BioMedical Engineering OnLine

RESEARCH

Open Access

Calculation of recovery coefficients for partial volume effect correction in PET/CT imaging using a customized anthropomorphic body phantom

Gunes Yavuz[1*], Bilal Kovan[2], Turkay Toklu[3], Tevfik F. Çermik[4] and Cengizhan Öztürk[1,5]

*Correspondence:
gunesyavuz75@gmail.com

[1] Institute of Biomedical Engineering, Boğaziçi University, Kandilli Campus, Feza Gürsey Bld., 34685 Çengelköy-Istanbul, Turkey
[2] Istanbul Faculty of Medicine, Department of Nuclear Medicine, Istanbul University, Turgut Özal Millet Cad., No:118U, 34104 Fatih-Istanbul, Türkiye
[3] Department of Nuclear Medicine, Yeditepe University, Kosuyolu Hospital, Kosuyolu Cad., No: 168, 34718 Kadıköy-Istanbul, Türkiye
[4] Department of Nuclear Medicine, Acıbadem Maslak Hospital, Büyükdere Cad. No: 40 Maslak, 34398 Sarıyer-Istanbul, Türkiye
[5] Center for Targeted Therapy Technologies (CT3), Boğaziçi University, Kandilli Campus, 34685 Çengelköy-İstanbul, Türkiye

Abstract

Positron Emission Tomography/Computed Tomography (PET/CT) combines metabolic and anatomical information improving the precision and accuracy of oncological diagnostics. The standardized uptake value (SUV) measures tumor metabolism, yet its accuracy is influenced by the partial volume effect (PVE), impacting small lesion detection. This study aims to refine PVE corrections for small lesions using an in-house customized, special anthropomorphic phantom. Scans of this phantom which contained spheres of different sizes were performed across four hospitals at different PET/CT systems from various manufacturers (Siemens and Philips analog PET/CT systems, GE analog and digital PET/CT systems). The phantom contained six custom-designed cylinders with embedded spheres simulating sub-centimeter (0.3, 0.5, 0.9) and centimeter (1.3, 1.9, 2.8) lesions. Scans were performed separately for each sphere in the thorax, abdomen, and pelvis regions at all sites. Recovery Coefficients (RCs) were calculated to correct SUV values, demonstrating that RCs vary by sphere size and anatomical region but not change significantly among scanners. RCs are approaching unity for larger spheres, ensuring accurate SUV measurements. However, small spheres (< 0.5 cm) exhibited significant measurement challenges due to PVE. The anthropomorphic phantom proved effective in obtaining realistic SUV-corrected values, offering a promising tool for enhancing the accuracy and standardization of PET imaging in oncology. This study underscores the necessity for advanced imaging technologies and standardized RC application in clinical settings to improve the quantification of PET imaging, particularly in small lesion detection.

Keywords: Partial volume effect, PVE, Partial volume effect correction, Anthropomorphic phantom, Recovery coefficients, PET/CT, Medical imaging

Introduction

PET/CT imaging integrates metabolic and anatomical information into a single scan, leveraging the strengths of both modalities. PET/CT studies, using the radiopharmaceutical 18F-FluoroDeoxyGlucose (18F-FDG) are primarily used in oncology, facilitating

tumor detection, staging, and the monitoring of treatment responses [1].

Evaluating PET/CT clinical images involves both qualitative visual assessment and quantitative analysis. Visual assessment aids in identifying anatomical and metabolic changes in structures, ranging from tissue necrosis to tumors. Quantitative analysis, on the other hand, evaluates tumor metabolism by measuring the SUV which aids in not only distinguishing between benign and malignant tumors, but also evaluation of therapy response and forecasting tumor aggressivity. The SUV is a semiquantitative metric of radiopharmaceutical activity concentration in a region of interest (ROI) compared to the overall distribution in the body, normalized by the patient's body weight and the injected dose [1].

The quantification of PET/CT images using SUV is influenced by multiple technical and physiological factors. PVE causes measurement inaccuracies in the concentration of radiopharmaceuticals within regions smaller than approximately 2–3 times the scanner's spatial resolution (4–5 mm) [2–4].

PET-NEMA (National Electrical Manufacturers Association) EIC (Electrotechnical International Commision) phantoms are designed to evaluate the performance of PET imaging systems but have limitations when it comes to accounting for different tissue densities and performing precise sub-centimeter measurements. Furthermore, PET NEMA EIC body phantom performs PVE corrections with water.

The water is used to create a homogeneous medium that mimics the average density of soft tissues, allowing for standardization in the evaluation of PET image quality, contrast recovery, and other performance metrics. However, the human body is inhomogeneous with varying tissue densities [5–7].

The PVE in PET imaging is influenced by various factors, including spatial resolution, image sampling, tumor size, tumor shape, and the method used to measure radiopharmaceutical concentration within the tumor [8]. This phenomenon causes some electron–positron annihilations originating within the tumor to be detected as if they had originated in surrounding tissues, a process referred to as "spill-out." Consequently, larger tumors may appear to have lower uptake in images than they do. Conversely, some annihilations originating in surrounding tissues are detected as coming from the tumor, known as "spill-in". The spill-in effect causes signal blurring. These effects partially compensate for each other, impacting the accuracy of quantitative measurements in PET imaging [8, 9].

Scientific studies are being conducted to enhance the detectability and measurement accuracy of small lesions in PET imaging systems. Bettinardi et al. emphasize that blurring caused by spatial resolution reduces image contrast, thereby limiting the detection of small lesions [10].

Alavi et al. strongly emphasize the necessity of PVE correction to obtain reliable measurements of tracer activity accumulated in the body, especially in the case of repeated PET scans. PVE correction is crucial for assessing treatment response, particularly in situations where significant shrinkage and reduction in tumor size are expected [11].

Multiple PVE correction strategies have been developed to compensate for PVE in PET imaging, but there is no consensus on standardized application in oncological PET studies. These strategies generally focus on producing accurate SUV_{max} measurements for lesions and allow for determining tumor metabolic activity. However, SUV_{max}

readings are strongly affected by noise, and highly susceptible to statistical fluctuations, because they rely on a single voxel. This reliance leads to challenges, including reduced accuracy in quantification and potential diagnostic errors due to overestimation or underestimation of tracer uptake [12].

Among all PVE correction methods, more common ones are based on RC model which involves applying a correction factor to the SUV measured in a region, as proposed by Grings et al. [13].

Srinivas et al. proposed a more realistic model that considers hot spheres in a phantom with hot background, acknowledging that real tumors are always surrounded by tissues with some back ground activity [14]. Hoffman et al. suggested PET experimental measurements of RC can be carried out using 18F-FDG radioactive spheres (hot spheres) [15]. RC coefficients were obtained by calculating the ratio between PET measured and the actual (known) radioactivity concentration within the hot spheres [16–18]. This approach was designed to be applied to the PVE correction of PET oncological lesions in real patients, as radioactive spheres effectively simulate metabolically active oncological lesions [2]. This method is particularly advantageous because it requires SUV correction only within the tumor region. This ensures accurate quantification of metabolic activity while minimizing the need for broader, more complex corrections across non-target tissues. This targeted approach enhances the precision of PET imaging in evaluating tumor characteristics and treatment response [19].

The main objective of this study is to achieve a high accuracy rate in RC calculations applied to PVE corrections using our uniquely designed heterogeneous anthropomorphic human body phantom. These phantoms mimic human anatomical conditions to evaluate the performance of correction algorithms (RCs) and determine their effectiveness and limitations. The research aims to improve PVE correction strategies to enhance the accuracy and reliability of PET imaging, especially in small lesions.

Results

Sphere and background activities are measured in each site with six different sizes of spheres (0.3, 0.5, 0.9, 1.3, 1.9 and 2.8 cm) on 4:1 and 8:1 ratio and for each three anatomical regions (Thorax, Abdomen, Pelvis). Accordingly, RC values are calculated for each sphere. While doing all these calculations, according to F-18 FDG injected time, to eliminate half time effect, activity decay calculations taken into consideration to get more realistic results. The calculated RCs are shown in Tables 1, 2, 3 and 4 for each site separately.

The RCs measure how well the imaging system recovers the true activity concentration within a ROI, such as a lesion or a tumor. After doing all calculations, RC data were analyzed and fitted to the Asymptotic Regression Model (ARM) function with the following Eq. 1

$$Y = a - bc^{x}. \tag{1}$$

The software facilitated precise curve fitting and data analysis, enabling a thorough examination of the RC characteristics. The utilization of the ARM provided a robust framework for understanding the asymptotic behavior of the data. This approach

Table 1 For acquisitions at $n=4{:}1$ and $n=8{:}1$, RC values for each anatomical region and each sphere for Site-1

Sphere (cm)	4:1 Thorax	4:1 Abdomen	4:1 Pelvis	8:1 Thorax	8:1 Abdomen	8:1 Pelvis
0.30	0.122	0.133	0.154	0.075	0.048	0.086
0.50	0.372	0.313	0.297	0.097	0.089	0.075
0.90	0.425	0.478	0.473	0.360	0.347	0.306
1.30	0.616	0.500	0.555	0.379	0.361	0.392
1.90	0.629	0.671	0.592	0.556	0.491	0.451
2.80	0.677	0.714	0.799	0.590	0.744	0.624

Table 2 For acquisitions at $n=4{:}1$ and $n=8{:}1$, RC values for each anatomical region and each sphere for Site-2

Sphere (cm)	4:1 Thorax	4:1 Abdomen	4:1 Pelvis	8:1 Thorax	8:1 Abdomen	8:1 Pelvis
0.30	0.110	0.133	0.115	0.115	0.093	0.152
0.50	0.251	0.218	0.262	0.243	0.217	0.228
0.90	0.430	0.404	0.421	0.295	0.305	0.291
1.30	0.549	0.549	0.527	0.581	0.606	0.564
1.90	0.648	0.679	0.635	0.717	0.645	0.615
2.80	0.767	0.847	0.819	0.878	0.815	0.871

Table 3 For acquisitions at $n=4{:}1$ and $n=8{:}1$, RC values for each anatomical region and each sphere for Site-3

Sphere (cm)	4:1 Thorax	4:1 Abdomen	4:1 Pelvis	8:1 Thorax	8:1 Abdomen	8:1 Pelvis
0.30	0.085	0.116	0.116	0.118	0.107	0.100
0.50	0.135	0.198	0.116	0.212	0.239	0.187
0.90	0.307	0.336	0.290	0.406	0.349	0.397
1.30	0.533	0.530	0.637	0.470	0.507	0.514
1.90	0.745	0.805	0.704	0.494	0.552	0.534
2.80	0.862	0.812	0.818	0.821	0.750	0.746

Table 4 For acquisitions at $n=4{:}1$ and $n=8{:}1$, RC values for each anatomical region and each sphere for Site-4

Sphere (cm)	4:1 Thorax	4:1 Abdomen	4:1 Pelvis	8:1 Thorax	8:1 Abdomen	8:1 Pelvis
0.30	0.087	0.129	0.124	0.112	0.102	0.118
0.50	0.186	0.211	0.182	0.227	0.202	0.192
0.90	0.476	0.407	0.423	0.336	0.305	0.335
1.30	0.497	0.517	0.567	0.467	0.439	0.405
1.90	0.623	0.671	0.668	0.600	0.517	0.534
2.80	0.782	0.738	0.768	0.740	0.726	0.706

ensured the reliability and accuracy of the results (RCs) when comparing clinical data, allowing for a deeper insight into the underlying trends and patterns.

Figure 1 shows respectively for each site; the curves generated from RCs as a function of the sphere diameters, for 3 anatomical regions for $n=4{:}1$ and $n=8{:}1$ concentration ratio.

We analyzed the relationship between sphere diameter and RC (Fig. 1). The x-axis represents the sphere diameter used in the study, ranging from very small diameters (a few millimeters) to larger ones (several centimeters), while the y-axis represents the RCs, typically ranging from 0 to 1. An RC of 1 indicates perfect recovery, where the measured activity equals the true activity, whereas an RC less than 1 indicates underestimation. The graphs (Fig. 1) show that for small spheres, the RC is low due to partial volume effects, where the imaging system struggles to capture the true activity concentration accurately. As the sphere diameter increases, the RC generally increases because larger objects are better resolved by the imaging system, leading to more accurate activity measurements. Eventually, the RC may plateau and approach 1 for larger spheres, indicating that the imaging system can accurately recover the true activity concentration for these objects. In addition, the spatial resolution of the imaging system plays a crucial role, with higher resolution systems yielding better RCs for smaller diameters.

When plotting RCs against the diameters of spheres, the graph typically shows how well the imaging system can recover the true activity concentration for objects of different sizes.

Similar curves are generated from RCs as a function of the sphere diameters, for $n=4{:}1$ and $n=8{:}1$, for all sites, to observe how the RCs varied with tumor sizes (spheres) for 3 anatomical regions; in the thorax (Fig. 2), abdomen (Fig. 3), and pelvis (Fig. 4). These graphs

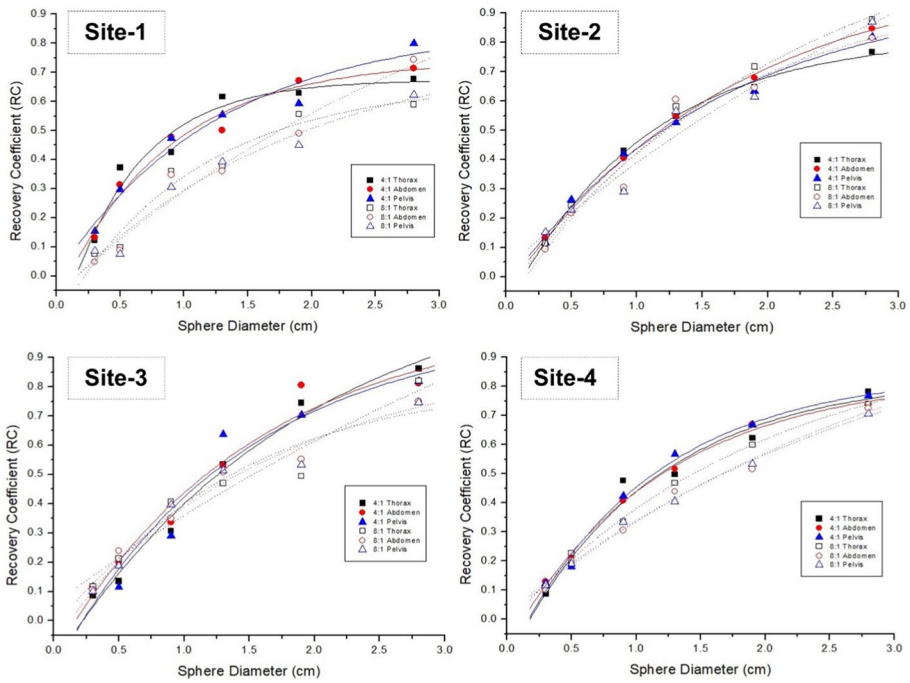

Fig. 1 RCs as a function of the diameter of spheres for all sites

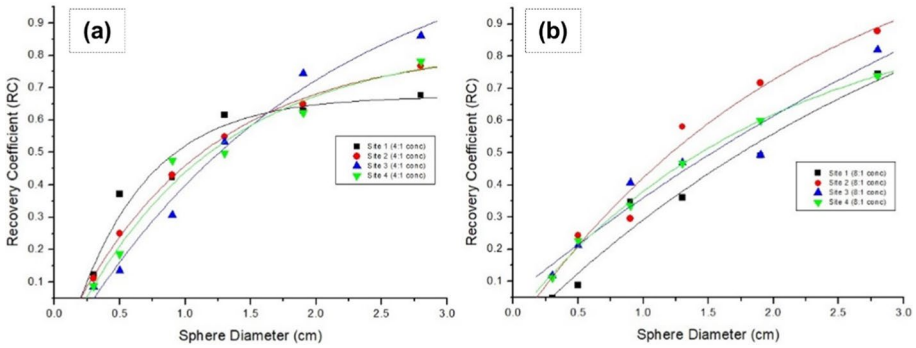

Fig. 2 Variation of RCs with tumor sizes (spheres) in the thorax region for *n* values **a** *n* = 4:1 and **b** *n* = 8:1 for all four sites

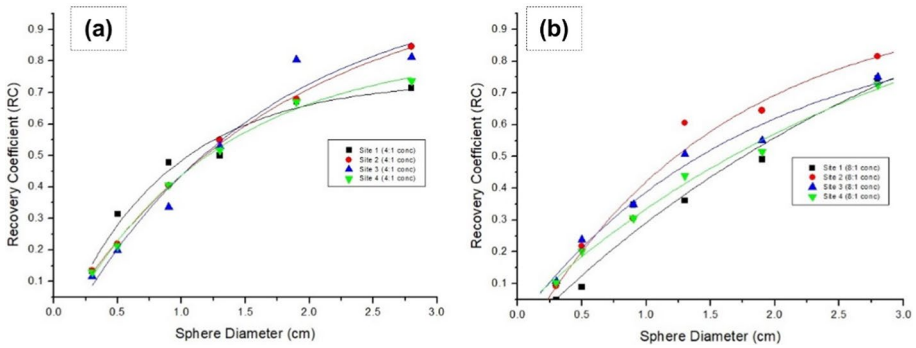

Fig. 3 Variation of RCs with tumor sizes (spheres) in the abdomen region for *n* values **a** *n* = 4:1 and **b** *n* = 8:1 for all four sites

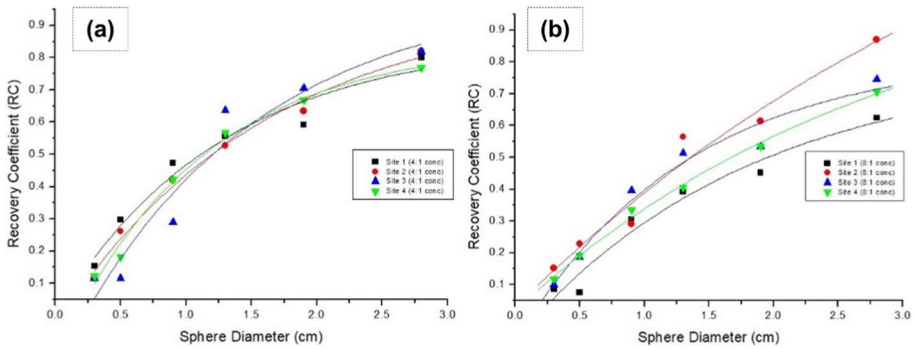

Fig. 4 Variation of RCs with tumor sizes (spheres) in the pelvis region for *n* values **a** *n* = 4:1 and **b** *n* = 8:1 for all four sites

provided a visual representation of the RC's dependency on tumor dimensions, offering insights into the quantitative relationships across different anatomical regions and varying "*n*" values.

Discussion

Accurate SUV measurements are vital for many tumors where metabolic activity measurement allows for the assessment of treatment response. Correcting the uptake obtained from the patient with RC can help to eliminate measurement discrepancies between devices.

Anthropomorphic phantoms with known properties and corrected SUV values can assist in calibrating PET scanners, ensuring that the equipment is functioning correctly and producing reliable data [10, 20].

The use of SUV-corrected values in phantom imaging is essential for ensuring accuracy, consistency, and standardization in PET imaging. These values are integral to quality control, equipment calibration, research, clinical decision-making, and collaborative studies across multiple centers. Employing these standardized measures allows clinicians and researchers to obtain precise and comparable data from PET imaging, which in turn enhances patient outcomes and drives progress in medical imaging technology [20].

One major factor contributing to inaccurate SUV measurements is the small size of the lesion, which frequently leads to underestimation of the measured values. As the diameter of the sphere (or tumor) decreases, the impact of the PVE becomes more pronounced. This situation highlights how scatter radiation outside the FOV (Field of View) significantly reduces counts in smaller-diameter regions by increasing the dead time of the detection system [21]. Conversely, for larger spheres, the influence of PVE is less pronounced, and scatter radiation outside the FOV does not significantly reduce counts in these regions. These findings are consistent with studies by Matheoud et al. and Krempser et al. and our results are similar [1, 22].

In PET and SPECT imaging, accurate calculation of RCs necessitates accounting for activity outside the FOV, which is essential for improving the precision of RC-based PVE correction techniques. This approach is particularly crucial as it addresses the influence of scattered radiation and spill-over effects, leading to a more comprehensive correction methodology [22].

In the clinical setting, estimating scatter from outside the FOV is challenging. Nevertheless, the use of anthropomorphic phantoms represents a valuable approach for optimizing the determination of RCs in clinical acquisitions [23].

Analysis of this study's results, consistent with Krempser et al.'s investigation [1], indicates that the variability in sphere diameter among RCs is more pronounced than the variability related to the n value. Notably, RCs exhibit significant variation as the sphere diameter ranges from 0.3 cm to 0.9 cm. This observation is critical as it demonstrates that smaller spheres are more susceptible to variability, which must be accounted for in quantitative analyses. Furthermore, all RC curves tend to stabilize at the unit value of 1, indicating that the measured activity concentration equals the known activity concentration as the sphere diameter increases from 0.3 cm to 2.8 cm. This stabilization highlights a threshold beyond which further increases in sphere diameter do not significantly impact RCs, simplifying the interpretation of results for larger spheres.

In addition, it was observed that RCs increase with higher n values, indicating a correlation between these n values and RC stability. This relationship enhances our understanding of how sphere diameter, n values, and RCs interact, offering valuable insights

into the complexities of PVE in PET imaging. This nuanced understanding can enhance PET imaging accuracy by better accounting for PVE, leading to more reliable quantification of metabolic. Acknowledging these factors emphasizes the importance of considering sphere diameter and *n* values in PET study design and interpretation to mitigate PVE and improve the robustness of imaging results.

In our study involving the unique anthropomorphic phantom, the primary aim was not to compare the RC values with those obtained from the NEMA gold standard phantom, due to differences in spherical sizes and phantom characteristics. However, upon reviewing reference studies that utilized the NEMA phantom, it was observed that all of its spherical lesion sizes exceed 1 cm. Notably, the 1.3 cm and 2.8 cm lesion sizes from both phantoms can be directly compared.

In our study, RC determined reached close to 0.9 (equivalent to 90% of recovery counts) for sphere diameters of 2.8 cm with anthropomorphic phantom studies. In contrast, Srinivas et al. reported RCs up to 0.8, and Krempser et al. achieved up to 1 in their investigations with PET-NEMA phantoms in 1:8 ratio.

Furthermore, when examining the 1.3 cm lesion size in both reference studies, we observed that the RC values differed significantly. In Srinivas' study, the RC value was 0.25, while in Krempser's study it was 0.65. In contrast, our study yielded an RC value of 0.6 for the same lesion size. These discrepancies may be attributed to differences in phantom characteristics, imaging techniques, or other experimental factors [1, 14].

RCs exhibit a notable specificity based on the unique characteristics of a given scanner and the specific radiopharmaceuticals used. Consequently, they cannot be universally applied for PVE correction across different scanners or for images obtained with the same scanner but using a radiopharmaceutical other than F18-FDG, as emphasized by Gallivanone et al. [19].

In our study, we extended the experimental replication to assess RCs, considering scanners even from the same manufacturer (Site 1 and Site 4) but of different models. This approach was motivated by the recognition of the distinct influence exerted by each detection system and the parameters of image reconstruction algorithms. The findings underscore the need for a tailored and scanner-specific approach when addressing PVE in quantitative analyses, taking into account both the technological specifications and reconstruction algorithms of the imaging systems employed [24–26]. This consideration becomes particularly relevant in multi-site or multi-model studies where harmonizing RCs across different systems is essential for accurate and consistent quantitative assessments.

We observed variability in RCs across different PET/CT systems in different detector Technologies (digital SiPM, Silicon Photo Multiplier, vs analog PMT, Photo Multiplier Tube), driven by key factors. Detector technology plays a major role in sensitivity and temporal resolution. TOF (Time of Flight) technology reduces noise and enhances spatial resolution, while non-TOF systems may struggle with accuracy. Variations in reconstruction algorithms (e.g. GE's Q. Clear, Siemens' HD-PET), axial field of view, system sensitivity, and equipment age or maintenance of the equipment also contribute to the observed differences.

We obtained RC values for six different sphere sizes in three regions: thorax, abdomen, and pelvis. These RC values have allowed us to standardize the mean SUV

values of tumors of equivalent sizes obtained from real clinical patients, making them independent of factors related to the device and patient attenuation. This enables the elimination of differences in uptake values obtained using different devices, particularly during patient follow-up, providing the opportunity for accurate comparison and interpretation of PET results conducted at different centers.

Krempser et al. stated in their study [1] that as the *n* values increase, the RC also increases. However, in the study conducted at Site-1, it was observed that the RC values decreased. This discrepancy is thought to be due to the time differences between the loading of activity and the initiation of imaging.

In delayed imaging, as the activity count within the phantom decreases, the noise in the images increases, leading to greater uncertainty in the measurements. This rise in noise levels can significantly affect the accuracy and precision of the RC values, potentially skewing the results. The increased uncertainty might be attributed to the diminished signal-to-noise ratio (SNR), which poses a challenge in accurately quantifying the RC, especially in lower activity regions. Future studies should consider optimizing the timing of activity loading and imaging initiation to mitigate these effects and ensure more reliable and consistent RC measurements.

The primary objective of our study, conducted using standard activity within phantoms, was not to compare the performance of different devices. Instead, our focus was on calculating RC values for various sphere sizes and body localizations specific to each device. For each system, we generated a curve representing the RC values, which we have detailed and illustrated in the study.

As observed from the graphs generated in this study, it is evident that the RC values measured in the thorax, abdomen and pelvis are similar across all centers. This result indicates that the RC values do not vary significantly with the position of the body. The consistency of RC values across different anatomical regions suggests that the RC is not influenced by the specific location within the body. This finding is important as it underscores the robustness of the RC as a parameter, implying that it can be reliably used for quantitative assessments regardless of the anatomical region being imaged. Such uniformity enhances the utility of RC measurements in clinical and research settings, ensuring that comparisons and interpretations of RC data are valid and not confounded by positional differences. Further studies could expand on this by exploring other potential factors that might affect RC values, thereby providing a more comprehensive understanding of the parameters influencing RC measurements.

Gallivanone et al. applied the statistical *t*-test to obtain *p*-values for comparing lesion sizes and SUV values in oncological patient groups [2]. Similarly, Bettinardi et al., in their review article, highlighted the statistical significance of correlations between PET biomarkers and biological prognostic indices or survival endpoints [10]. Furthermore, Koopman et al. also employed the *t*-test in their statistical analysis to compare patient and scan characteristics between two scanning groups, using *p*-values for these comparisons [25]. In this study, *p*-values were calculated using linear regression analysis; where the RC values calculated at different sites for all three anatomical regions and both *n* (1:4, 1:8) ratios were processed.

Two separate linear regression models were applied using R: a Language and Environment for Statistical Computing to evaluate the impact of various factors on the dependent variable (RC).

First model involves all parameters and considers direct effects only. Results of this first model, parallel to our study, demonstrate that while region does not appear to influence the dependent variable ($p > 0.1$), Site-2 ($p < 0.001$) and Site-3 ($0.01 < p < 0.1$) do have a statistically significant impact.

Second model was employed to understand combined effect of sphere size and sites, where region parameter was removed. The results from this second model confirms combined effect of sphere size and Site-2 ($0.001 < p < 0.01$) and Site-3 ($0.001 < p < 0.01$) has statistically significant effect. These results also show effect of site increases with increasing sphere size, as all direct site parameters have $p > 0.1$.

Analyzing the pixel sizes of the PET/CT systems at each center reveals that, for a system with a resolution of 5 mm and a pixel size of 4 mm, placing a 3-mm sphere simulating a lesion for measurement illustrates the challenges of detecting activity uptake in small-scale lesions.

This finding underscores the limitations of current PET/CT technology in accurately quantifying RC values for very small lesions. The discrepancy likely arises due to the resolution constraints, where the partial volume effect becomes more pronounced, and the activity distribution within such small volumes is not accurately captured. Consequently, this leads to increased uncertainty and variability in the measured RC values, making them unreliable for clinical or research applications in small lesions.

These limitations highlight the need for advancements in imaging technology to improve spatial resolution, detector sensitivity and reduce pixel size, thereby enabling more accurate detection and quantification of activity in smaller lesions. Future research should focus on developing enhanced imaging techniques and algorithms that can better account for these limitations, potentially incorporating high-resolution imaging systems or advanced image reconstruction methods to achieve more reliable results in small-scale lesion measurements.

As indicated above, in all devices, the obtained RC value for 0.3 cm is quite low. The highest value is achieved for the Thorax region at Site-1, while the lowest results obtained at Sites 3 and 4 are close to each other. Across centers, as the lesion size increases, the RC values approach each other. Therefore, the sphere simulating the lesion, with a size of 0.3 cm, is below the resolution limit for all devices, and semi-quantitative values obtained from lesions of this size are not reliable for diagnosis and follow-up treatment response.

At Site 4, where the digital PET/CT system is located, although the image quality is very high, and lesion detection is at an advanced level (since system has TOF and Q-Clear as well), the low RC values compared to other sites (even in large spheres, around 0.7, 0.8 ratios) may indicate a potential error in SUV calibration. For each device and each size, RCs have been obtained that will enable the standardization of quantitative values. Our results indicate that similar studies could contribute to the assessment of patients in the future, especially the ones with small tumors. A larger scale chart study needs to be conducted to provide a more extensive dataset for all

targeted tumor sizes using similar specialized anthropomorphic phantoms and the methodology described in this study.

The gold standard for RC calculations is the NEMA phantom, and the spheres within the NEMA phantom are typically used for measurements. However, since the sphere sizes in the NEMA phantom are larger than the sub-centimeter spheres used in our study, a direct comparison of our results with the NEMA phantom was not possible. In our study, the primary reason for using an anthropomorphic phantom was to take body habitus into account, which is not considered in NEMA, and to use it as a parameter. The phantom was designed to account for variations in tissue types, including adipose tissue, bone tissue, and body shape, organ structures. This was partially achieved with this study.

While our anthropomorphic phantom simulates a standard body type, future studies incorporating a wider range of phantom sizes or patient-specific simulations could provide further insight into the influence of body habitus on PVE correction.

Tumor heterogeneity is outside the scope of this study. While our anthropomorphic phantom study partially addresses factors affecting body habitus, we believe that tumor heterogeneity is not an area that can be evaluated with the phantom we used but requires a distinct and separate research study.

In this study, we have taken a significant step by utilizing an anthropomorphic phantom as a novel approach with the aim of achieving standardization across all systems in the future. The primary goal is to introduce the anthropomorphic phantom into the field of nuclear medicine (NM).

Our objective is not to compare devices, but rather to develop advanced versions of the anthropomorphic phantom, including organ-specific models and—in the future—even respiration-triggered ones. Our initial aim is to raise awareness of the potential of the widespread use of anthropomorphic phantoms.

We acknowledge the continuous evolution of PET/CT technology, particularly with recent advancements in digital PET systems, machine learning-based image reconstruction, and other innovations. In light of these developments, we believe that the framework for calculating RCs and correcting PVE, as outlined in our study, remains adaptable to future technological advancements. The core methodology, specifically the use of anthropomorphic phantoms, is designed to be robust and flexible, allowing for adjustments as new systems and techniques emerge.

To further enhance the generalizability of our findings, we recommend validation across a broader range of PET/CT systems, including next-generation scanners. By applying our methodology to newer systems, and potentially incorporating artificial intelligence-driven reconstruction methods, the proposed PVE correction approach could be extended and further validated.

Conclusion

This study is pioneering as it marks the first application of an anthropomorphic phantom in lesions smaller than 1 cm (0.3, 0.5, 0.9) in molecular imaging. A primary objective for future research is to develop phantoms that simulate various life stages—such as pediatric, adolescent, adult, and elderly patient groups—along with different body volumes. Ultimately, our aim is to create a virtual phantom using AI-supported

software to enhance and standardize quantitative data across patient groups by utilizing recovery coefficient (RC) values derived from these phantoms.

In this preliminary study, we believe we have reached significant initial results to determine the RC values necessary for accurately measuring the tumoral radiopharmaceutical uptake in patients.

There are two main reasons for determining corrected SUV values. The first reason is that it allows for the development of a perspective to correct the significant numerical discrepancies caused by different PET/CT systems operating worldwide. This correction will guide clinicians on the percentage changes in SUV to consider, particularly regarding treatment response. The second reason is that it will enable the use of corrected SUV parameters, independent of the device, in algorithms and evaluations used in the rapidly proliferating artificial intelligence applications, facilitating their use in patient assessments in a much shorter time.

The standardization of PET imaging results is a fundamental requirement for the accurate evaluation of molecular imaging outcomes. Therefore, we believe that the RC values derived from the results of our study using the first prototype of the anthropomorphic phantom, which we consider to be the most realistic, could significantly contribute to the increased confidence in the numerical results obtained from molecular imaging in the coming years. However, there is still a need for the development of more realistic anthropomorphic phantoms and the use of these phantoms in a multicenter setting across numerous PET/CT systems to establish generally accepted RC values.

Materials and methods/methodology

PET/CT scanners

Scans were conducted at four different hospitals using four PET/CT systems from three vendors, incorporating two distinct technologies. Table 5 provides the technical; specifications of these systems [24, 27, 28].

At these four sites, a total of 144 scans were performed across the thorax, abdomen, and pelvis regions for six different sphere sizes, using both 4:1 and 8:1 activity ratios (72 scans per ratio). The 18-F FDG radioactivity dose was diluted in saline and injected into the cylinders and spheres under standardized conditions as outlined in Table 6.

After phantom imaging at each hospital, the effects of different detector technologies were examined by calculating RCs for PVE correction.

Table 5 Technical specifications of the PET/CT systems used in this study

Site	System	Crystal	Technology	Matrix	Pixel sizes	Resolution
1	GE Discovery IQ-4 Ring	BGO	Q-Clear	192 × 192	3.65	5.1 mm @1 cm CFOV
2	Siemens Biograph mCT Excel 20	LSO	TOF, UltraHD	200 × 200	4.07	4.4 mm @ CFOV
3	Philips GEMINI TF	LYSO	TOF	144 × 144	4	4.7 mm @1 cm CFOV
4	GE Discovery MI Digital—(SiPM) PET	LYSO	TOF, Q-Clear	256 × 256	2.73	3.95 mm

Table 6 Applied concentration ratio and dose for spheres and cylinders

	Saline volume	Concentration ratio	Concentration dose (FDG)
Cylinders (Background)	1000 ml	1	2 mCi
Spheres 4:1	250 ml	4	2 mCi
Spheres 8:1	125 ml	8	2 mCi

Anthropomorphic human body phantom

This study used a custom-designed anthropomorphic body phantom to simulate varying tissue attenuation across the thorax, abdomen, and pelvis regions, as illustrated in Fig. 5A [29]. The phantom, cylinders, and spheres were produced in-house. Six spheres of varying diameters (0.3, 0.5, 0.9, 1.3, 1.9, and 2.8 cm) were positioned in cylinders placed within the phantom (Fig. 5B and C).

Cylinders were filled with an F18-FDG solution with a background concentration of 2 mCi. Sphere-to-background ratios of 4:1 and 8:1 were achieved by filling the spheres with higher activity concentrations.

Image acquisition, reconstruction, and quantification

Anthropomorphic Body Phantom positioned in PET/CT System as illustrated in Fig. 6A. Imaging was performed under uniform conditions across all PET/CT systems. Whole-body PET/CT scans were conducted using static imaging protocols. For each set of spheres, scans were performed in the thorax, abdomen, and pelvis regions, with the spheres positioned in successive orders to ensure comprehensive data acquisition.

Images were reconstructed using system-specific algorithms, including TOF, Ultra-HD, and Q-Clear, according to each manufacturer's technology. During each acquisition,

Fig. 5 In-house customized anthropomorphic human body phantom (**a**), cylinders (**b**) and spheres (**c**)

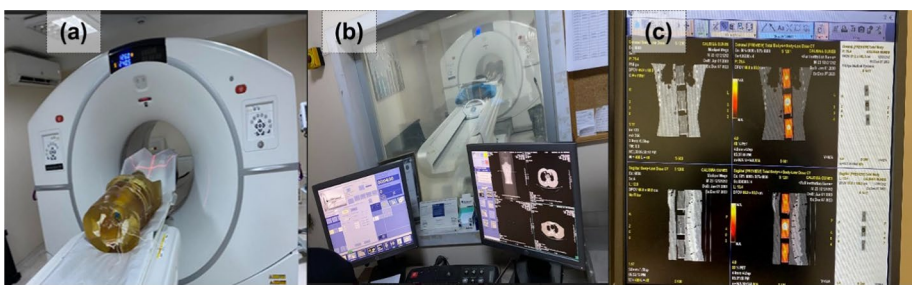

Fig. 6 Anthropomorphic human body phantom positioning, image acquisition and reconstruction

the number of bed positions varied based on the system's FOV, with each bed position lasting 2 min (Fig. 6B). Attenuation correction was performed using a CT scan with standardized parameters across all sites. Image evaluation was performed using post-processing reconstruction workstation (Fig. 6C).

For quantification, the central slice corresponding to the center of the spheres was used. Circular volumes of interest (VOIs) were manually drawn over each sphere image to measure maximum activity concentrations. RCs for each sphere size were calculated using Eq. 2, taking into account activity decay to ensure accurate quantification:

$$\mathrm{RC} = \frac{\text{Measured Activity Conc.(Sphere)} - \text{Measured Activity Conc.(BackGround)}}{\text{Known Activity Conc.(Sphere)} - \text{Known Activity Conc.(BackGround)}} \tag{2}$$

The RCs, dimensionless values ranging from 0 to 1, reflect the scanner's accuracy in measuring activity concentrations within the spheres, thus accounting for varying tissue attenuation and enhancing the reliability of the results [1].

We employed linear regression analysis to evaluate the influence of these variables on the Recovery Coefficient (RC). This enabled us to examine the relationships and statistical significance of predictors such as sphere size, site, and region on RC values. The results underscore significant predictors, such as sphere size and specific site effects, while indicating that region-related variables may have minimal or negligible impact.

Abbreviations

EIC	International Electrotechnical Commission
FDG	Fluorodeoxyglucose
FOV	Field of view
NEMA	National Electrical Manufacturers Association
PET/CT	Positron Emission Tomography/Computed Tomography
PVE	Partial volume effect
RC	Recovery coefficients
ROI	Region of interest
SNR	Signal-to-noise ratio
SUV	Standardized uptake value
TOF	Time of flight
VOI	Volume of interest

Acknowledgements
Not applicable.

Author contributions
Conceptualization: G. Yavuz, T.F. Cermik, C. Ozturk; methodology and experimentation: G. Yavuz, B. Kovan, C. Ozturk; post-processing and reconstruction: G. Yavuz, T. Toklu; writing—original draft: G. Yavuz; writing—review and editing: C. Ozturk, T.F. Cermik. All the authors have read and agreed to the submitted version of the manuscript. All authors have contributed to the material presented, and this work has not been published elsewhere.

Funding
This research was funded by Boğaziçi University Research Fund, Grant Number 19944.

Availability of data and materials
No datasets were generated or analyzed during the current study.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 7 August 2024 Accepted: 6 January 2025
Published online: 15 February 2025

References

1. Krempser AR, Ichinose RM, de Sá AMM, de Oliveira SMV, Carneiro MP. Recovery coefficients determination for partial volume effect correction in oncological PET/CT images considering the effect of activity outside the field of view. Ann Nuclear Med. 2013;27:924–30.
2. Gallivanone F, Canevari C, Gianolli L, Salvatore C, Della Rosa P, Gilardi M, et al. A partial volume effect correction tailored for 18F-FDG-PET oncological studies. Biomed Res Int. 2013;2013:780458.
3. Boellaard R. Standards for PET image acquisition and quantitative data analysis. J Nucl Med. 2009;50:11S-20S.
4. Wu Z, Guo B, Huang B, Hao X, Wu P, Zhao B, et al. Phantom and clinical assessment of small pulmonary nodules using Q. Clear reconstruction on a silicon-photomultiplier-based time-of-flight PET/CT system. Sci Rep. 2021;11:10328.
5. Lu S, Zhang P, Li C, Sun J, Liu W, Zhang P. A NIM PET/CT phantom for evaluating the PET image quality of microlesions and the performance parameters of CT. BMC Med Imaging. 2021;21:1–13.
6. Adler S, Seidel J, Choyke P, Knopp MV, Binzel K, Zhang J, et al. Minimum lesion detectability as a measure of PET system performance. EJNMMI Phys. 2017;4:1–14.
7. Øen SK, Aasheim LB, Eikenes L, Karlberg AM. Image quality and detectability in Siemens Biograph PET/MRI and PET/CT systems—a phantom study. EJNMMI Phys. 2019;6:1–16.
8. Soret M, Bacharach SL, Buvat I. Partial-volume effect in PET tumor imaging. J Nucl Med. 2007;48:932–45.
9. Kessler RM, Ellis Jr JR, Eden M. Analysis of emission tomographic scan data: limitations imposed by resolution and background. LWW; 1984.
10. Bettinardi V, Castiglioni I, De Bernardi E, Gilardi M. PET quantification: strategies for partial volume correction. Clin Transl Imaging. 2014;2:199–218.
11. Alavi A, Werner TJ, Høilund-Carlsen PF, Zaidi H. Correction for partial volume effect is a must, not a luxury, to fully exploit the potential of quantitative PET imaging in clinical oncology. Mol Imaging Biol. 2018;20:1–3.
12. Driscoll B, Shek T, Vines D, Sun A, Jaffray D, Yeung I. Phantom validation of a conservation of activity-based partial volume correction method for arterial input function in dynamic PET imaging. Tomography. 2022;8:842–57.
13. Grings A, Jobic C, Kuwert T, Ritt P. The magnitude of the partial volume effect in SPECT imaging of the kidneys: a phantom study. EJNMMI Phys. 2022;9:18.
14. Srinivas SM, Dhurairaj T, Basu S, Bural G, Surti S, Alavi A. A recovery coefficient method for partial volume correction of PET images. Ann Nucl Med. 2009;23:341–8.
15. Hoffman EJ, Cutler PD, Guerrero TM, Digby WM, Mazziotta JC. Assessment of accuracy of PET utilizing a 3-D phantom to simulate the activity distribution of [18F] fluorodeoxyglucose uptake in the human brain. J Cereb Blood Flow Metab. 1991;11:A17-25.
16. Hoffman EJ, Huang S-C, Phelps ME. Quantitation in positron emission computed tomography: 1. Effect of object size. J Computer Assist Tomogr. 1979;3:299–308.
17. Rousset O, Rahmim A, Alavi A, Zaidi H. Partial volume correction strategies in PET. PET clinics. 2007;2:235–49.
18. Zaidi H, Ruest T, Schoenahl F, Montandon M-L. Comparative assessment of statistical brain MR image segmentation algorithms and their impact on partial volume correction in PET. Neuroimage. 2006;32:1591–607.
19. Gallivanone F, Stefano A, Grosso E, Canevari C, Gianolli L, Messa C, et al. PVE correction in PET-CT whole-body oncological studies from PVE-affected images. IEEE Trans Nucl Sci. 2011;58:736–47.
20. Lennie E, Tsoumpas C, Sourbron S. Multimodal phantoms for clinical PET/MRI. EJNMMI Phys. 2021;8:62.
21. Meechai T, Tepmongkol S, Pluempitiwiriyawej C. Partial-volume effect correction in positron emission tomography brain scan image using super-resolution image reconstruction. Br J Radiol. 2015;88:20140119.
22. Matheoud R, Secco C, Della Monica P, Leva L, Sacchetti G, Inglese E, et al. The effect of activity outside the field of view on image quality for a 3D LSO-based whole body PET/CT scanner. Phys Med Biol. 2009;54:5861.
23. Fedrigo R, Kadrmas DJ, Edem PE, Fougner L, Klyuzhin IS, Petric MP, et al. Quantitative evaluation of PSMA PET imaging using a realistic anthropomorphic phantom and shell-less radioactive epoxy lesions. EJNMMI Phys. 2022;9:2.
24. Hsu DF, Ilan E, Peterson WT, Uribe J, Lubberink M, Levin CS. Studies of a next-generation silicon-photomultiplier–based time-of-flight PET/CT system. J Nucl Med. 2017;58:1511–8.
25. Koopman D, van Dalen JA, Stevens H, Slump CH, Knollema S, Jager PL. Performance of digital PET compared with high-resolution conventional PET in patients with cancer. J Nucl Med. 2020;61:1448–54.
26. Salvadori J, Odille F, Verger A, Olivier P, Karcher G, Marie P-Y, et al. Head-to-head comparison between digital and analog PET of human and phantom images when optimized for maximizing the signal-to-noise ratio from small lesions. EJNMMI Phys. 2020;7:1–14.
27. Reynés-Llompart G, Gámez-Cenzano C, Romero-Zayas I, Rodríguez-Bel L, Vercher-Conejero JL, Martí-Climent JM. Performance characteristics of the whole-body discovery IQ PET/CT system. J Nucl Med. 2017;58:1155–61.
28. Rausch I, Cal-González J, Dapra D, Gallowitsch HJ, Lind P, Beyer T, et al. Performance evaluation of the Biograph mCT Flow PET/CT system according to the NEMA NU2-2012 standard. EJNMMI Phys. 2015;2:1–17.
29. Kovan B. İnternal Radyonüklid Tedavisinde Organ Doz Hesaplama Parametrelerinin Optimizasyonu. Istanbul University, Institute of Science Studies; 2021.

Publisher's Note