Deep Learning-enhanced Voxel-Based Personalized Dosimetry in ^{177}Lu-DOTATE Radiopharmaceutical Therapy: Implications for Radiation Protection

Daniel Juarez Luna

May 8, 2025

Abstract

This paper reviews and assesses a transformer-based deep learning model developed by Mansouri et al. (2024) for personalized voxel-level dosimetry in [^{177}Lu]Lu-DOTATATE radiopharmaceutical therapy, with a particular focus on radiation protection implications. Although organ-averaged dose estimates are provided by conventional MIRD-based dosimetry techniques, they are computationally constrained and do not account for intra-organ heterogeneity. In order to increase dosimetric precision while lowering processing requirements, Mansouri's hybrid technique combines deep learning with Monte Carlo-based reference data. We reviewed the hybrid transformer-based model proposed by Mansouri et al., highlighting its methodological innovations, such as their comparison to Single and Multiple Voxel S-value approaches to deep learning based workflow and clinical potential. Their findings showed quicker dose map creation, and more accurate dosage calculation for vital organs and lesions. We also examine the effects of such processes on treatment uniformity, occupational exposure, and the developing role of AI in clinical decision-making. This evaluation affirms the potential of deep learning models to support ALARA compliance and treatment optimization in modern nuclear medicine.

1 Introduction

Soon after the initial discovery of radioactivity by Henri Becquerel in 1896 and of radium by Marie and Pierre Curie in 1898, radioactivity was used for medical treatment. The initial interest in Radiopharmaceutical therapy (RPT) arose from the exciting and novel opportunities offered by radioactivity for addressing diseases, even in the absence of a full understanding of its biological effects.

In modern days, RPT is considered a targeted cancer treatment meant to bind specifically with a tumor receptor it gives radionuclides the ability to "label" elements and be tracked in the body [1]. In general, a radioactive isotope is linked to a pharmaceutical compound that serves as the transport and bridge towards a particular set of cells. These compounds are usually administered

intravenously; upon administration, the radiopharmaceutical circulates around the body and potentially binds with cancerous tissue, then delivers the cytotoxic radiation directly [2]. In simple terms, a radioactive source plus a chemical part can target specific cancerous cells and get rid of them.

This method minimizes the exposure to healthy tissue. The idea behind it is solid and makes this methodology particularly well suited for treating metastatic and disseminated cancers. It showed early success with radio-phosphorous and the proposed use of radio-iodine and offers an advantage over external beam radiotherapy [1].

Since the early 2000s, lutetium-177 (^{177}Lu), a β^- emitting radionuclide with perfect qualities for targeted radiotherapy and imaging, has been one of the most often employed isotopes in RPT. Its half-life is 6.7 days, and it releases low-energy gamma rays that enable SPECT imaging after treatment, as well as medium-energy β^- particles that are appropriate for small-to-medium tumor lesions [1, 2] Figure 1 shows the principle of linking chemically a radioactive isotope to a targeting vector and how it is used on therapies such as [^{177}Lu]Lu-DOTATATE for neuroendocrine tumors by administrating it to the patient then recompiling the relevant information with medical imaging modalities. Source: *Frontiers in Nuclear Medicine* (2024)

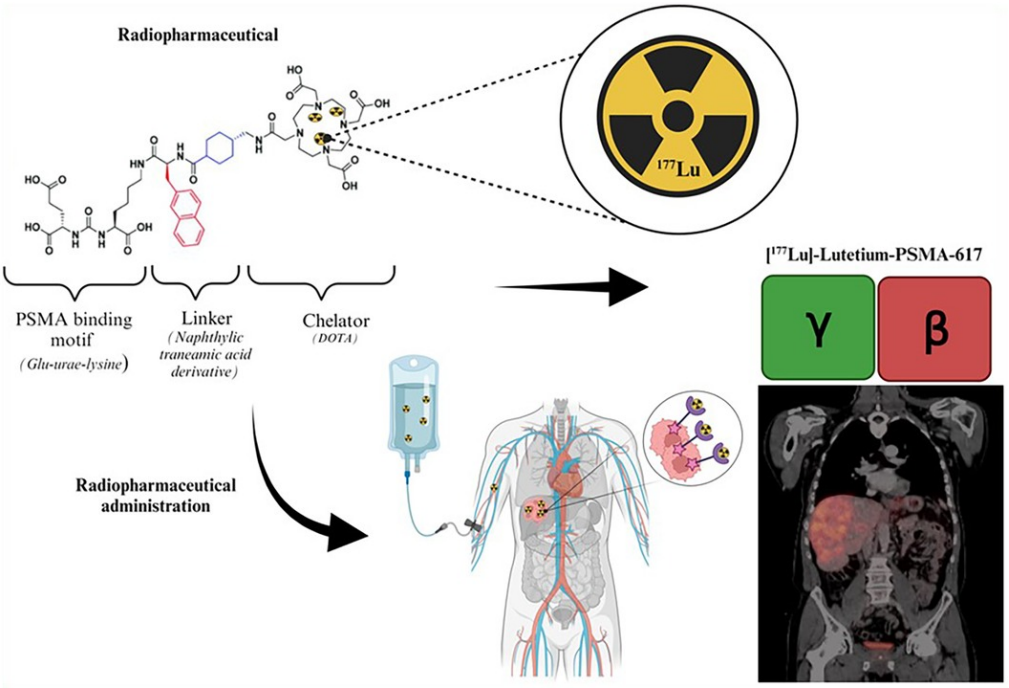

Figure 1: Illustration of a theranostic radiopharmaceutical approach using [^{177}Lu]Lutetium-PSMA. [3].

Traditional methods to quantify internal dose, such as the Medical Internal Radiation Dose (MIRD) schema, estimate average doses to entire organs using tabulated organ-level S-values derived from standard phantoms. These methods form the basis of modern clinical dosimetry in nuclear medicine and have been widely used for several decades now [1]. Yet it lacks the capacity to

capture intra-organ heterogeneity in radiopharmaceutical uptake, nor can they provide spatially resolved dose distributions [4]. As a result, treatment plans based solely on organ-level metrics may underdose tumors or exceed toxicity thresholds in organs at risk.

Voxel-based personalized dosimetry has emerged as an alternative. Using Monte Carlo dose calculation model and quantitative imaging (such as SPECT/CT), this method estimates dose deposition at the voxel level, which are small 3D units that represent distinct patient volumes. However, its extensive computational requirements make it impractical for routine clinical use. Deep learning (DL) has been successfully employed for different computational medical imaging tasks, particularly convolutional neural networks (CNNs) like U-Net. While th abovementioned studies benefited from utilizing CNNs, the application of transformer architectures remains limited in dosimetry tasks. Transformers benefit from self-attention and can process the input data in parallel, which leads to increased efficacy. In addition, they can achieve adequate performance with very limited training data. [5]

This paper builds upon the work of Mansouri et al. (2024), which developed and validated a hybrid deep transformer-based model for improving voxel-level dosimetry in [^{177}Lu]Lu-DOTATATE radiopharmaceutical therapy. It evaluates the clinical and radiation protection implications of replacing conventional MIRD formalisms with deep learning approaches.

Specifically, we ask, *How can transformer-based deep learning models improve voxel-level dosimetry accuracy in [^{177}Lu]Lu-DOTATATE therapy compared to conventional MIRD methods, and what are the implications for radiation protection?*

We aim to clarify the dosimetry and radiation protection consequences of using voxel-based DL enhanced dosimetry to streamline therapeutic workloads in modern nuclear medicine.

2 Dosimetry Methods

Internal dosimetry presents a challenge that external beam therapy does not face, dosimetrist need to calculate absorbed dose from the activity of administered radiopharmaceutical.

In order to ensure safety and efficacy in RPT for patients, we need to stimate the absorbed dose to both normal tissues and tumors in the most accurate way [5]. The methodology includes, establishing optimal absorbed dose tolerance levels for various tissues and comparing dose-response patterns between different radiopharmaceuticals and patients. [2, 3]. This is done by quantifying the radiation dose delivered and adjusting the administered activity to maximize the dose-to-tumor while minimizing toxicity to organs at risk (OAR).

The way to measure these parameters comes from nuclear imaging modalities, either PET or SPECT. Usually combined with CT imaging systems to obtain the functional and anatomical information from both, and with that know the localization of radiotracer uptake in the body. For dosimetry, it allows us to correlate physiological activity with anatomical detail that supports accurate quantification.

It is important to highlight that the CT component enables attenuation correction for later used on the reconstructed SPECT images. Second, it allows for scatter correction using dual-energy window methods that reduces the impact of Compton-scattered photons that would otherwise degrade image fidelity. Although this integration of CT into SPECT/CT imaging can appear to be an increase in the overall dose due to the external x-ray beam, it is often regulated and justified as it serves to improve image quality and dosimetric accuracy. This additional exposure aligns with the ALARA (As Low As Reasonably Achievable) principle, as it enables better attenuation and scatter correction.

This information not only anticipates the biologic effects of radiation but is now used to personalize and optimize therapy for individual patients. So, we can prevent under- or over-dosing associated with the use of fixed administered activity regimens. To promote this goals, standards that ensure radiation exposure for therapeutic purposes are pushed to individually plan and verify minimum doses to non-target volumes and tissues [3].

In addition to improving visualization, the combination of functional imaging (PET/SPECT) and anatomical imaging (CT) directly contributes to the quantitative framework needed for dosimetry. It correlates voxel data with patient anatomy and facilitate the creation of customized dosage maps.

The traditional approach used in RPT to calculate the absorbed radiation dose to different organs is know as Medical Internal Radiation Dose (MIRD) formalism. To compute absorbed dose using MIRD, we must first quantify the time-dependent behavior of the radiopharmaceutical within source and target organs [6].

From SPECT/CT region of interest (ROI) corrected for attenuation, scatter, and partial volume effects, we obtain the Time-Activity Curve (TAC). Then the area under TAC, known as the Time-Integrated Activity (TIA) or $\tilde{A}$, representing the cumulative decay events in each region over the imaging period, is computed and used with the S-values to get the absorbed dose to target organ.

These S-value are tablulated and represent the absorbed dose to a target region per unit time-integrated activity in a source for various radiounucleids and it is a standarized reference for anatomical models by age and gender.

In figure 2 we see an example of a planar count rate and SPECT-derived activity being used to reconstruct the time-activity curve and calculate the time-integrated activity (TIA). [2].

TIA is typically calculated by fitting a mono-exponential decay model to later time points of the TAC (e.g., $t > 24\,\mathrm{h}$) and applying trapezoidal integration to earlier time points [2]. Mathematically, TIA is expressed as:

$$\tilde{A} = \int_0^{\infty} A(t)dt \approx \int_0^{t^*} A(t)dt + \frac{A(t^*)}{\lambda} \tag{1}$$

where $A(t)$ is activity at time t, t^* is the final measurement time, and λ is the clearance rate from the fitted exponential. This integral gives the cumulated activity and serves as input to the

MIRD formalism when multiplied by the appropriate S-value [**?**].

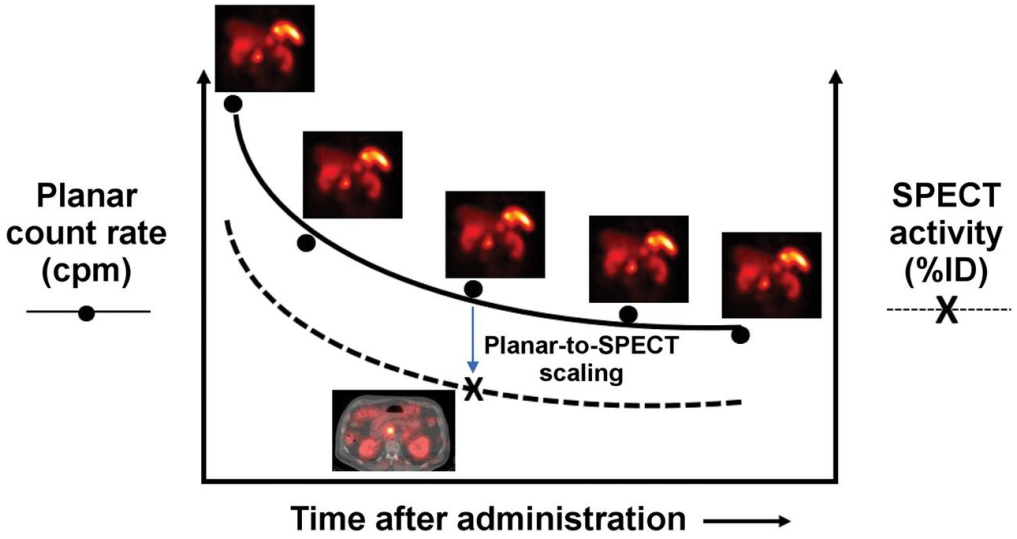

Figure 2: Hybrid planar/SPECT imaging method for acquiring time–activity curves [2].

It is a robust technique that follows the one size fits all kind of phylosophy. While this allowed us to get useful ratios tabulated it has some limitations. Medical physicists used phantoms that represent average patient anatomy and physiology but lack internal anatomical differences in organ size, shape and position that are present in patients. The technique is also based on some assuptions that simplify the work, for example it assumes uniform activity distribution within each organ. Figure 3 represents the interaction of the organs in terms on the potential activity transfer or dose. [4].

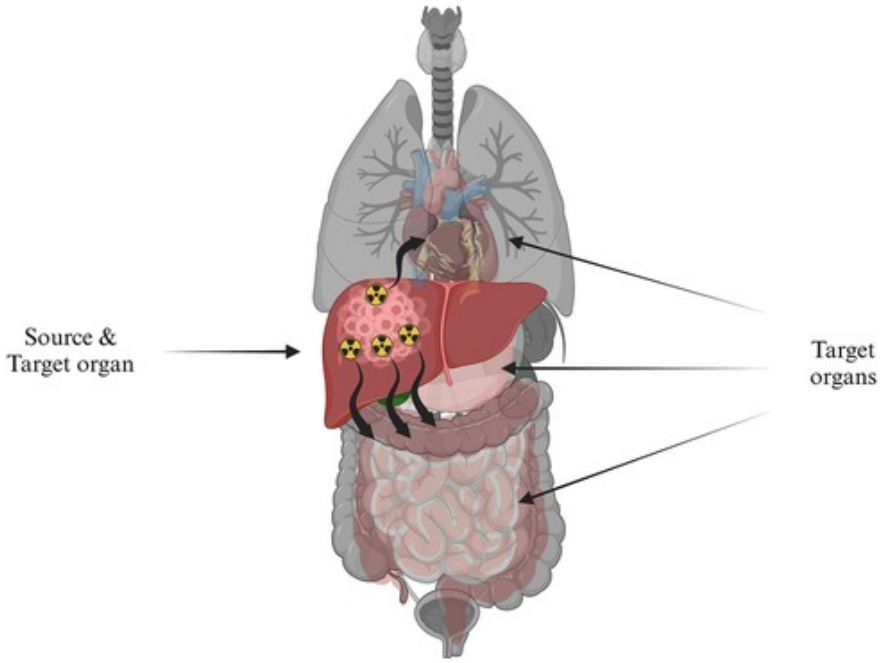

Figure 3: Schematic representation of radiation emission from a source organ (liver) affecting nearby target organs, highlighting absorbed dose distribution. [3].

While in reality, the uptake of radiopharmaceuticals can be highly heterogeneous within organs and tumors [2, 7]. It also lacks the ability to identify activity variations inside organs, as the absorbed dose is averaged across the entire organ, suborgnan or subtumor level information is lost. Furthermore, it might hinder the establishment of accurate dose-effect relationships for tumors and normal tissues, as it does not account for dose heterogeneity [4].This means greater accuracy was needed in the estimation of how each organ handles the radiopharmaceutical.

A more detailed depiction of radiopharmaceutical dispersion and energy deposition was made possible by voxel-based techniques, which were developed in response to the spatial and anatomical constraints of organ-level models. Instead of averaging for each organ we take the activity distribution from SPECT or PET imaging and encapusalte its information in voxels, a volume element that represent a specific location in 3D within a patient, with imaging modalities like CT or MRI we are able to find the spatial distribution. [2, 8]. In the case of PET/SPECT these values come in units of activity, reflecting concentration of the radiopharmaceutical in a small volume of tissue.

Two main methods are commonly used, Single Voxel S-value (SSV) and Multiple Voxel S-value (MSV). SSV method applies a homogeneous water-equivalent kernel to estimate absorbed dose, simplifying the physical modeling at the expense of accuracy in heterogeneous tissues. MSV method overcomes this by using precomputed kernels that reflect differences in density and composition among tissues. These kernels, are generated using Monte Carlo simulations in anatomically realistic voxelized phantoms.

Voxel-Based Dosimetry helps overcome some of the previously mentioned limitations as it can model non-uniform distribution and can be adapted for each patient. However it also adds some challenges, for example, computational burden for complicated dose calculations with the use of Monte Carlo, image and activity registration issues from SPECT or PET and tumor delimitation accuracy.

Studies have shown that voxel-based methods, particularly Monte Carlo (MC) and Multiple Voxel S-value (MSV), provide more accurate dose estimates compared to organ-level MIRD, especially in regions with tissue inhomogeneities like the lung-liver interface or bone-soft tissue boundaries [3, 5, 8]. This reflects the modern push, toward personalized dose estimation based on patient-specific anatomy rather than averaged models.

Researchers contribute to the refinement of dosimetric methodologies and the investigation of combined-modality therapies. The incrementally better computation power now allows to further develop methodologies with the use of Deep learning and neural networks.

Despite their accuracy, voxel-level approaches require a lot of work and effort. A single-bed SPECT scan using MC simulations, the gold standard for voxel-level dosimetry, may take up to 48 hours, while manual organ and tumor segmentation can take several hours. Even MSV techniques, which require less computing power than MC, can take more than an hour on a typical CPU. [3, 5].

Deep learning (DL) has been used by researchers to overcome these limitations. It has been demonstrated that the time needed for organ contouring may be significantly decreased by using

automatic segmentation tools that employ DL algorithms, such as TotalSegmentator and Moose [7]. Furthermore, compared to standard MC simulations, DL-based dose map prediction models may provide voxel-level dose distributions in less than a minute, which is orders of magnitude quicker [8].

In addition to increased efficiency, the incorporation of DL into the voxel-based dosimetry process holds promise for real-time treatment planning and increased clinical accessibility. Mansouri et al. (2024) blend the learning capabilities of transformer networks with the physical roots of voxel S-value approaches, to investigate how this technologies will impact dosimetry planning in the near future.

3 Hybrid Deep Learning Model for Voxel Dosimetry

Building on the conceptual foundation and motivation for integrating deep learning into voxel-based dosimetry, the following section summarizes the methodology employed by Mansouri et al. (2024) to develop and validate a hybrid transformer-based model. This includes the clinical imaging workflow, the construction of dose maps through MIRD-based techniques, and the training of a deep learning model using Monte Carlo simulations as the reference standard. Their goal was to improve the speed and accuracy of dose map generation, moving towards personalized treatment planning in routine clinical settings.

The primary radionucleus they utilized was lutetium-177 (^{177}Lu), specifically [^{177}Lu]Lu-DOTA-TATE also known as ^{177}Lu-Oxodotreotide or [^{177}Lu]Lu-DOTA-(Tyr3)-octreotate. They attach the somatostatin affine peptide (Tyr3)-octreotate to ^{177}Lu using a somatostatin analogue peptide linked with the bifunctional chelator DOTA. [9]

With a halflife of 6.7 days (162 hours), the radioisotope ^{177}Lu decays to hafnium 177 (^{177}Hf), which releases β^- particles with maximum energies of 497 keV (78.6%), 384 keV (9.1%), and 176 keV (12.2%). The electrons' maximum (mean) soft-tissue penetration is 1.7 mm (0.23 mm). As a result, it is categorized as an emitter of short-range β particles. Low-to-medium energy γ-rays with energies of 113 keV (6.6%) and 208 keV (11%) accompany the β decay. (as reported in [9], citing data from Dash et al. and Zalutsky)

The retrospective study utilized SPECT/CT images from 50 sessions of 22 patients with neuroendocrine tumors (NETs) undergoing up to four cycles of [^{177}Lu]Lu-DOTATATE therapy. The injected activity was personalized based on kidney function, body habitus, and prior dosimetry results.

Imaging was conducted at approximately 4, 24, 69, and 120 hours post-injection using a Siemens Symbia T16 SPECT/CT camera. The SPECT images were reconstructed with a 3D-OSEM algorithm, an iterative image reconstruction algorithm that refines the image over multiple cycles to improve quality, and included CT-based attenuation correction and dual-energy window scatter correction. Futhermore, activity calibration was validated using a cylindrical phantom.

That is, they took multiple images spaced in time after the patient was injected with [^{177}Lu]Lu-DOTATATE to capture how the radiopharmaceutical distributes and clears from different tissues over time and provide the TIA and TACs for dose calculation.

The Multiple Voxel S-value (MSV) method was employed to generate dose maps using a MIRD-based framework, relying on multiple pre-calculated kernels to reflect tissue heterogeneities. This approach mitigates the limitations of Single Voxel S-value (SSV) method, which assumes uniform water-equivalent tissue. As discussed before, the limitations of generalized anatomical models, where the use of pre-tabulated S-values, ignores anatomical variation. Especially in the context of organ shape, density, and adjacent tissues.

Time-activity curves were created using the multi-timepoint SPECT data in order to produce the Time-Integrated Activity (TIA) maps. A modified UNET deep learning model (RESUNET) was used to segment OARs anatomically before curve fitting, and the results were manually corrected. An expert physician manually identified the tumor volumes on SPECT. Based on a hybrid curve fitting method, TACs were built voxel-wise.

Radiotracer uptake was modeled using trapezoidal interpolation for early time points ($t < 24$ h). Later time points ($t > 24$ h) were fitted with a mono-exponential decay model:

$$A(t) = C \cdot e^{-\lambda t} ... \tag{2}$$

where C is the scaling constant and λ the biological clearance rate. TIA was computed as the area under each fitted TAC. To minimize noise, the effective half-life for a given volume of interest was refined by taking the mean λ across voxels [5].

This is consistent with pharmacokinetic modeling discussed in class, where radioactive decay and biological clearance often follow first-order kinetics, particularly for radiopharmaceuticals with predictable biodistribution and elimination profiles.

In figure 4 the workflow followed by the authors is shown. It clearly depicts the extraction of SPECT/CT images accross time, with the density map for corrections. Organs at risk (OARs) are segmented using deep learning, and then areas of interest (ROIs) are registered and delineated. The time-integrated activity (TIA) for voxel-level dose estimates is then calculated by fitting time-activity curves (TACs) using mono-exponential models.

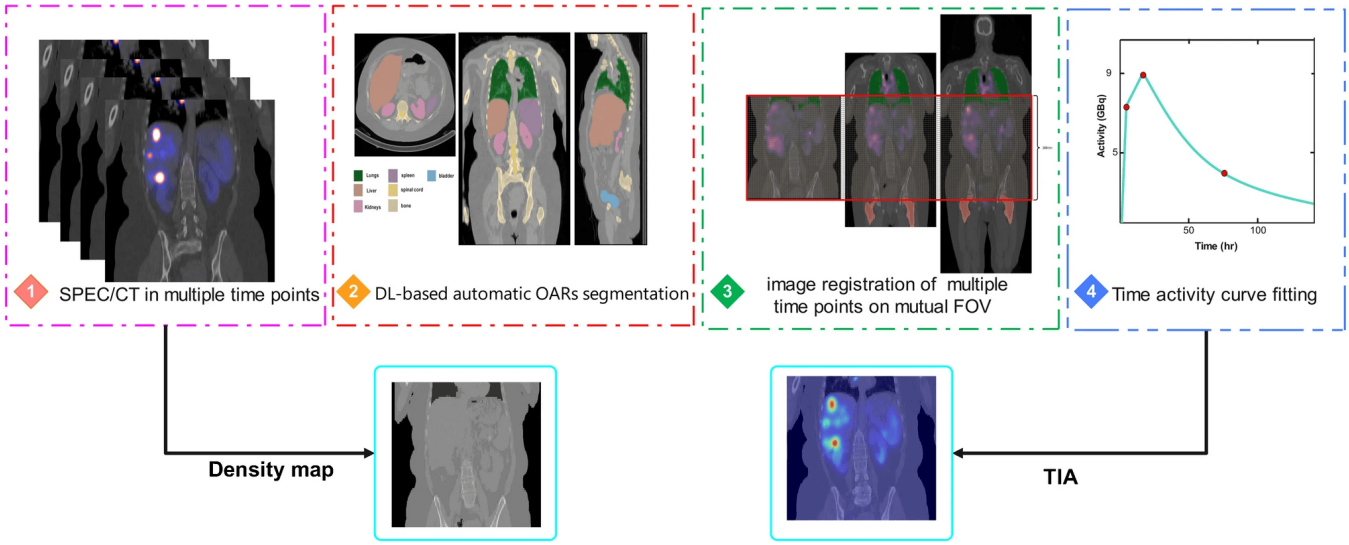

Figure 4: Workflow diagram of the hybrid DL model proposed by Mansouri et al. (2024), illustrating sequential SPECT/CT acquisition, DL-based segmentation of organs-at-risk, image registration, and time-activity curve fitting to compute voxel-level time-integrated activity (TIA).

The deep learning (DL) model used in this study was based on a modified UNETR (UNet-Transformer) architecture, originally proposed by Hatamizadeh et al. [10], and implemented using PyTorch on an NVIDIA GeForce RTX 3080 GPU. The model was trained to learn the residual difference between MC simulations and MSV-derived dose estimates, which is the goal value MC − MSV, instead of explicitly predicting Monte Carlo (MC) dosage maps. The inputs for the model were co-registered CT scans and MSV dosage maps. The result showed voxel-wise residuals, which were then added back to the MSV baseline to provide the complete MC dose estimate.

The network concentrates on fixing localized mistakes (such as boundary mismatches and heterogeneities) while maintaining the underlying dosage physics recorded in the MSV calculation thanks to this residual learning approach.

The Monte Carlo simulations, generated via a validated MCNP-based dose kernels, served as the ground truth in this framework due to their superior accuracy in modeling energy deposition across heterogeneous tissues.

The network focuses on correcting localized errors (such border mismatches and heterogeneities) while preserving the underlying dose physics documented in the MSV calculation. Illustrated in Figure 5. The transformer encoder processes co-registered CT images and MSV dose maps through embedded patches and self-attention. Residual connections feed into a series of convolutional layers, refining predictions to approximate the Monte Carlo (MC) dose map.

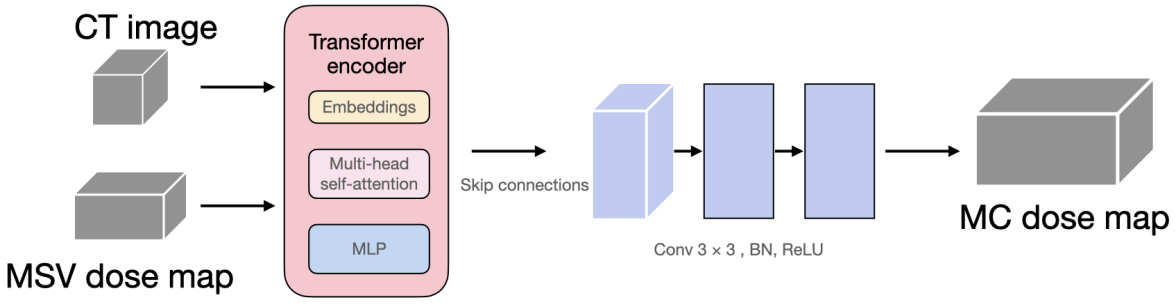

Figure 5: Simplified architecture of the modified UNETR network used in this study.

Monte Carlo methods offer stochastic modeling of particle transport with fine anatomical granularity, outperforming deterministic schemes in settings where tissue density and shape vary significantly.

The hybrid transformer-based deep learning model demonstrated efficient and accurate voxel-wise dosimetry predictions, with accuracy approaching the Monte Carlo gold standard.

The perfomance metrics used in the study include, Mean Absolute Error (MAE), organ-level dosage errors, and voxel-wise Root Mean Squared Error (RMSE). In areas of diverse tissue where the MIRD assumptions frequently fail, the model performed better than the SSV and MSV formalisms. Voxel-wise relative absolute error (RAE) maps for the three assessed dosimetry methodologies (SSV, MSV, and DL) with regard to MC are shown in Figure 6.

In thicker tissues like bone, the SSV formalism underestimates the absorbed dosage, whereas in low-density tissues like the lungs, it overestimates it. The DL model minimizes spatial inaccuracies, especially in diverse locations. Comparing with the MSV model it is a little gain with the DL in accuracy and a great gain in computation time.

To quantify these improvements, Figure 7 shows violin plots for key quantitative metrics across the three models. Such as SSIM, PSNR, Gamma pass rates, RMSE, MAE, MSE, ME, RE, and RAE. Structural Similarity Index Measure **SSIM**, ranges from 0 to 1, this metric evaluates the perceived similarity between two dose maps by comparing luminance, contrast, and structural patterns. Peak Signal-to-Noise Ratio **PSNR** Expressed in decibels (dB), PSNR evaluates the fidelity of the predicted dose maps by penalizing large voxel-wise deviations. Higher PSNR indicates better preservation of dose gradients. **Gamma pass rate** is a quality assurance metric that calculates the percentage of voxels that meet both dose difference and spatial tolerance criteria (e.g., 3%/3mm). DL achieved 97.35% within lesion volumes, slightly higher than MSV (97.2%) and SSV (96.8%). In term of **RMSE and MAE** DL yielded the lowest values, indicating minimal average deviation per voxel from MC predictions.

To further contextualize voxel-level performance, Table 1 presents the absolute percent error across key organs and lesion sites for each dosimetry method.

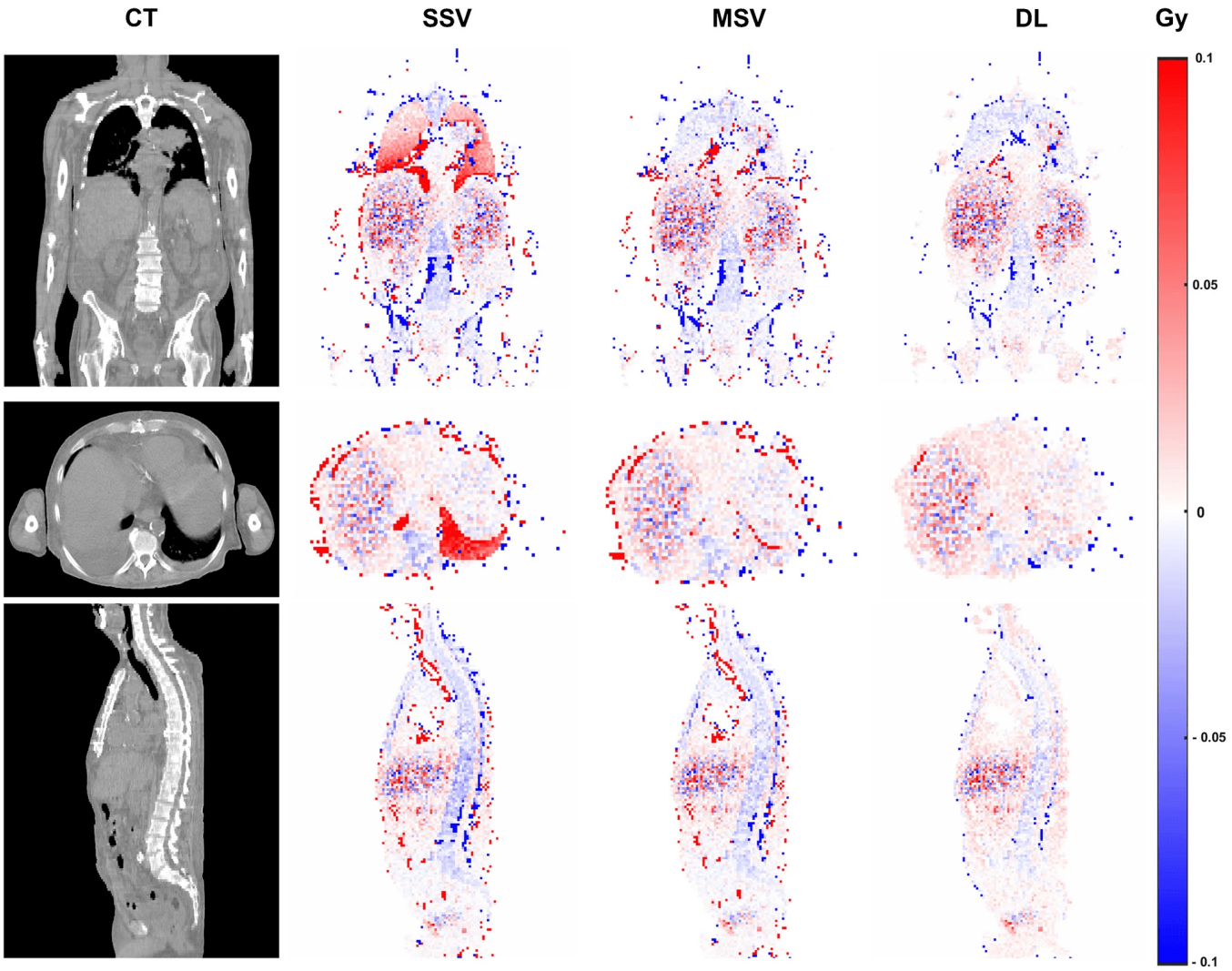

Figure 6: Comparison of dose maps across Single Voxel S-value (SSV), Multiple Voxel S-value (MSV), and Deep Learning (DL) methods.

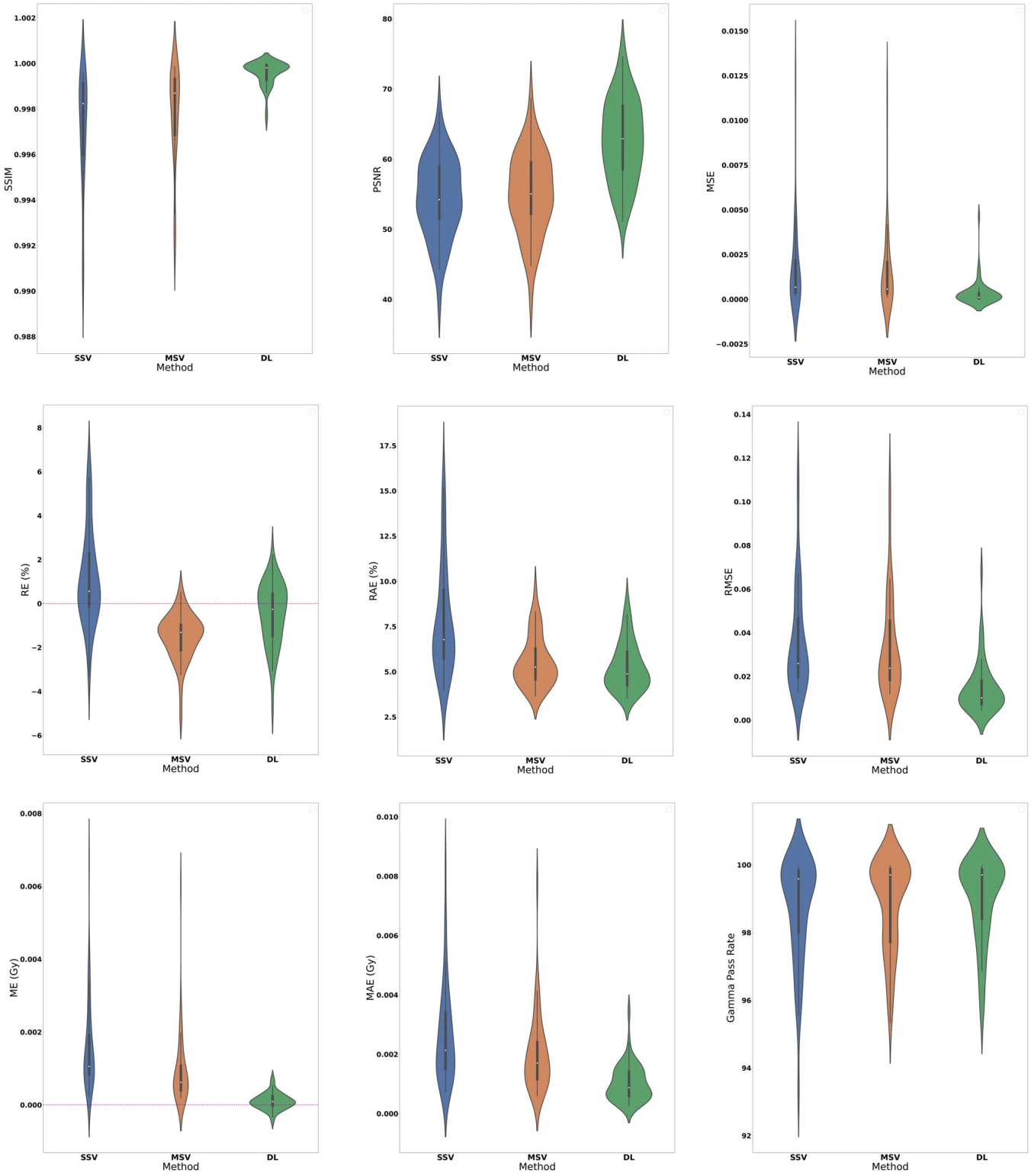

Figure 7: Violin plots comparing dosimetric accuracy metrics across SSV, MSV, and DL methods, including SSIM, PSNR, RMSE, MAE, and Gamma pass rate.

Table 1: Absolute percent error (%) for lesions and organs-at-risk (OARs) across three dosimetry approaches, with respect to Monte Carlo (MC) ground truth. Values are reported as mean ± standard deviation.

Organ / Region	**SSV**	**MSV**	**DL**
Lesions	1.44 ± 3.05	1.18 ± 2.65	1.15 ± 2.50
Left Kidney	0.16 ± 0.05	0.14 ± 0.05	0.30 ± 0.10
Right Kidney	0.45 ± 1.90	0.35 ± 1.26	0.49 ± 1.59
Liver	0.90 ± 0.93	0.84 ± 0.86	0.61 ± 0.40
Spleen	1.11 ± 2.10	1.06 ± 2.05	0.22 ± 0.20
Bones	17.07 ± 9.32	11.47 ± 6.40	8.21 ± 5.92
Lungs	45.00 ± 39.10	4.84 ± 5.64	2.60 ± 2.40
Spinal Cord	10.10 ± 11.44	9.96 ± 11.27	7.42 ± 14.25
Bladder	2.44 ± 7.03	2.42 ± 7.04	1.08 ± 1.81

The discussion emphasized that the DL approach offers a practical path to accelerate clinical adoption of voxel-based personalized dosimetry. It enhances therapeutic precision and supports patient safety by better resolving dose distributions in both tumors and organs at risk (OARs).

This directly relates to radiation protection principles from class—specifically optimization and justification—as this model supports the individualization of therapeutic plans, helping ensure that dose to healthy tissue is minimized while therapeutic efficacy is maintained.

From a clinical operations standpoint, the ability to generate dose maps within minutes—as opposed to hours with MSV or days with Monte Carlo—makes this approach promising for routine implementation. In terms of practical impact, this could significantly reduce the time and resources needed for personalized planning, aligning with the ALARA (As Low As Reasonably Achievable) principle not just for patient safety, but for staff exposure as well, by minimizing time spent in contact with radioactive materials.

This work aligns with the broader goals of radiation protection and individualized medicine in radiopharmaceutical therapy. It contributes to the growing body of research exploring AI-driven improvements to dosimetry workflows, demonstrating the potential for both improved accuracy and enhanced accessibility in modern nuclear medicine practice. In the following section we develope some ideas on what could be the implications on radiation protection when using this approach of dose estimation for both patients and staff.

4 Implications for Radiation Protection

When dealing with patients protection the Deep learning-enhanced voxel-based dosimetry offers significant advantages. It improves precision in absorbed dose estimates, particularly for organs at risk (OARs). Precision in this regard supports not only the individualization of treatments, but also the broader goals of radiation protection under the ICRP's justification and optimization principles.

For example, improved dose prediction is archived by estimating absorbed dose at the voxel level. This approach helps limit exposure to sensitive substructures. In particular, the hybrid deep learning model in [5] was trained to correct MSV dose maps with accuracy comparable to Monte Carlo simulations. With it, clinicians can better identify localized hotspots that may otherwise go unnoticed in organ-level approaches and reduce the risk of toxicity. With high-resolution dosimetry radiologists can identify patients at risk of exceeding critical thresholds, and make dose adjustments before adverse events occur [9].

Personalized voxel-level dose maps improve tumor targeting, potentially enhancing therapeutic outcomes while minimizing risk to non-target regions and allows to optimize therapeutic efficacy and establish dose-response relationships where comparison with older data will help refine safery protocols and thresholds. In line with that, Personalized dosimetry permits adjustments to administered activity based on patient-specific factors, improving therapeutic index and reducing likelihood of exceeding organ tolerance thresholds [3, 9].

The Mansouri et al. deep learning model was especially successful in reducing prediction error for vital organs such as the liver, lungs, and kidneys. For instance, the absolute percent dosage errors in kidneys were decreased to less than 0.5%, while in complicated heterogeneous tissues like lungs DL reached 2.6% error [5], much excelling both SSV and MSV approaches.

This precision supports safer personalized thresholds and enables a tighter control over absorbed doses to organs at risk, especially when standard population-based S-values might under- or over-estimate exposure.

Clinicians may preserve safety margins in treatment regimens with the support of this degree of accuracy, which allows for tighter control over absorbed dosage in important structures. This attention to detail recalls long-standing radiation safety concepts, even though patient-specific dose tolerances are not controlled in the same way as occupational limits. For instance, according to Riveira-Martin et al., occupational dose limitations include a maximum of 59 sessions per year based on extremities dose ($41.5 - 45.2 \mu Sv/GBq$ to the hands) and $100 - 197$ sessions per year based on eye or whole-body dose limits (e.g., $2.02 \mu Sv/GBq$ to the eye lens) [9]. The reasoning is the same, even if these are applicable to employees; radiation protection depends on knowing, measuring, and adhering to dosage margins. The same idea is supported by DL enhanced dosimetry; however, it is now more and more practical for every patient due to voxel-level speed and precision.

Now, with Staff in mind, Voxel-based DL workflows indirectly improve occupational radiation protection through efficiency and safety protocol reinforcement.

Most the assessment we can make on the work of Mansouri et al. in terms of radiation protection revolves around the proposed workflow. Although the authors did not study occupational exposure directly their hybrid DL model supports faster and automated voxel-level dose calculations, contributing to reduced on-site preparation and planning time. Protection must follow systemic strategies aimed to minimize exposure windows through procedure streamlining, planning standardization, and reduction of manual dosimetry tasks [5, 16, 17].

Standardization efforts include shielding strategies and pre-treatment contamination control. Mansouri et al. support this direction by presenting what appears as a clinically feasible transformer-based model designed to accelerate voxel-level dosimetry without compromising accuracy. [5, 16]. The automatization for contouring and dose calculation reduces exposure time by minimizing staff interaction windows during active radiopharmaceutical periods [7]. Additionally DL systems support real-time dosimetry estimation, allowing for faster infusion decisions which also leads to reduced staff proximity during therapy [9].

With the DL workflow archiving fast, accurate, and reproducible voxel-level dose maps, a reduction on the need for additional imaging sessions or late-stage manual recalculations follows. This indirectly supports protection strategies highlighted by Riveira-Martin et al., who emphasize minimizing repeat contact with patients and radioactive materials through efficient planning workflows [5, 9]. Additionally, it backs Wang's view, which outlines the institution's objectives to reduce procedural radiation risks in accordance with internal monitoring procedures such whole-body, lens, and extremity dosimetry that are utilized to enforce ALARA compliance and identify exposure vulnerabilities early [17].

Finally, Mansouri et al.'s hybrid DL model allows for precise, customized dose estimations without adding to patient burden or staff exposure, which is consistent with ethical principles of justification and optimization. Their work reinforces the fundamental principles of radiation protection as outlined by EURATOM and NRC rules, supporting the ethical duty to administer the bare lowest dosage required to produce therapeutic benefit [5, 16].

5 Challenges and Future Directions

Although voxel-based dosimetry augmented by deep learning offers encouraging benefits for patient-specific radiation therapy, a number of obstacles must be overcome before these techniques can be completely implemented in clinical settings. The reliance on diverse and high-quality training data is a significant drawback. According to Mansouri et al. [5], the size and diversity of the dataset may still have an impact on the transformer-based model's performance, even though it attained high accuracy during five-fold cross-validation. It may be difficult for models that were trained on a small number of cases or from a single institution to generalize to patients with unusual anatomical or pathological traits or to larger clinical settings.

Another concern is the interpretability of deep learning models, commonly referred to as the "black box" problem. This problem is important because the lack of transparency in model decisions could restrict clinician trust and slow down regulatory approval. Although Mansouri's model emphasizes residual learning to maintain physical dose restrictions, it is still challenging to completely understand the inner workings of the corrective process in the absence of further model explainability tools.

Furthermore, more thorough assessment is needed to confirm these model's validity and gener-

alizability. Dose accuracy may be affected by differences in scanner calibration, patient placement, and clinical protocols amongst centers [9]. External validations and multi-center trials are necessary to verify the reliability of DL-based dosimetry techniques across various imaging systems and patient populations.

Broadly speaking, regulatory concerns provide a significant impediment. It will take convincing proof of safety, repeatability, and therapeutic value to include AI tools into standard medical operations. Standardized monitoring and protocol creation ensure consistent practice and adherence to safety limits [17]. This idea should and must immediately be applied to all AI-augmented or assited dosimetry systems.

6 Conclusion

In this paper, we set out to answer how transformer-based deep learning models can improve voxel-level dosimetry accuracy in [^{177}Lu]Lu-DOTATATE therapy compared to conventional MIRD methods, and to explore the implications for radiation protection.

Our assessment of Mansouri et al.'s work, in which they used deep learning a modified UNETR residual architecture trained on co-registered CT and MSV dose maps, confirmed that DL can generate accurate voxel-wise estimates within minutes, rivaling the Monte Carlo gold standard.

The hybrid model achieved lower mean absolute errors across critical organs and minimized spatial inaccuracies in heterogeneous tissues such as lung and bone. It also demonstrated structural fidelity in predicted dose distributions, as evidenced by favorable SSIM and gamma pass rate metrics. From a radiation protection perspective, this improvement supports more accurate risk assessments and tighter dose constraints for organs at risk during treatment planing. Although the DL method does not directly modify occupational exposure protocols, it enhances planning efficiency and supports systemic safety goals, which can translate later to minimize staff interaction times and enabling workflow standardization.

Continued examination is necessary for the incorporation of AI into medical dosimetry. Interpretability, strong cross-institutional validation, and cautious regulatory alignment are necessary to moderate the use of transformer models like Mansouri's, even if they provide promise tools for optimization and personalization. Medical physicists and physicians should continue to actively define the role of AI as computing power increases, making sure that these technologies complement rather than replace patient safety frameworks and physical reasoning. Future research must concentrate on open, broadly applicable, and morally sound AI-driven processes that complement both new developments and accepted radiation safety guidelines.

References

[1] Jonathan Gear. Milestones in dosimetry for nuclear medicine therapy. *Br J Radiol*, 95(1135):20220056, Jul 2022.

[2] Joe O'Donoghue, Pat Zanzonico, John Humm, and Adam Kesner. Dosimetry in radiopharmaceutical therapy. *Journal of Nuclear Medicine*, 63(10):1467–1474, 2022.

[3] Keamogetswe Ramonaheng, Milani Qebetu, Honest Ndlovu, Cecile Swanepoel, Liani Smith, Sipho Mdanda, Amanda Mdlophane, and Mike Sathekge. Activity quantification and dosimetry in radiopharmaceutical therapy with reference to 177lutetium. *Frontiers in Nuclear Medicine*, Volume 4 - 2024, 2024.

[4] Greta S P Mok and Yuni K Dewaraja. Recent advances in voxel-based targeted radionuclide therapy dosimetry. *Quant Imaging Med Surg*, 11(2):483–489, Feb 2021.

[5] Zahra Mansouri, Yazdan Salimi, Azadeh Akhavanallaf, Isaac Shiri, Eliluane Pirazzo Andrade Teixeira, Xinchi Hou, Jean-Mathieu Beauregard, Arman Rahmim, and Habib Zaidi. Deep transformer-based personalized dosimetry from spect/ct images: a hybrid approach for [177lu]lu-dotatate radiopharmaceutical therapy. *European Journal of Nuclear Medicine and Molecular Imaging*, 51(6):1516–1529, 2024.

[6] Oncology Medical Physics. Mird internal dosimetry, 2025. Accessed: 2025-04-20.

[7] Laure Vergnaud, Yuni K. Dewaraja, Anne-Laure Giraudet, Jean-Noël Badel, and David Sarrut. A review of 177lu dosimetry workflows: how to reduce the imaging workloads? *EJNMMI Physics*, 11(1):65, 2024.

[8] Keon Min Kim, Min Sun Lee, Min Seok Suh, Gi Jeong Cheon, and Jae Sung Lee. Voxel-based internal dosimetry for (177)lu-labeled radiopharmaceutical therapy using deep residual learning. *Nucl Med Mol Imaging*, 57(2):94–102, Apr 2023.

[9] Mercedes Riveira-Martin, Lara Struelens, José Muñoz Iglesias, Werner Schoonjans, Olga Tabuenca, JoséManuel Nogueiras, Francisco Javier Salvador Gómez, and Antonio López Medina. Radiation exposure assessment of nuclear medicine staff administering [177lu]lu-dota-tate with active and passive dosimetry. *EJNMMI Physics*, 10(1):70, 2023.

[10] Amirhossein Hatamizadeh, Yucheng Tang, Vishwesh Nath, Dong Yang, Andriy Myronenko, Bennett Landman, Holger R Roth, and Daguang Xu. Unetr: Transformers for 3d medical image segmentation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 574–584, 2022.

[11] NVIDIA Research. Understanding ssim. NVIDIA Research Blog, 2020. `https://research.nvidia.com/publication/2020-understanding-ssim`.

[12] MathWorks. Psnr - compute peak signal-to-noise ratio. MathWorks Documentation, 2025. https://mathworks.com/help/vision/ref/psnr.html.

[13] Jim Frost. Root mean square error (rmse). Statistics By Jim, 2022. https://statisticsbyjim.com/regression/root-mean-square-error/.

[14] Wikipedia contributors. Mean squared error. Wikipedia, 2025. https://en.wikipedia.org/wiki/Mean_squared_error.

[15] U.S. Department of Energy and Office of Science. Gamma passing rate: Topics by science.gov. Science.gov, 2025. https://www.science.gov/topicpages/g/gamma+passing+rate.

[16] Derek John Cappon, Susanna Fang, Kendall Berry, Gina Capone, Gina L Carlton, Mario Chrétien, John Gough, Jacob Kamen, Anne Khoorshed, Andy Miller, Stephen Nelli, M Peter Petric, Frank Tourneur, and Josip J Zic. Clinical best practices for radiation safety during lutetium-177 therapy. *Health Phys*, 124(2):139–146, Feb 2023.

[17] Hong-Bo Wang, Qing-Zhao Zhang, Zhen Zhang, Chang-Song Hou, Wen-Liang Li, Hui Yang, and Quan-Fu Sun. Necessity of internal monitoring for nuclear medicine staff in a large specialized chinese hospital. *International Journal of Environmental Research and Public Health*, 13(4):418, April 2016.