT images. This tool was used to obtain the total number of counts in the VOI on the SPECT images. The SUV was calculated from the activity concentration (kBq/ml) divided by the total administered activity (kBq) within the phantom background chamber and normalised to the weight (g) of the solution in the phantom. SUV_{mean}, SUV_{max}, $SUV_{0.6\ max}$, and $SUV_{0.75\ max}$ metrics were used for each sphere in different phantom contrast ratios and activity concentrations, similar to the study conducted by Lee *et al*. (*15*). The $SUV_{0.6\ max}$ and $SUV_{0.75\ max}$ were calculated based on the assumption of the average values of all regional voxels with values being between 60% and 75% to the maximal voxel value, respectively. The statistical variance or noise was

Accuracy Assessment of SUV

determined using the coefficient of variation (COV). For each measurement, the COV for each sphere was calculated by dividing the standard deviation by the mean of the reading for each sphere.

Statistical analysis

Statistical analysis was performed using SPSS (version 24, IBM Corp., Armonk, NY, USA) software. An independent sample T-test was used to verify whether significant differences existed between i) high and low activity concentrations and ii) 1:4 and 1:10 ratios. Meanwhile, a sample T-test was used to compare the SUVs with the reference values for different factors, such as the different activity concentrations, ratios, SUV metrics, and reconstruction. In this study, multiple metric SUVs were tested to find the value nearest to the reference value, and the results were proven through one-way analysis of variance (ANOVA).

RESULTS

From Figure 1, high iteration numbers increased the contrast along with the noise of an image. The use of post-filter suppressed the noise, but it caused an over smoothing effect (low-contrast) on the image (rows b and d). By using a filter, the contrast between the sphere and background was reduced, resulting in the less qualitative enhancement of the images. However, the smooth effect caused by the filter introduced additional blurring to the image and hence eliminated details of the structure within the image**.**

Accuracy Assessment of SUV

Descriptive analysis

All the SUV metrics (SUV_{mean}, SUV_{max}, $SUV_{0.6\ max}$, and $SUV_{0.75\ max}$) for different activity concentrations and ratios were plotted on the graph against the function of sphere size and background. Based on the plotted graph, the SUV at low activity concentration indicated an overestimated value compared with the actual value. Meanwhile, the high activity concentration underestimated the SUV for the sphere-to-background ratios of 1:4 and 1:10 (Figure 2, Supplementary Figures 1, 2, and 3). Meanwhile, overestimated values were observed at low activity concentration, and a sphere-to-background ratio of 1:4 showed a higher accuracy in low activity concentration (Supplementary Figure 1) than in high activity concentration (Figure 2).

The various diameters of spheres inside the Jaszczak phantom (9.9, 12.4, 15.6, 19.7, 24.8, and 31.2 mm) were analysed across the activity concentrations and sphere-to-background ratios of 1:4 and 1:10. The results showed that at the sphere-to-background ratio of 1:4 with high activity concentration, the SUVs (mean, maximum, 0.6 maximum, and 0.75 maximum) increased with the increase in sphere diameter (Figure 2).

When we focused on the largest sphere, the SUV_{mean} reduced by about 38% of the actual SUV at a low activity concentration, whereas SUV_{max} revealed an increase of about 24%. Given that $SUV_{0.6\ max}$ and $SUV_{0.75\ max}$ showed no major variations, the reduction biases of about −6% and 7% were observed, respectively.

In this experimental work, the fixed subset of 10 was used for the reconstruction parameter, whereas the iteration increase from 2 to 10 and 20 indicated an overestimation of the SUV when the filter was not utilised. In the comparison of the difference between

using and not using filters, we discovered that the SUVs were underestimated when using filters regardless of the differences in i) iteration numbers, ii) SUV threshold, iii) and sphere diameter and iv) different activity concentrations and v) sphere-to-background ratios. Such result was due to the filter function of simultaneously removing noise while reducing the counts in the final reconstructed images.

COV was used to demonstrate the noise level; it can be achieved by dividing the standard deviation with the average activity concentration times by 100%. Figure 3 shows the noise for various reconstructions for the three largest spheres in the phantom. The higher the iteration number, the higher the COV (Figure 3 and Supplementary Figure 4). However, when using filters, the COV values were the same for all the three iterations used.

Statistical analysis

All the data obtained were analysed using the SPSS software. The *p*-values for different SUV metrics at the sphere-to-background ratios of 1:4 and 1:10 with different activity concentrations significantly showed a difference. Theoretically, the true mean value at the 1:4 sphere-to-background ratio is 4 for all the spheres inside the phantom, whereas the true mean value at the 1:10 ratio is 10. The SUV_{max} at low activity concentration was the nearest with the theoretical value for both ratios.

One-way ANOVA was conducted to determine whether the SUVs differed for groups with various SUV metrics. The SUV was classified into four groups: SUV_{mean}, SUV_{max}, $SUV_{0.6\ max}$, and $SUV_{0.75\ max}$. The SUV showed statistically significant difference between different SUV metrics, where $p < 0.001$. Dunnett T3 post-hoc analysis revealed

Accuracy Assessment of SUV

that the SUV_{mean} was statistically significant ($p < 0.05$) when compared with the other three groups, but no other group differences were statistically significant between SUV_{max}, $SUV_{0.6\ max}$, and $SUV_{0.75\ max}$.

The varying spherical sizes inside the phantom represent the lesion of the patients' body. Significant differences were observed for all the sphere sizes, with the true mean values of $\mu = 4$ and $\mu = 10$, except for the 24.8 mm sphere in low activity concentration and at a sphere-to-background ratio of 1:4 ($p = 0.104$).

The effects of different reconstruction iterations on the true mean value of the phantom showed a significant difference with the true mean ($\mu = 4$ and $\mu = 10$) under different reconstruction methods, except for 2 iterations with 10 subsets (2i10s) without post-filtering in low activity concentration for the sphere-to-background ratios of 1:4 ($p = 0.595$) and 1:10 ($p = 0.268$).

DISCUSSION

Two different activity concentrations of backgrounds (60 and 10 kBq/ml) were used in this study. These backgrounds were intended to compare the two concentrations to obtain an accurate reading relative to the theoretical value. Increasing the activity concentration retained the SUV based on the ratios. However, the results showed that the SUV was the closest to the actual value in low activity concentration. When we observed the phantom images at different concentrations and sphere-to-background ratios, we determined that the spill-in effect appeared at a low activity concentration at 1:4 sphere-to-background ratio for the smallest sphere diameter (9.9 mm). This effect is known as the

partial volume effect, which is caused by the limited partial resolution of the imaging system (*16,17*).

In the comparison of the four graphs based on the SUV in Figure 2 and Supplementary Figures 1, 2, and 3, the SUV_{mean} showed an underestimation, whereas the SUV_{max} exhibited an overestimation from the real SUV. Given that $SUV_{0.6\ max}$ and $SUV_{0.75\ max}$ showed no considerale difference, the $SUV_{0.75\ max}$ was closer to the theoretical value.

Supplementary Figure 1 indicates that a low activity concentration with a low sphere-to-background ratio may be disastrous if the SUV_{max} metric is used. Given the high background value, proper background subtraction is an important measure in quantification. The number of iterations is the primary variable affecting image quality. Although a high number of iterations will generally result in a high spatial resolution, the noise level will also increase. The high the number of iterations will also hasten image convergence.

The algorithm cannot fully converge if the iterations number of is inadequate, which will eventually result in a blurry image with an inadequate contrast. Meanwhile, if the number of iterations is extremely large, then the reconstructed image shows over-sharpening with an elevated level of noise (Figure 1). The selection of reconstruction parameters, such as iteration number, filtering, attenuation correction, scattered correction and resolution recovery are important to produce good-quality images with the minimum level of noise.

In this study, an image was reconstructed using 10 subsets with 2, 10, and 20 iterations. The 2i10s without post-filtering in low activity concentration for both spheres-

to-background ratios showed the most suitable parameter based on the SUV obtained from this study. The result indicates that the obtained value approached the true theoretical value.

The post-filter reduced the noise but also produced a smoothing effect on the final reconstructed images. Filters can have a major effect on the quality of clinical images due to their degree of smoothing. Proper filter selection and appropriate smoothing allow doctors to interpret the results and perform diagnosis accurately (*18*). Gaussian smoothing demonstrates a low image resolution and thus reduces the accuracy of the SUV.

In this experimental study, when we applied the 4 mm Gaussian filter, the accuracy of SUV dramatically reduced, and not a single value under various ratios and spheres had reached the true theoretical value. Thus, the selection of filters can affect the quantitative value of SPECT/CT images.

Clinical bone SPECT/CT studies usually have high sensitivity but low specificity. This imaging procedure typically reconstructs images using three or five iterations and eight to ten subsets with post-smoothing using a Gaussian or Butterworth filter (*19*). For quantitative analysis, the value for filter application must be appropriately selected to compensate between image quality and noise. Given that the SUV_{max} provides the most accurate value in various ratios and low or high activity concentration, it should be considered for use in clinical settings.

In nuclear medicine imaging, quantification offers a great advantage. Although SUV may have initially been a framework for PET imaging, it is now found equally ideal for SPECT imaging. The use of SUV in SPECT imaging offers a wide range of radiopharmaceuticals and applications. In this study, we compared different SUV metrics by creating various factors to determine the accuracy of SUV readings. SUV_{max}

demonstrated the reading that correlated most accurately with the clinical setting commonly used for SUV reporting.

CONCLUSION

The concentration of activity ratio (high or low activity concentration) plays a role in the determination of accurate SUV. Based on the analysis results, the low activity concentration under both ratios provided the more accurate value compared with the high activity concentration. The SUV_{max} was the closest to the actual theoretical values. From the aspect of reconstruction, the use of 2i10s without post-filtering is the optimal protocol for accurate quantification and optimal overall image quality with a compromising noise level.

Acknowledgement: This study is supported by Universiti Sains Malaysia (USM), Short Term Grant 304/CIPPT/6315160 and Fundamental Research Grant Scheme 203/CIPPT/6711730

REFERENCES

1. Bailey DL, Willowson KP. An evidence-based review of quantitative SPECT imaging and potential clinical applications. *J Nucl Med.* 2013;54:83–89.

2. Alzimami KS, Sassi SA, Spyrou NM. A Comparison Between 3D OSEM and FBP Image Reconstruction Algorithms in SPECT. In: Ao S-I, Gelman L, eds. *Advances in Electrical Engineering and Computational Science*. Dordrecht: Springer Netherlands; 2009:195–206.

3. Zeng GL. A filtered backprojection algorithm with characteristics of the iterative landweber algorithm. *Med Phys.* 2012;39:603–607.

4. Ramírez J, Górriz JM, Gómez-Río M, et al. Effective Emission Tomography Image Reconstruction Algorithms for SPECT Data. Paper presented at: Computational Science–ICCS 2008; 2008//, 2008; Berlin, Heidelberg.

5. Meysam T, Marian N. Quantitative evaluation of the effect of attenuation correction in SPECT images with CT-derived attenuation. Paper presented at: Proc.SPIE, 2019.

6. Katua AM, Ankrah AO, Vorster M, van Gelder A, Sathekge MM. Optimization of Ordered Subset Expectation Maximization Reconstruction for Reducing Urinary Bladder Artifacts in Single-photon Emission Computed Tomography Imaging. *World J Nucl Med.* 2011;10:3–8.

7. Hudson HM, Larkin RS. Accelerated image reconstruction using ordered subsets of projection data. *IEEE Trans Med Imaging.* 1994;13:601–609.

8. Ismail FS, Mansor S. Impact of Resolution Recovery in Quantitative (99 m)Tc SPECT/CT Cardiac Phantom Studies. *J Med Imaging Radiat Sci.* 2019;50:449–453.

9. Knoll P, Kotalova D, Köchle G, et al. Comparison of advanced iterative reconstruction methods for SPECT/CT. *Zeitschrift für Medizinische Physik.* 2012;22:58–69.

10. Frey EC, Humm JL, Ljungberg M. Accuracy and precision of radioactivity quantification in nuclear medicine images. *Semin Nucl Med.* 2012;42:208–218.

11. Francis H, Amuasi JH, Kwame KA, Vangu MD. Quantitative Assessment of Radionuclide Uptake and Positron Emission Tomography-computed Tomography Image Contrast. *World J Nucl Med.* 2016;15:167–172.

12. Adams MC, Turkington TG, Wilson JM, Wong TZ. A systematic review of the factors affecting accuracy of SUV measurements. *AJR Am J Roentgenol.* 2010;195:310–320.

13. Ritt P, Vija H, Hornegger J, Kuwert T. Absolute quantification in SPECT. *Eur J Nucl Med Mol Imaging.* 2011;38 Suppl 1:S69–77.

14. Loening AM, Gambhir SS. AMIDE: a free software tool for multimodality medical image analysis. *Mol Imaging.* 2003;2:131–137.

15. Lee JR, Madsen MT, Bushnel D, Menda Y. A threshold method to improve standardized uptake value reproducibility. *Nucl Med Commun.* 2000;21:685–690.

16. Soret M, Bacharach SL, Buvat I. Partial-volume effect in PET tumor imaging. *J Nucl Med.* 2007;48:932–945.

17. Du Y, Madar I, Stumpf MJ, Rong X, Fung GS, Frey EC. Compensation for spill-in and spill-out partial volume effects in cardiac PET imaging. *J Nucl Cardiol.* 2013;20:84–98.

18. Lyra M, Ploussi A. Filtering in SPECT Image Reconstruction. *International Journal of Biomedical Imaging.* 2011;2011:693795.

19. Van den Wyngaert T, Strobel K, Kampen WU, et al. The EANM practice guidelines for bone scintigraphy. *Eur J Nucl Med Mol Imaging.* 2016;43:1723–1738.

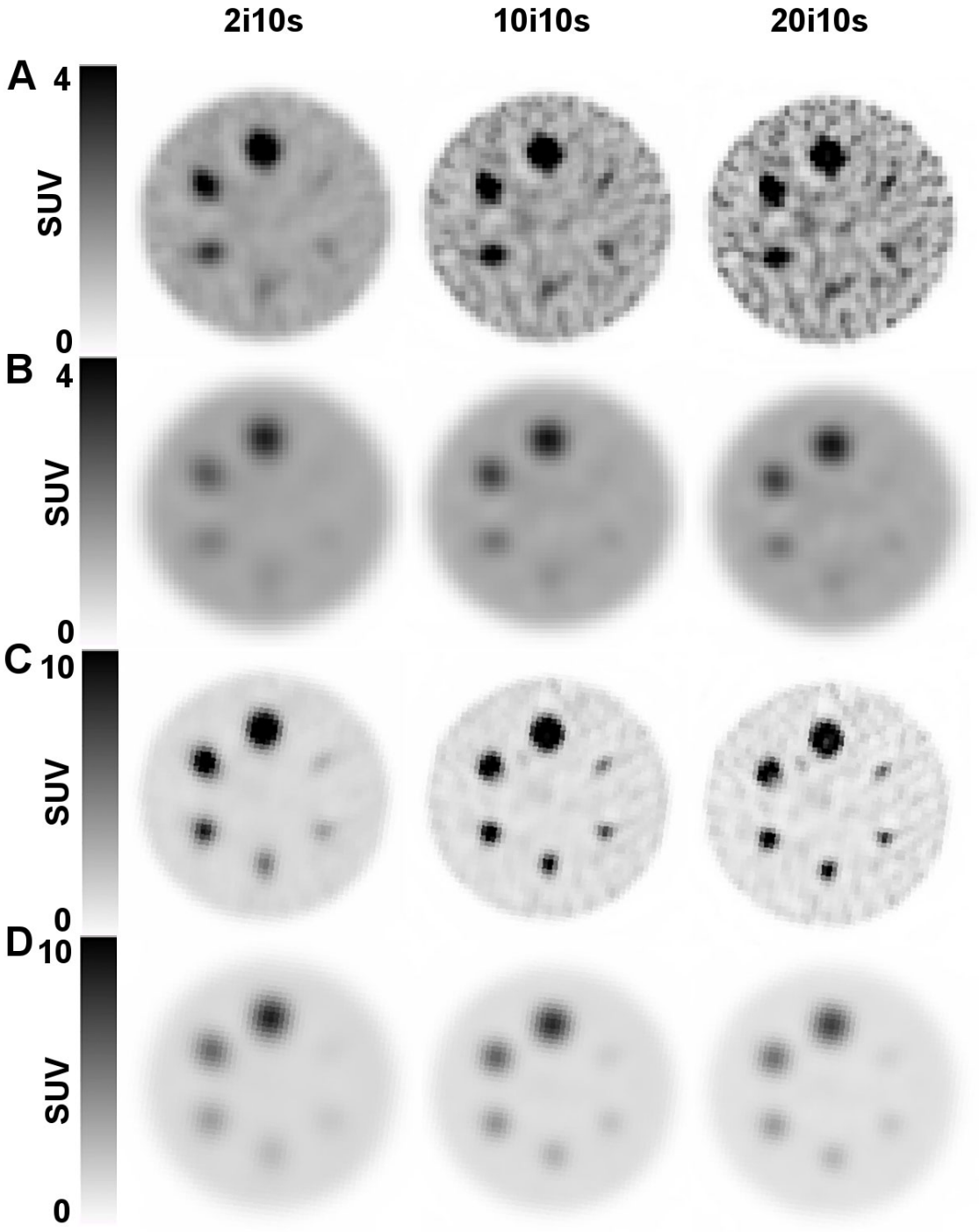


Figure 1. SUV images from the phantom study without (rows a and c) and with (rows b and d) 4 mm Gaussian post-filter. Reconstructed images achieved with 3D-OSEM reconstruction algorithm with 2, 10, and 20 iterations with 10 subsets. The image was obtained with the background-to-sphere ratios of 1:4 and 1:10 with a high activity concentration.

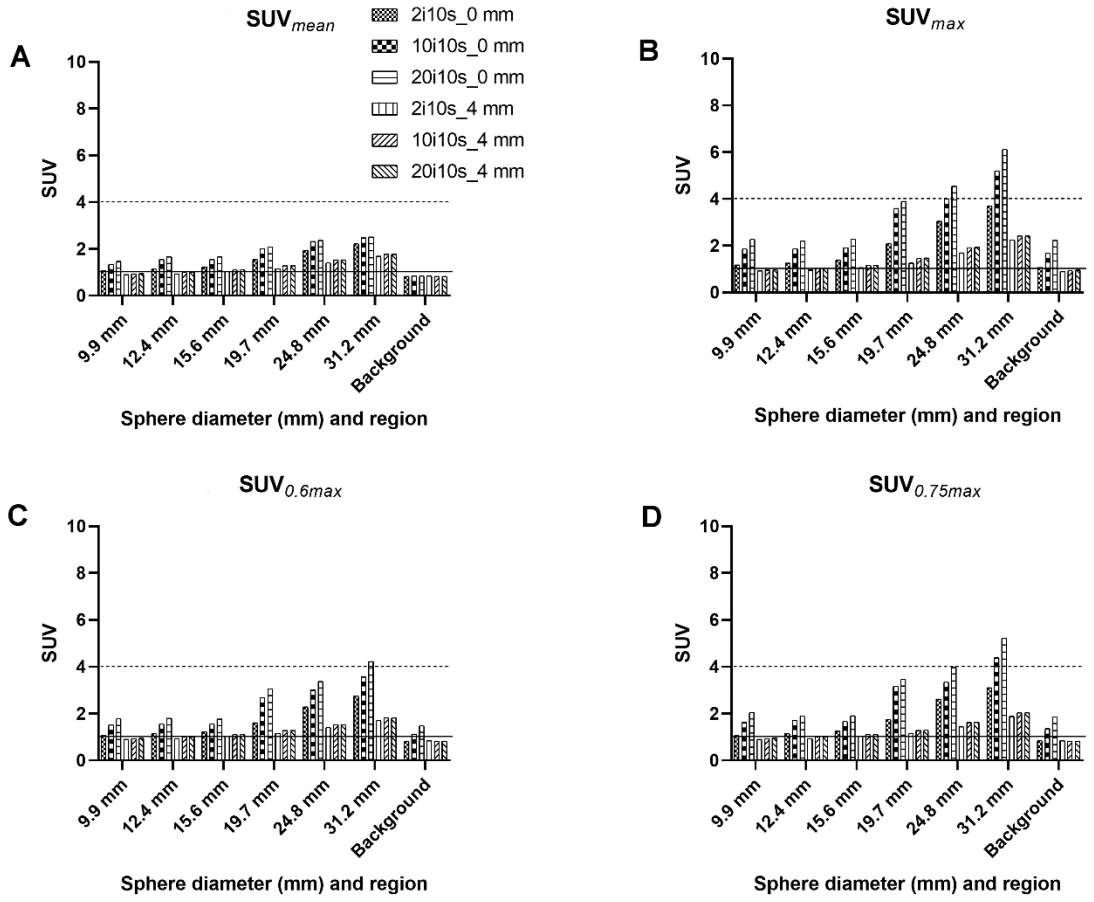


Figure 2. Sphere-to-background ratio of 1:4 with high activity concentration. The dotted lines show the true SUV for the respective spheres and background.

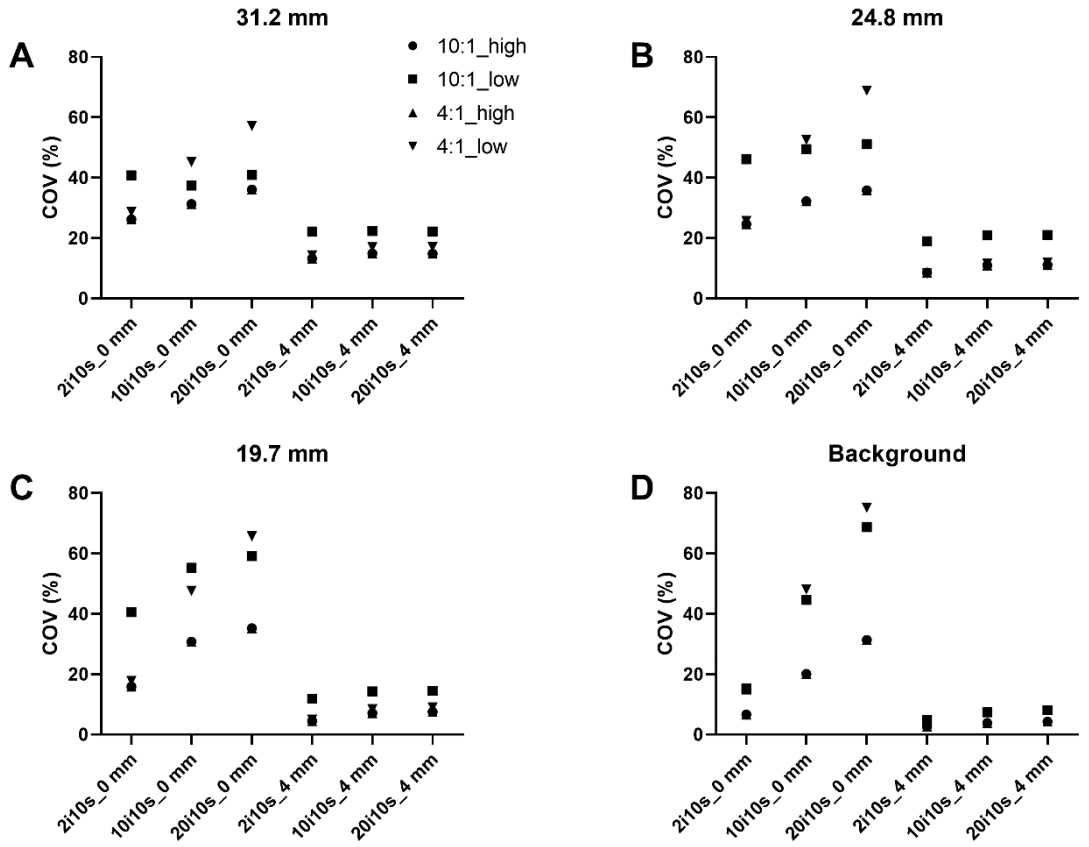
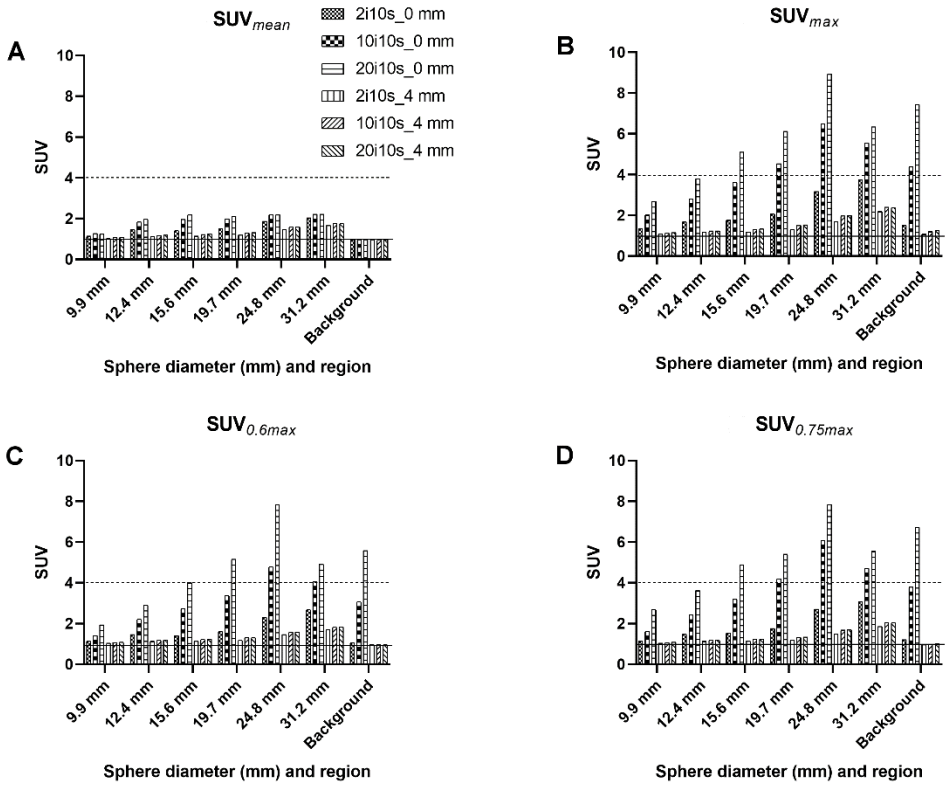
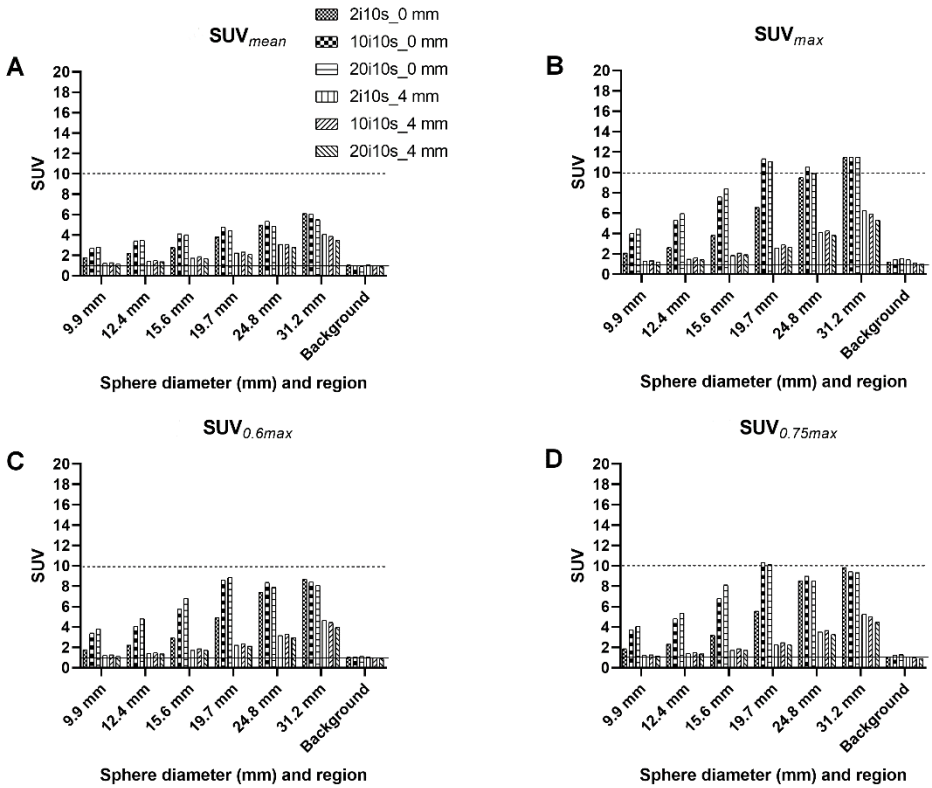


Figure 3. Noise (COV %) for various reconstructions for the three largest spheres and background.

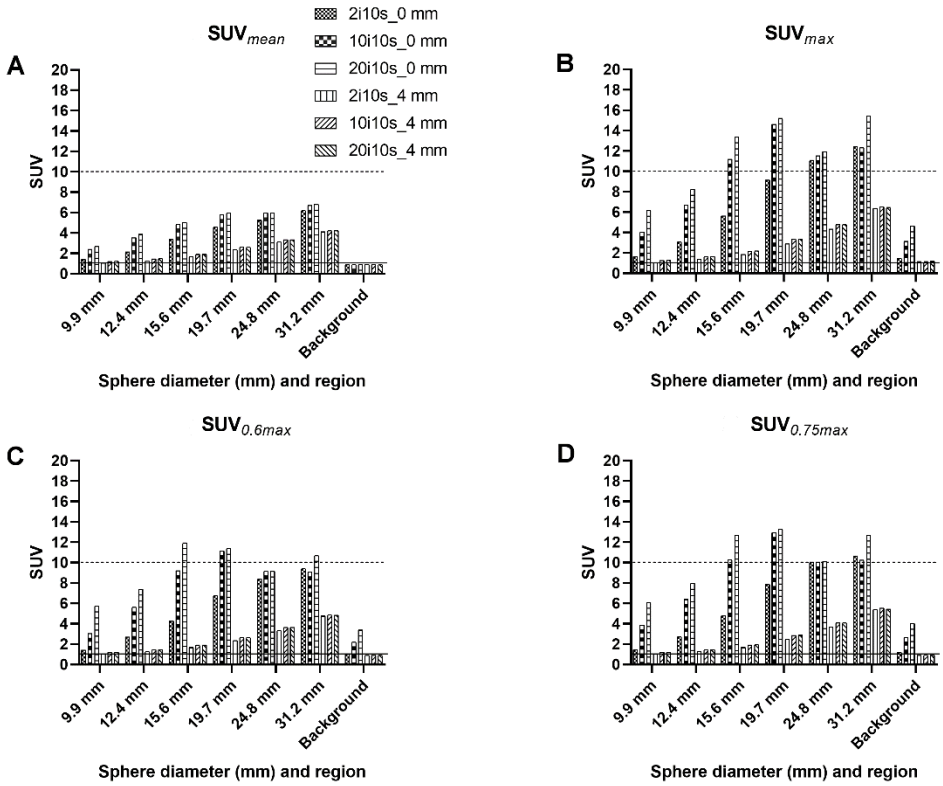
Supplemental Figure 1



Supplemental Figure 2



Supplemental Figure 3



Supplemental Figure 4