Radiation Physics and Chemistry 237 (2025) 113028

Contents lists available at ScienceDirect

Radiation Physics and Chemistry

journal homepage: www.elsevier.com/locate/radphyschem

Deep learning-based segmentation of Jaszczak ACR phantom images for optimized Radium-223 dosimetry

Cristian Felipe Griebler [a,*], Leanderson Pereira Cordeiro [b], Luis Felipe Lima [c], Vagner Bolzan [d], Vitor Dutra [d], Lidia Vasconcellos De Sá [a], Daniel Alexandre Baptista Bonifacio [e]

[a] Instituto de Radioproteção e Dosimetria, Brazil
[b] Universidade Federal de Pernambuco, Brazil
[c] Universidade Federal Do Rio de Janeiro, Brazil
[d] Instituto Federal de Santa Catarina, Brazil
[e] Energy and Nuclear Research Institute (IPEN), Brazil

ARTICLE INFO

Handling Editor: Piotr Ulanski

Keywords:
Dosimetry
Quantification
SPECT calibration
Artificial intelligence
Monte Carlo method

ABSTRACT

Precise and personalized absorbed dose estimation in radionuclide therapy is crucial for optimizing treatment efficiency while minimizing harm to healthy tissues. Radium-223 dichloride (Ra-223), an alpha emitter used in treating metastatic castration-resistant prostate cancer, has shown positive results in extending patient survival. However, the current practice of uniform Ra-223 activity administration based solely on patient weight can lead to suboptimal treatment outcomes. Treatment efficacy evaluation involves quantifying activity and absorbed dose through image quality analysis, revealing potential areas for optimization. This work introduces an innovative approach that integrates a deep learning-based model for automated segmentation of the Jaszczak ACR phantom—a tool for image quality analysis in nuclear medicine—with Monte Carlo simulation for dosimetry. The model exhibits efficient segmentation, surpassing 83.7 % in class-wise Dice coefficients, offering a time-efficient alternative to manual segmentation. The study highlights the superior performance of the 89 keV energy window in image quality parameters, emphasizing its role in lesion detection. Additionally, it addresses challenges in achieving accurate quantitative outcomes in nuclear medicine applications, particularly in Ra-223 therapy. These insights contribute to refining dosimetry protocols for Ra-223, enhancing the precision of quantitative outcomes in nuclear medicine. The practical implications extend to improving daily routines for clinical professionals in nuclear medicine applications, showcasing the potential of advanced imaging techniques and computational tools in optimizing Ra-223 therapy.

1. Introduction

Precise dosimetry is essential in radionuclide therapy to achieve optimal therapeutic outcomes while minimizing radiation exposure to healthy tissues. Radium-223 (Ra-223), an alpha-emitting radionuclide, has become a prominent option for the treatment of metastatic castration-resistant prostate cancer (mCRPC), especially in cases involving bone metastases. Ra-223 mimics calcium and localizes to areas of high bone turnover, where it emits high-energy alpha particles. These characteristics make it effective for targeted radiotherapy, providing a survival benefit while limiting damage to adjacent healthy tissue due to the short range of the alpha particles (Danieli et al., 2022; Morris et al., 2019).

Despite these benefits, the current administration of Ra-223, which is based solely on patient weight (at a standardized dose of 55 kBq/kg), does not account for patient-specific or lesion-specific attributes. This uniform approach can lead to inconsistent absorbed doses, resulting in either suboptimal or excessive radiation exposure across patients and lesion sites (Deshayes et al., 2017; Parker et al., 2013; Higano et al., 2023; Sgouros et al., 2004).

Personalized dosimetry is essential for optimizing the therapeutic efficacy of Ra-223, as it allows for accurate quantification of the absorbed dose to the tumor and critical organs, such as the bone marrow, while reducing unnecessary radiation exposure to non-target tissues (Morris et al., 2019; Sandström et al., 2020). The European Council Directive (2013/59/EURATOM) emphasizes the importance of

* Corresponding author.
E-mail address: cristiangrb@gmail.com (C.F. Griebler).

https://doi.org/10.1016/j.radphyschem.2025.113028
Received 6 October 2024; Received in revised form 31 March 2025; Accepted 2 June 2025
Available online 6 June 2025

dose optimization in radiation therapy, highlighting the limitations of weight-based dosing strategies and the need for individualized treatment protocols (Council, 2014).

Single-photon emission computed tomography (SPECT) combined with computed tomography (CT) has become a key tool in nuclear medicine, providing a hybrid imaging modality for quantitative assessments of radiopharmaceutical distribution. However, accurate quantification of the absorbed dose relies on precise segmentation of volumes of interest (VOIs) to assess image quality. The Jaszczak ACR phantom is commonly used in this context for calibration and validation of quantitative.

SPECT imaging techniques (Dickson et al., 2023; Gear et al.). However, manual segmentation of the Jaszczak phantom is time-consuming and susceptible to variability.

Advances in artificial intelligence (AI) have introduced the possibility of automated image segmentation, offering significant improvements in efficiency and accuracy. AI-based models, particularly deep learning approaches, have demonstrated potential in medical image analysis by enhancing lesion detection, segmentation, and dosimetry (Jiadong, 2023).Convolutional neural networks (CNNs), such as UNET and its variants, have shown promising results in segmenting complex medical images, enabling precise volumetric quantification of VOIs and improving dosimetry accuracy (Diaz-Pinto et al., 2022, 2022).

This study presents an approach that integrates a deep learning-based model for automated segmentation of the Jaszczak ACR phantom with Monte Carlo-based dosimetry. Automating phantom segmentation aims to reduce the time and variability associated with manual segmentation, thereby improving the accuracy of Ra-223 dosimetry. The dosimetric implications of this approach are evaluated by assessing the impact of accurate recovery coefficient determination on dosimetry calculations for Ra-223 therapy. The combination of AI-based segmentation and advanced dosimetry tools has the potential to refine current Ra-223 treatment protocols and improve therapeutic outcomes for patients undergoing radionuclide therapy.

2. Materials and methods

Dosimetry workflow is represented by Fig. 1. Imaging procedures were conducted using the Symbia T2 SPECT/CT equipment (Siemens Medical Solution Inc., USA) with a medium-energy general-purpose (MEGP) collimator, as recommended in previous studies (Benabdallah et al., 2019; Owaki et al., 2017).

2.1. Jaszczak ACR phantom

This study employed the Jaszczak ACR phantom for image quality analysis. This phantom consists of four fillable plastic cylinders designed to mimic lesion uptake, with volumes of 2, 4.5, 8 and 19 ml and corresponding diameters of 8, 12, 16 and 26 mm. Additionally, three supplementary cylinders, each with a 19 ml volume, were incorporated into the phantom, composed of varying materials to simulate bone tissue (Teflon), water-filled plastic, and air-filled plastic. The total volume of the Jaszczak ACR phantom, inclusive of all cylinders, amounts to 6.815 L (Hindorf et al., 2012).

2.2. Administered activity

According to the pharmaceutical company (Pacilio et al., 2016), up to 77 % of the administered activity is absorbed by the bone tissue, leaving the remaining 23 % circulating throughout the body. Considering the prescribed radiopharmaceutical dosage of 55 kBq/kg (Pacilio et al., 2016) and taking into account the total volume of the Jaszczak phantom (6815 L), the activity to be administered is 374.82 kBq.

Relying on the proportion of absorption by bone tissue, the experimental activity concentration within the cylinders was 8.14 kBq/mL, while it was 0.0179 kBq/mL within the Jaszczak ACR phantom body. In the sensitivity test, the experimental activity introduced into the syringe was also 8.14 kBq/mL.

2.3. Image reconstruction

The energy windows were set at 89 keV (25 % width), 154 keV (20 % width), 280 keV (20 % width), and a triple energy window (TEW) configuration. SPECT acquisitions employed a circular orbit with a 2.8126° rotation step at 31-s intervals, totaling 180° with 64 views for each head. This configuration was chosen to ensure a balanced photon statistics profile and a feasible total acquisition time of 33 min for prospective patient studies. Image reconstructions utilized OSEM/MLEM with 8 subsets with 4 iterations based on a previous study (Diaz-Pinto et al., 2022). A Gaussian filter with 4 mm FWHM (full width at half maximum) was applied. Attenuation correction was performed through attenuation maps obtained from the CT acquisition. Reconstructed images had a matrix size of 128 × 128, as previous studies (Benabdallah et al., 2019; Owaki et al., 2017; Hindorf et al., 2012; Pacilio et al., 2016; Lima et al., 2022; Xofigo, 2018).

2.4. Deep learning-based Jaszczak ACR phantom segmentation

A set of 27 CT images was used for model training, with an additional set of five CT images employed specifically for synthetic data generation. To enhance dataset diversity and increase the robustness of the model, data augmentation techniques were applied (Goceri, 2023). These techniques involve the processes of rotation, translation, and flipping to create synthetic data points. The original data acquisition was performed using the Symbia T2 Siemens SPECT/CT.

For the training process, MONAI Label (Diaz-Pinto et al., 2024) was used in conjunction with Slicer 3D software (Fedorov et al., 2012). The MONAI Label tool supports automated segmentation as an annotation approach, primarily using a non-interactive algorithm based on a standard convolutional neural network (CNN), such as UNET (Yin et al., 2022), nnU-Net (Isensee et al., 2021), or UNETR (Hatamizadeh et al., 2022). In contrast, Monai offers an interactive segmentation model, DeepGrow, in which the user actively participates by providing positive and negative clicks. Positive clicks are used to expand the segmentation to include areas of interest, whereas negative clicks are employed to contract the segmentation by excluding specific regions from the area of

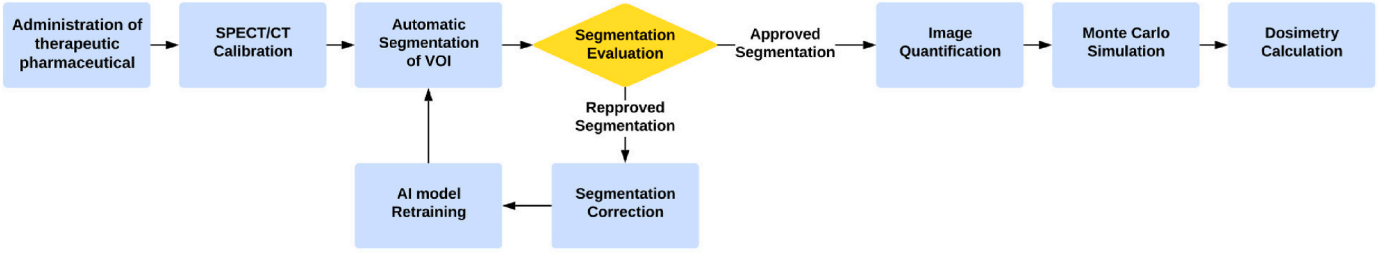

Fig. 1. Dosimetry workflow with automatic segmentation and retraining of Artificial Intelligence model step.

interest (Diaz-Pinto et al., 2024).

DeepEdit extends the concept of DeepGrow's click-based segmentation by enabling both click-free segmentation inference and click-based segmentation editing (Diaz-Pinto et al., 2022a). The key distinctions lie in how this model is trained and the composition of the input tensor's channels. During the training process, the input tensor can take one of two forms: It can either consist of the image with zeroed tensors in automatic segmentation mode, or it may incorporate tensors representing user-provided label and background clicks in interactive mode (Diaz-Pinto et al., 2022a).

The DeepEdit model's algorithms were adapted to recognize five distinct labels corresponding to the structure of the ACR phantom. These labels include four designations for the fillable cylinders, which are referred to as cylinder 1 (2 ml), 2 (4.5 ml), 3 (8 ml), and 4 (19 ml), as well as a background label. The dataset comprised 27 CT images and a split ratio of 0.2. Consequently, 21 images were designated for the training phase, while the remaining 5 images were set aside for evaluation. All 27 images used in the model were manually labeled using the Segment Editor tool in Slicer 3D (Fedorov et al., 2012).

The used optimizer was Adam with an initial learning rate of 0.001. Dice coefficient and mean Dice were employed as key metrics to evaluate the accuracy of the model (Diaz-Pinto et al., 2024). These metrics are adopted in the field of medical image analysis for their efficacy in quantifying the overlap between predicted and ground truth segmentations (Diaz-Pinto et al., 2022a).

2.5. Sensitivity test

A sensitivity test was conducted to evaluate the system's responsiveness to varying input activities. This involved injecting a known activity (8.14 kBq/mL) into the 8 ml syringe to simulate a lesion, allowing for the quantitative assessment of the system's ability to detect and quantify low activity levels.

2.6. Image quality analysis

Image quality analysis was undertaken to assess the following reconstruction parameters: Contrast (C_i, Signal-to-Noise Ratio (SNR_i), and Recovery Coefficient (RC). VOIs were delineated with diameters corresponding to the physical inner diameters of the four hot cylinders through the deep learning-based Jaszczak ACR Phantom segmentation. For each VOI size, 10 VOIs were designated to obtain background counts to ensure a consistent background noise metric across all four hot cylinders. Fig. 2 illustrates the VOIs for image quality analysis of a cylinder with 8 mm of diameter.

The SNR_i and C_i values for each hot cylinder and each energy window configuration were calculated. The SNR_i and C_i are defined as:

$$SNR_i = \left(\frac{P_i - B_i}{D_i}\right)$$

$$C_i = \frac{P_i}{B_i}$$

where D_i represents the standard deviation value within the background (BG) circle corresponding to the *ith* BG circle, B_i denotes the average count value within the BG circle associated with the *ith* cylinder and P_i is the average count value within the *ith* cylinder.

Furthermore, to quantify the equipment sensitivity and the contribution of each energy window, a syringe with 8 ml of volume filled with a solution of 8.14 kBq/ml of Ra-223 was used as a lesion simulator. The sensitivity is defined as

$$S_i = \left(\frac{Counts}{A^*t}\right),$$

where Counts is the total number of counts measured in a VOI within the radioactive volume; t is the acquisition duration (seconds); and A is the activity in the cylinder (kBq). According to the MIRD pamphlet 22 (Sgouros et al., 2010), the RC is used to study the accuracy of activity quantification in SPECT/CT images. The RC is defined as:

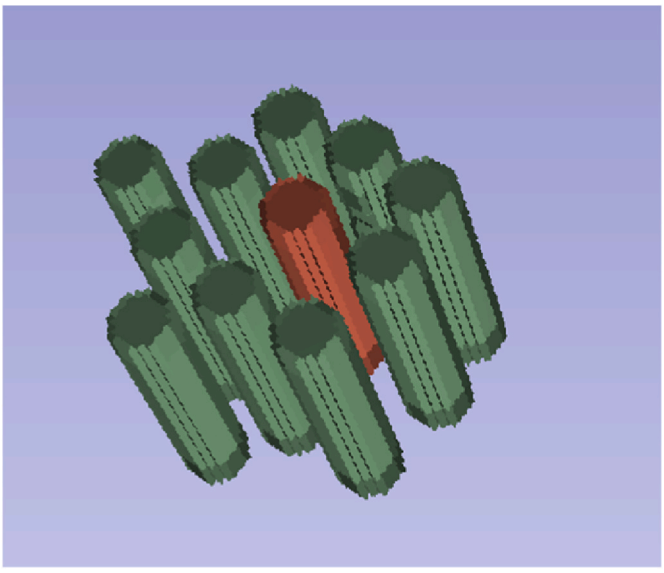

Fig. 2. Volume rendered in Slicer 3D for image quality analysis. Ten green cylinders for background counts and the red cylinder for targeted counts.

$$RC = \frac{A_{deti} - A_{det,BGi}}{A_{insi} - A_{ins,BGi}},$$

where A_{deti} is the identified activity within the *ith* cylinder, $A_{det,BGi}$ is the background activity detected within a cylinder of the *ith* volume, A_{insi} is the actual activity present inside the *ith* cylinder and $A_{ins,BGi}$ is the genuine background activity within a cylinder of the *ith* volume and a specified background activity concentration. For example, the activity estimated for the 89 keV energy window is:

$$A_{deti} = \left(\frac{C_{measured}}{S_{89keV}{}^*t}\right)$$

Because spatial resolution degrades as collimator-to-patient distance increases, the protocol distance from phantom image acquisition was the same as the patient image acquisition. The functionality of automated border contour was activated during the images acquisitions.

2.7. Dosimetry

The dosimetry study employed GATE (GEant4 Application for Emission Tomography) version 9.3 (Grevillot et al., 2011). The simulation used the Standard Option 4 physics list, which includes electromagnetic interactions tailored for medical physics applications (Grevillot et al., 2011). In line with established practices in Monte Carlo dosimetry, the QGSP_BIC_EMZ physics list was loaded to enable high-precision electromagnetic modeling, accounting for low-energy processes and multiple scattering—crucial for accurate alpha-particle transport (Grevillot et al., 2011).

The phantom geometry was defined using regular geometric volumes and its materials. This approach, which uses precise xyz-based positioning, provides a higher-fidelity and controlled representation of the phantom compared to the use of SPECT images, which are inherently limited by resolution and noise (Grevillot et al., 2011; Costa et al., 2017; Villoing et al., 2017).

Range-based production cuts were defined separately for each particle type and region. A 2 mm cut was set for gamma, electron, and positron particles within the ACR (phantom) region to ensure high

dosimetric accuracy, while a 10 mm cut was applied in the surrounding volume to reduce computational burden. These range-based cutoffs are internally converted to material-specific energy thresholds by Geant4 and provide a balance between precision and efficiency, as recommended in prior studies (Cordeiro et al., 2023).

Dose was calculated using a Dose Actor file, which was linked to the phantom volumes to generate three-dimensional voxelized dose maps. These maps served as the source for extracting dose values (Cordeiro et al., 2023). The simulated time was set to a 30-min acquisition period using the command 'setTimeStop 1800s'.

The simulation had approximately $5*10^8$ events. Dosimetric evaluations were based on activity data extracted from the 89 keV energy window, considering both scenarios with and without recovery correction (RC), and were compared with the ground truth dosimetry derived from the inserted activity in the cylinders.

3. Results

The Counts (Fig. 3a) indicate the number of detected counts according to cylinder diameter size. In Fig. 3b, background counts correspond to the average counts measured within each background VOI. TEW combines counts from three energy windows, so both counts and background counts would be higher compared to the other energy windows.

Across all four cylinder volumes, the 89 keV energy window exhibited higher count values and lower background counts than the 156 keV and 280 keV energy windows, impacting both Contrast and SNR performance. The contrast of the 89 keV energy window displays a significant difference for cylinder diameters exceeding 12 mm when compared to other energy windows, as illustrated in Fig. 3c.

In particular, the 280 keV energy window exhibited higher background count values compared to the 156 keV and 89 keV energy windows, leading to the lowest SNR values for all cylinders (see Fig. 3d). Conversely, the 89 keV energy window displayed superior performance, yielding the highest SNR compared to alternative energy windows.

To determine the RC, the detected activity - derived from the counts' rate measured in the SPECT image-was evaluated across cylinder volumes and energy windows, as illustrated in Fig. 4a. The RC serves as a quantification of the disparity between the detected and inserted activity within each cylinder.

The RC for the 26 mm cylinder ranges from 1.08 for the 89 keV window to 1.21 for the 280 keV window. These findings suggest that the activity of 26 mm lesions may be overestimated by up to 8 ± 1.4 %.

RC value less than one shows a partial volume effect, as observed in the 8 mm and 12 mm cylinders in Fig. 4b. Specifically, the RC for the 12 mm cylinder ranges from 0.56 for the 89 keV energy window to 0.47 for the 280 keV energy window, while the 8 mm diameter cylinder exhibits RC values from 0.28 for the 89 keV energy window to 0.19 for the 156 and 280 keV energy window. These observations imply that the activity of 12 and 8 mm lesions may be underestimated by up to 44 ± 8.9 % and 73 ± 6.8 %, respectively.

3.1. Deep learning-based Jaszczak ACR phantom segmentation

The deep learning model achieved reasonable performance after 500 epochs of training, reaching 10,500 iterations (Diaz-Pinto et al., 2022a). The best overall metric achieved during training was a Dice coefficient of 0.9717, which occurred at epoch 492, which is comparable to state-of-the-art models such as nnU-Net and UNETR when applied to similar volumetric datasets (Isensee et al., 2021; Hatamizadeh et al.,

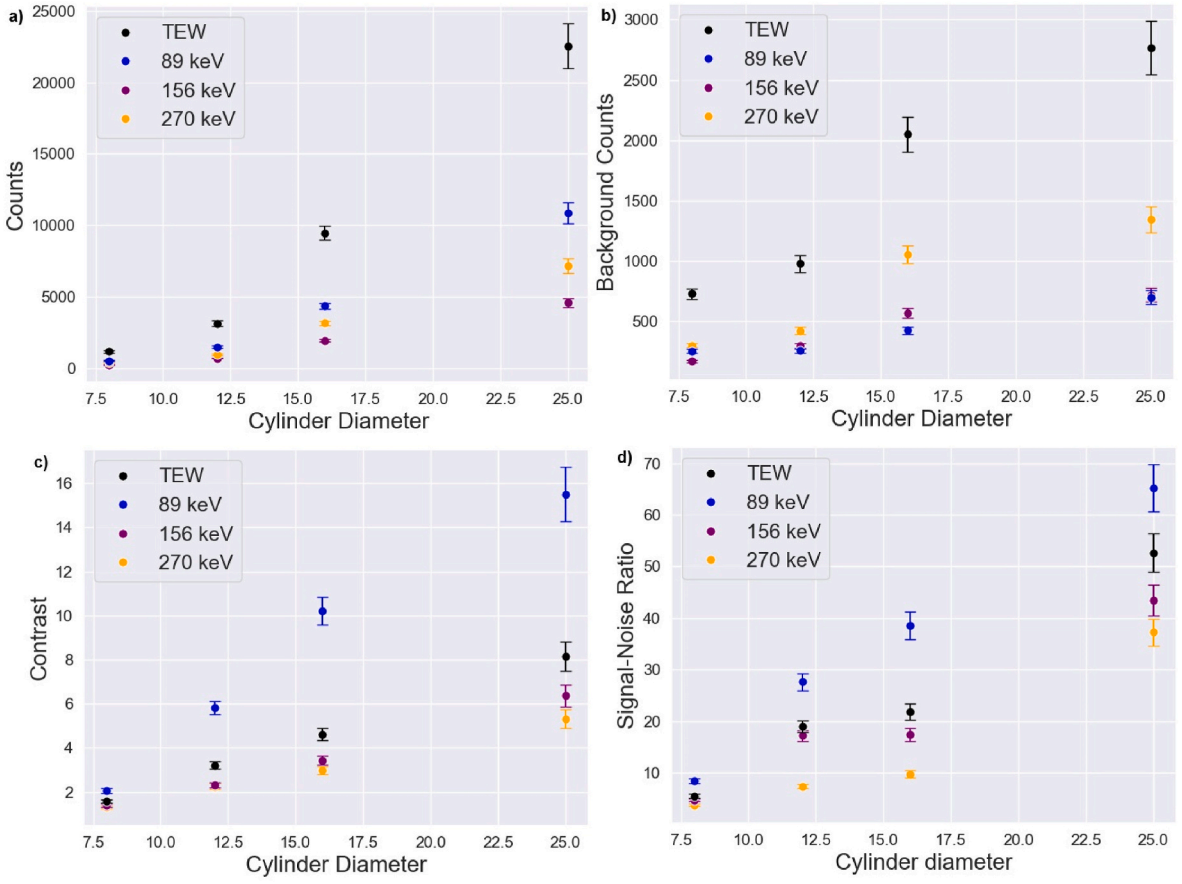

Fig. 3. Image quality parameters. a) Average counts inside each cylinder b) average counts inside background cylinders c) Contrast versus cylinder diameter size d) Signal-Noise-Ratio versus cylinder diameter size.

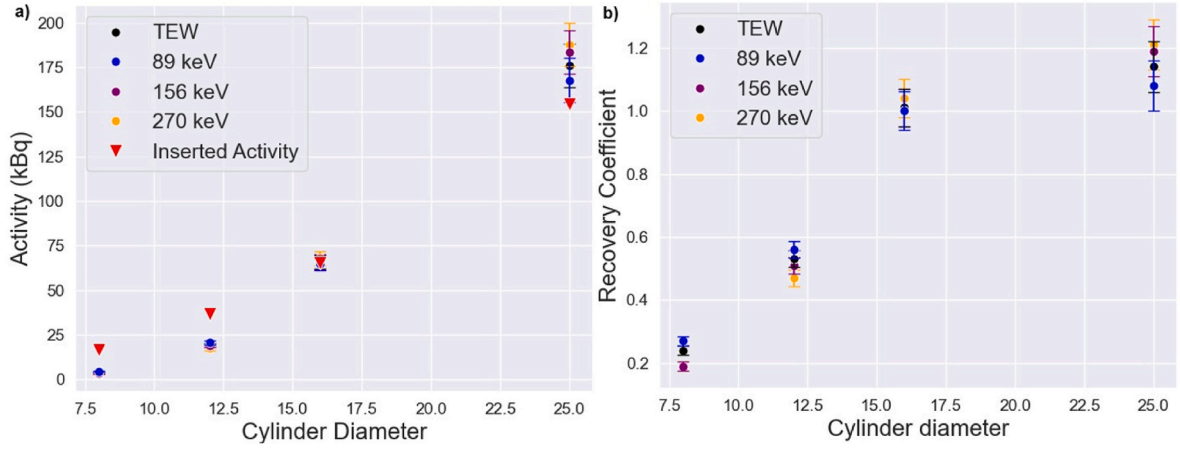

Fig. 4. Activities and Recovery Coefficient. a) Detected activities and inserted activity inside each cylinder b) Recovery coefficient of each energy window and cylinder size.

2022). The specific class-wise Dice coefficients for the training set, such as "cylinder 1", "cylinder 2", "cylinder 3" and "cylinder 4" ranged from 0.9210 to 0.9691, demonstrating high accuracy in segmenting these objects. In the evaluation phase, the model achieved a mean Dice coefficient of 0.8371, with class-wise Dice coefficients ranging from 0.7423 to 0.8933. The best overall evaluation metric was a Dice coefficient of 0.8754, occurring at epoch 483. These results highlight the model's robust performance in segmenting the cylinders in the given dataset, showing promising potential for practical applications (Isensee et al., 2021; Hatamizadeh et al., 2022). The automatic segmentation results are illustrated in Fig. 5.

3.2. Dosimetry

Concerning the dosimetry study, Table 1 presents the fluctuations and corrections in both detected activity and dosimetric parameters across cylinder sizes.

When computing absorbed doses for bone metastatic lesions in SPECT images, it is essential to correct activity detection by applying the suitable recovery coefficient. The results of Table 1 imply that lesions with a size of 26 mm might be overestimated by 8 %. In contrast, lesions measuring 12 mm and 8 mm could potentially be underestimated by 44 % and 73 %, respectively.

4. Discussion

Studies have employed a diverse array of Ra-223 concentrations, ranging from 0.05 μCi/ml (Owaki et al., 2017) to 0.75 μCi/ml (Owaki et al., 2017; Yue et al., 2015). In line with the established medical protocol (Xofigo, 2018), this study utilized a concentration of 0.26 μCi/ml, consistent with the approach adopted by Lima et al. (2022).

In this work, the MEGP collimator was used, as in previous studies (Benabdallah et al., 2019; Hindorf et al., 2012). However, Owaki et al. (2017) found the high-energy general-purpose HEGP collimator to outperform the MEGP in lesion recognition, especially when evaluating the number of lesions detected during SPECT examinations with Tc-99 m HMDP. Additionally, LEHR (Low Energy High Resolution) collimator has been proven able to visualize and quantify image quality parameters (Benabdallah et al., 2019).

The application of a Gaussian filter in the study has demonstrated enhancement in the consistency of activity concentration measurements (Gear et al.). However, the results of SNR indicate that the use of a Gaussian filter may lead to a decrease in image quality for energy windows of 154 keV and 280 keV, where low count statistics are observed. Additionally, Lima et al. (2022) reported that Gaussian filters resulted in relative differences when compared to the 89 keV window. Therefore, while a Gaussian filter with a FWHM of 4 mm can be employed for the 89 keV (+-25 %) energy protocol, it is not recommended for use in the triple window energy configuration.

The selection of energy windows for Ra-223 has been debated in the literature (Benabdallah et al., 2019; Owaki et al., 2017; Hindorf et al., 2012; Pacilio et al., 2016; Lima et al., 2022; Simões et al., 2018). While certain studies have employed 82 keV ± 20 % (Owaki et al., 2017; Hindorf et al., 2012; Pacilio et al., 2016), 84 keV ± 20 % (Lima et al., 2022; Yue et al., 2015) or 85.0 keV ± 20 % (Owaki et al., 2017), previous research (Benabdallah et al., 2019; Hindorf et al., 2012) has demonstrated the advantages of adopting the 89 keV photopeak with a 25 % width window. This energy window is aligned with photon emission probabilities exceeding 1 % (Benabdallah et al., 2019) and effectively avoids the characteristic X-rays from lead at 72 keV and 75 keV with probabilities of 28.7 % and 46.2 %, respectively (Hindorf et al., 2012).

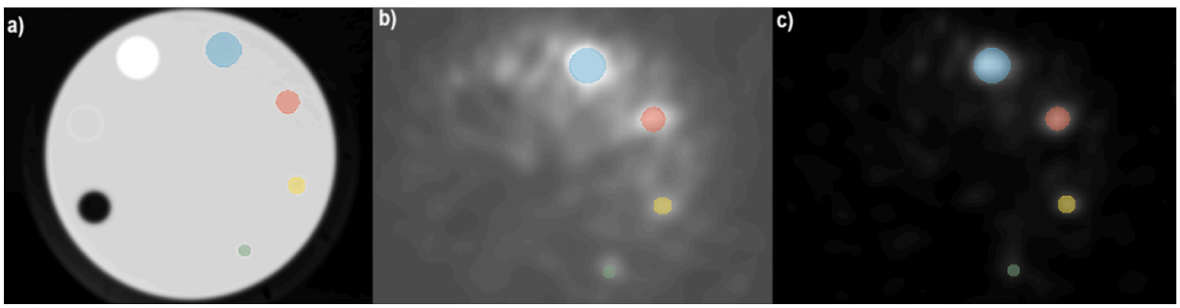

Fig. 5. Automatic segmentation by deep learning-based model. a) Generated segmentation in CT image. b) Applied segmentation in SPECT image of triple energy window. c) Applied segmentation in SPECT image of 89 keV energy window.

Table 1
Results of 89 keV energy window dosimetry.

Cylinder Size	$A_{det}(kBq)$	$A_{det}^{RC}(kBq)$	$D_{abv}(10^{-2}Gy)$	$D_{abv}^{RC}(10^{-2}Gy)$	$D_{True}(10^{-2}Gy)$
26 mm	165.54	152.64	5.30 (0.72)	4.88 (0.67)	4.95 (0.68)
16 mm	64.53	64.84	5.04 (0.69)	5.05 (0.70)	5.08 (0.69)
12 mm	20.01	36.79	2.77 (0.31)	4.95 (0.53)	5.07 (0.54)
8 mm	4.27	15.62	1.33 (0.14)	4.85 (0.50)	5.05 (0.53)

A_{det} is detected activity concentration, A_{det}^{RC} is activity concentration corrected with recovery coefficients, D_{abv} is absorbed dose of A_{det}, D_{abv}^{RC} is absorbed dose of A_{det}^{RC} and D_{True} is the true dose calculated from the actual activity inserted into the cylinders. Uncertainties' simulations are in parentheses.

At a concentration of 8.14 kBq/ml within an 8 ml cylinder, a sensitivity of 62.54 ± 5.3 counts per second per megabecquerel (cts/s/MBq) was observed. This result aligns with a prior study by Benabdallah et al. (2019), where, in the Ra-223 concentration range of 6.5 kBq/ml to 22.8 kBq/ml, sensitivities were reported in the range of 73.7 ± 6.2 to 43.4 ± 5.6 ct/s/MBq.

The 89 keV energy window exhibited significantly higher contrast compared to other windows (Fig. 2c), primarily due to its alignment with the main photopeak of Ra-223, which minimizes scatter contribution and enhances photoelectric interactions. This energy also showed the highest SNR across all cylinder sizes, reflecting an optimal balance between photon statistics and noise reduction. Overall, 89 keV provided the most favorable image quality metrics, driven by improved true event selection and reduced background contamination.

RC values for the 26 mm cylinder ranged from 1.08 to 1.21 across different energy windows, suggesting an overestimation of up to 8 % in the activity of these lesions. This aligns with other references (Collarino et al., 2018; Peters et al., 2020), which also observed an overshoot in RC values higher than 1, attributed to the resolution recovery algorithm used during reconstruction, emphasizing its role in influencing quantitative outcomes.

Partial volume effect is present for small lesions, specifically the 8 mm and 12 mm cylinders, with RC values less than 1.0, deprecating image quality parameters. In contrast, Benabdallah et al. (2019) discusses the RC for 5.6 ml spheres, with background activity, which were mainly overestimated. This discrepancy is attributed to the spill-in and spill-out effects, emphasizing the impact of background activity that, unlike our methodology, was not reduced from the VOIs of the spheres (Benabdallah et al., 2019).

In another study (Pacilio et al., 2016), an average lesion size of 122 ml (across 24 lesions ranging from 9 to 378 ml in 9 patients) was reported for osteoblastic bone metastasis of prostate cancer. Accordingly, this protocol will facilitate the quantification of osteoblastic bone metastasis with clinical uptake. However, it is acknowledged that implementing partial volume correction through RC will be imperative for achieving a more reliable quantification.

A thorough 20-year literature review shows that error rates for clinically significant or major errors in radiology usually range from 2 % to 20 %, depending on the specific radiological study (Goddard et al., 2001). Implementing systematic improvements, such as automatic segmentation, has the potential to significantly reduce these error rates.

Manual segmentation of Jaszczak ACR Phantom can take 15–31 min for experienced and non-experienced clinical professionals, respectively. The approach of this work reduces the process to less than 2.5 min–31 s for segmentation and an additional 2 min for inspection and correction. The time improvement is six times for experienced professionals and twelve times for non-experienced ones.

The deep learning-based Jaszczak ACR Phantom segmentation performance, in the evaluation phase, surpassed 91,23 % in class-wise Dice coefficients. The ability for clinical experts to make real-time adjustments to the segmentation ensures that the model's output aligns with specific clinical needs and purposes. This dynamic interaction enhances the practicality and adaptability of the segmentation tool in a clinical setting, where customization and fine-tuning are often necessary for optimal results.

In addition to its high performance in ACR Phantom segmentation, the potential applicability of the model to various medical imaging modalities such as SPECT (Owaki et al., 2017), PET (Silosky et al., 2019), and MRI (Palesi et al., 2022) highlights its role as a valuable tool for medical image analysis tasks. The interactive mode allows clinical professionals to adapt the segmentation to specific purposes. Moreover, the ability for clinical experts to provide corrections to automatic segmentation and then retrain the model with these new annotations fosters a collaborative approach to refining and enhancing the model's performance.

Concerning the dosimetry study, the 26 mm cylinder size indicates an 8,45 % difference between A_{det} and A_{det}^{RC}, as well as between D_{abv} and D_{abv}^{RC}, emphasizing the impact of the reconstruction technique on the measured activity. The associated variations in dose values further underscore the importance of carefully considering the impact of reconstruction techniques when interpreting results in the context of image quantification.

For smaller cylinder sizes, the percentage difference between D_{abv} and D_{abv}^{RC} became more pronounced due to the partial volume effect. Overall, these results highlight the importance of considering and applying recovery correction, especially in scenarios involving smaller structures, to enhance the accuracy of both activity concentration and absorbed dose estimations.

Future research should prioritize advancing techniques for mitigating partial volume effects, particularly in the context of smaller lesions. Additionally, new strategies to reduce error and improve workflow efficiency should be explored. The continued focus on systematic improvements, including the integration of automated segmentation and artificial intelligence, promises to significantly reduce error rates in radiological studies. This dual emphasis on partial volume effect mitigation and workflow efficiency will elevate the reliability of nuclear medicine studies, contributing to more accurate and standardized clinical applications.

5. Conclusion

This study presents a deep learning-based segmentation model to enhance the precision and efficiency of ACR phantom segmentation in SPECT imaging.

These findings underscore the role of recovery correction in the accuracy of activity concentration and absorbed dose estimations. This investigation brought significant refinement of dosimetry protocols specifically pertinent to Ra-223. Moreover, the practical implications derived from this study extend beyond the realm of Ra-223 dosimetry, offering valuable avenues for enhancing the precision of quantitative outcomes, particularly in scenarios characterized by structural diversity such as metastasis environments.

CRediT authorship contribution statement

Cristian Felipe Griebler: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation,

Conceptualization. **Leanderson Pereira Cordeiro:** Software, Methodology, Data curation, Formal analysis. **Luis Felipe Lima:** Validation, Methodology, Data curation. **Vagner Bolzan:** Methodology, Data curation, Software. **Vitor Dutra:** Software, Methodology, Data curation. **Lidia Vasconcellos De Sá:** Writing – review & editing, Supervision, Funding acquisition. **Daniel Alexandre Baptista Bonifacio:** Supervision, Funding acquisition, Writing – review & editing.

Funding

The study was funded by Coordination for the Improvement of Higher Education Personnel (CAPES).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

All authors declare that they have no known conflicts of interest in terms of competing financial interests or personal relationships that could have an influence or are relevant to the work reported in this paper. We are grateful to the Bionuclear Clinic who allowed us to use their SPECT/CT system and to the local staff for their helpful assistance. Also, we thank them for providing modifications to their anthropomorphic phantom ACR in order to be more suitable to our study.

Data availability

Data will be made available on request.

References

Benabdallah, N., Bernardini, M., Bianciardi, M., et al., 2019. 223Ra-dichloride therapy: optimization of SPECT images. EJNMMI Res. 9 (1), 20. https://doi.org/10.1186/s13550-019-0488-7.

Collarino, A., Pereira Arias-Bouda, L.M., et al., 2018. Validation of absolute SPECT/CT quantification. Med. Phys. 45 (5), 2143–2153.

Cordeiro, L.P., de Sá, L.V., Kitamikado, R.A., Sapienza, M.T., Da, B.B., 2023. Optimized Monte Carlo simulations for voxel-based internal dosimetry. Phys. Med. Biol. 68. https://doi.org/10.1088/1361-6560/acd2a1.

Costa, G.C.A., Bonifácio, D.A.B., Sarrut, D., et al., 2017. Optimization of GATE simulations for whole-body planar scintigraphic acquisitions using the XCAT male phantom with 177Lu-DOTATATE biokinetics in a Siemens Symbia T2. Phys. Med. 42, 292–297.

Council, E.U., 2014. European Council directive 2013/59/euratom on basic safety standards. Off. J. Eur. Union L13, 1–73.

Danieli, R., Milano, A., Gallo, S., Veronese, I., Lascialfari, A., Indovina, L., Botta, F., Ferrari, M., Cicchetti, A., Raspanti, D., Cremonesi, M., 2022. Personalized dosimetry in targeted radiation therapy: a look to methods, tools and critical aspects. J. Personalized Med. 12 (2), 205. https://doi.org/10.3390/jpm12020205. PMID: 35207693; PMCID: PMC8874397.

Deshayes, E., Roumiguie, M., Thibault, C., Beuzeboc, P., Cachin, F., Hennequin, C., et al., 2017. Radium 223 dichloride for prostate cancer treatment. Drug Des. Dev. Ther. 11, 2643–2651. https://doi.org/10.2147/DDDT.S122417.

Diaz-Pinto, A., et al., 2022. DeepEdit: deep editable learning for interactive segmentation. In: DALI 2022. Springer, Cham. https://doi.org/10.1007/978-3-031-17027-0_2.

Diaz-Pinto, A., Alle, S., Nath, V., Tang, Y., Ihsani, A., Asad, M., Perez-Garcia, F., Mehta, P., Diaz-PintoLi, W., Flores, M., et al., 2024. MONAI Label: A framework for AI-assisted interactive labeling of 3D medical images. Med. Image Anal. 95, 103207.

Dickson, J.C., Armstrong, I.S., Gabiña, P.M., et al., 2023. EANM practice guideline for quantitative SPECT-CT. Eur. J. Nucl. Med. Mol. Imag. 50, 980–995. https://doi.org/10.1007/s00259-022-06028-9.

Fedorov, A., Beichel, R., Kalpathy-Cramer, J., et al., 2012. 3D Slicer as an image computing platform. Magn. Reson. Imaging 30 (9), 1323–1341. https://doi.org/10.1016/j.mri.2012.05.001.

Gear JI, Cox MG, Gustafsson J, Gleisner KS, Murray I, Glatting G, Konijnenberg M, Flux GD. EANM Practical Guidance.

Goceri, E., 2023. Medical image data augmentation: techniques, comparisons and interpretations. Artif. Intell. Rev. 56, 12561–12605. https://doi.org/10.1007/s10462-023-10453-z.

Goddard, P., Leslie, A., Jones, A., et al., 2001. Error in radiology. Br. J. Radiol. 74, 949–951. https://doi.org/10.1259/bjr.74.886.740949.

Grevillot, L., Bertrand, D., Dessy, F., et al., 2011. Monte Carlo pencil beam scanning model. Phys. Med. Biol. 56, 5203–5219.

Hatamizadeh, A., Tang, Y., Nath, V., et al., 2022. UNETR: transformers for 3D medical image segmentation. WACV 574–584.

Higano, C.S., George, D.J., Shore, N.D., Sartor, O., Miller, K., Conti, P.S., Sternberg, C.N., Saad, F., Sade, J.P., Bellmunt, J., Smith, M.R., Chandrawansa, K., Sandström, P., Verholen, F., Tombal, B., 2023. Clinical outcomes and treatment patterns in REASSURE: planned interim analysis of a real-world observational study of radium-223 in metastatic castration-resistant prostate cancer. eClinicalMedicine 60, 101993. https://doi.org/10.1016/j.eclinm.2023.101993. PMID: 37251627; PMCID: PMC10209672.

Hindorf, C., Chittenden, S., Aksnes, A.K., et al., 2012. Quantitative imaging of 223Ra-chloride. Nucl. Med. Commun. 33 (7), 726–732. https://doi.org/10.1097/MNM.0b013e328353bb6e.

Isensee, F., Jaeger, P.F., Kohl, S.A.A., et al., 2021. nnU-Net: a self-configuring method. Nat. Methods 18, 203–211. https://doi.org/10.1038/s41592-020-01008-z.

Jiadong, Z., et al., 2023. Recent advancements in artificial intelligence for breast cancer. Semin. Cancer Biol. https://doi.org/10.1016/j.semcancer.2023.09.001.

Lima, L.F.C., Pinto, G.M., da Silva, C.C.O., et al., 2022. Optimal theranostic SPECT imaging protocol. J. Med. Imag. Radiat. Sci. 53 (3), 374–383. https://doi.org/10.1016/j.jmir.2022.06.009.

Morris, M.J., Corey, E., Guise, T.A., et al., 2019. Radium-223 mechanism of action: implications for use in treatment combinations. Nat. Rev. Urol. 16, 745–756. https://doi.org/10.1038/s41585-019-0251-x.

Owaki, Y., Nakahara, T., Kosaka, T., et al., 2017. Ra-223 SPECT: phantom study and initial experience. EJNMMI Res. 7, 81.

Pacilio, M., Ventroni, G., De Vincentis, G., et al., 2016. Dosimetry of bone metastases in 223Ra-dichloride therapy. Eur. J. Nucl. Med. Mol. Imag. 43, 21–33. https://doi.org/10.1007/s00259-015-3150-2.

Palesi, F., Nigri, A., Gianeri, R., et al., 2022. MRI data quality assessment using ACR phantoms. Phys. Med. 104, 93–100.

Parker, C., Nilsson, S., Heinrich, D., Helle, S.I., O'Sullivan, J.M., Fosså, S.D., et al., 2013. Alpha emitter radium-223 and survival in metastatic prostate cancer. N. Engl. J. Med. https://doi.org/10.1056/NEJMoa1213755.

Peters, S.M.B., Meyer Viol, S.L., van der Werf, N.R., et al., 2020. Variability in lutetium-177 SPECT quantification. EJNMMI Phys. 7, 9. https://doi.org/10.1186/s40658-020-0278-3.

Sandström, M., Freedman, N., Fröss-Baron, K., Kahn, T., Sundin, A., 2020. Kidney dosimetry in 777 patients during 177Lu-dotatate therapy: aspects on extrapolations and measurement time points. EJNMMI Phys. 7, 1–15.

Sgouros, G., Kolbert, K.S., Sheikh, A., Pentlow, K.S., Mun, E.F., Barth, A., Robbins, R.J., Larson, S.M., 2004. Patient-specific dosimetry for 131 I-thyroid cancer therapy using 124I PET and 3D-ID software. J. Nucl. Med. 45, 1366–1372.

Sgouros, G., Roeske, J.C., McDevitt, M.R., et al., 2010. MIRD pamphlet No. 22 (abridged). J. Nucl. Med. 51, 311–328. https://doi.org/10.2967/jnumed.108.058651.

Silosky, M., Karki, R., Chin, B.B., 2019. 68Ga and 18F quantifcation, and detectability of hot spots using an ACR Phantom: contributions of radionuclide physical diferences to hot spot detectability. J. Nucl. Med. 60 (Suppl 1), 1200.

Simões, R.F.P., Da Silva, C.J., Da Silva, R.L., et al., 2018. Deconvolution method to split up X-ray peaks emitted by 223Ra. Braz. J. Radiat. Sci. https://doi.org/10.15392/bjrs.v6i1.353.

Villoing, D., Marcatili, S., Garcia, M.-P., Bardiès, M., 2017. Internal dosimetry with GATE: validation. Phys. Med. Biol. 62, 1885–1904.

Xofigo, 2018. Cloreto de Rádio (223 Ra). Bayer Health Pharm Inc., Anvisa.

Yin, X.X., Sun, L., Fu, Y., et al., 2022. U-Net-Based medical image segmentation. J. Healthc. Eng. https://doi.org/10.1155/2022/4189781.

Yue, J., Frey, E.C., Mauxion, T., et al., 2015. Potential for quantitative imaging of Ra-223. Int. J. Radiat. Oncol. Biol. Phys. 93, E237.