Journal of Nuclear Medicine, published on April 14, 2022 as doi:10.2967/jnumed.121.263738

A pipeline for automated voxel dosimetry: application in patients with multi-SPECT/CT imaging following 177Lu peptide receptor radionuclide therapy

Yuni K Dewaraja, PhD[1], David M Mirando, B.S.[2], Avery B Peterson, B.S.[1], Jeremy Niedbala, CNMT[1], John D Millet, MD[1], Justin K Mikell, PhD[3], Kirk A Frey, MD/PhD[1], Ka kit Wong, MD[1], Scott J Wilderman, PhD[1], Aaron S Nelson, MD[2]

[1]Department of Radiology, [3]Radiation Oncology, University of Michigan, Ann Arbor, MI; [2]MIM Software Inc., Cleveland, OH

Corresponding Author: Yuni K Dewaraja

1301 Catherine, 2276 Medical Science I/5610

Ann Arbor, MI 48109

Email: yuni@umich.edu

ORCID: https://orcid.org/0000-0002-3920-6925

Phone: 734-647-2324

DISCLOSURES

This work was supported by grant R01CA240706, an Academic-Industry Partnership award from the National Cancer Institute. Yuni Dewaraja is a consultant for MIM Software and David Mirando and Aaron Nelson are employees of MIM. No other potential conflicts of interest relevant to this article exist.

Word count: 6012

Running title: Automated pipeline for voxel dosimetry

Abstract

Patient-specific dosimetry in radiopharmaceutical therapy (RPT) is impeded by the lack of tools that are accurate *and* practical for the clinic. The aims were to construct and test an integrated voxel-level pipeline that automates key components (organ segmentation, registration, dose-rate estimation, and curve fitting) of the RPT dosimetry process and then to use it to report patient specific dosimetry in 177Lu-DOTATATE therapy. **Methods.** An integrated workflow that automates the entire dosimetry process, except tumor segmentation, was constructed. 1) Convolutional neural networks (CNNs) are used to auto-segment organs on CT of SPECT/CT; 2) Local contour intensity-based SPECT-SPECT alignment results in volume-of-interest propagation to other timepoints; 3) Dose-rate estimation is performed by explicit Monte Carlo (MC) using the fast, Dose-Planning Method code; 4) The optimal function for dose-rate fitting is automatically selected for each voxel. When reporting mean dose, partial volume correction is applied, and uncertainty is estimated by an empirical approach of perturbing segmentations. **Results**. The workflow was used with 4-timepoint SPECT/CT imaging data from 20 patients with 77 neuroendocrine tumors, segmented by a radiologist. CNN-defined kidneys resulted in high Dice values (0.91-0.94) and only small differences (2-5%) in mean dose when compared with manual segmentation. Contour intensity-based registration led to visually enhanced alignment and the voxel-level fitting had high R^2 values. Across patients, dosimetry results were highly variable, for example, the average (range) of the mean absorbed dose in Gy/GBq were: lesions, 3.2(0.2-10.4); L kidney, 0.49(0.24-1.02); R kidney 0.54(0.31-1.07) and healthy liver, 0.51(0.27-1.04) Patient results further demonstrated the high variability in the number of cycles needed to deliver 'hypothetical' threshold absorbed doses of 23 Gy to kidney and 100 Gy to tumor. The uncertainty in mean dose, attributable to variability in segmentation, was on average (range) 6% (3-17%) for organs and 10% (3-37%) for lesions. For a typical patient, the time for the entire process was ~ 25 minutes (~ 2 min manual time) on a desktop computer, including time for CNN organ segmentation, co-registration, MC dosimetry and voxel curve fitting.

Conclusion. A pipeline integrating novel tools that are fast and automated provides the capacity for clinical translation of dosimetry guided RPT.

Radiopharmaceutical therapy (RPT) is well-suited for the theranostic treatment approach because imaging after one cycle could be used to predict absorbed doses per unit activity that will be delivered from subsequent cycles. Thus, the injected activity or number of cycles can be modulated on an individualized basis to potentially enhance therapeutic efficacy while keeping organ toxicities at an acceptable level *(1-3).* Despite this potential, fixed activity protocols continue to be used in RPTs like ^{177}Lu-DOTATATE peptide receptor radionuclide therapy (PRRT) where the standard practice is to deliver 4 cycles at 7.4 GBq/*cycle (1).*

In RPT dosimetry, 3-D patient images coupled with voxel-level dose estimation using methods such as Monte Carlo (MC) radiation transport can account for spatial and temporal activity nonuniformity as well as tissue heterogeneity down to the resolution limit of the imaging system. This contrasts with traditional methods that approximate the anatomy by a reference phantom model and provide only the average absorbed dose to structures that may not provide sufficient information to predict potential biological effects *(4).* However, voxel dosimetry can be logistically difficult to implement due to: 1) the need for repeated SPECT imaging with relatively long acquisitions, typically 20 - 30 min *(5)*, 2) co-registration of sequential images and voxel-level curve fitting, 3) computational cost of accurate dose-rate estimation using direct MC radiation transport. The potential for reducing the imaging burden associated with dosimetry has been recently reported *(6).* However, there is much room to improve the efficiency and accuracy of the other steps above as well as volume-of-interest (VOI) segmentation, which is needed to report dose metrics to lesions/organs.

Recent advances in automated medical image segmentation using deep learning methods mostly focus on external beam radiotherapy applications *(7).* Despite the potential for substantially reducing the variability and labor associated with segmentation, well-validated deep learning tools are not yet routinely available with RPT dosimetry software. Although

segmentation in RPT dosimetry is often performed on emission images, segmentation on co-registered morphologic images (CT, MRI), exploits their superior spatial resolution. The non-contrast low-dose CT of integrated SPECT/CT and PET/CT systems is suitable for segmenting some organs and well-defined lesions.

Among voxel dosimetry methods, MC is the consensus reference standard that is superior to simpler methods such as dose voxel (or point) kernel convolution (DVK) in the presence of tissue heterogeneity. However, RPT dosimetry using general purpose MC codes (e.g., GEANT4, MCNP) involve long simulation times, typically hours, to achieve reasonable statistics *(8).* The Dose Planning Method (DPM) MC code, optimized specifically for dose computations in voxelated geometries derived from CT scans, was originally developed, benchmarked and experimentally validated for fast dose estimation in external beam radiotherapy *(9).* We previously adapted DPM for application in I-131 radioimmunotherapy by sampling decay locations internally within the voxelated geometry and benchmarked it *(10).* Recently we performed measurements with radiochromic film to directly validate DPM for ^{177}Lu and ^{131}I *(11).* DPM achieves its significant acceleration via transport mechanics and electron multiple scattering distribution functions that permits long transport steps across heterogeneity boundaries *(9).*

Our aim is to construct and test a voxel-dosimetry pipeline that integrates deep learning- based segmentation, fast MC dose estimation and other automated processes and to use it to report highly patient specific dosimetry using multi-SPECT/CT imaging after 177Lu-PRRT.

MATERIALS AND METHODS

Dosimetry pipeline

The dosimetry pipeline (Fig 1) combining automated tools was implemented as a workflow within a commercial software platform (MIM Software Inc, Cleveland, Ohio). The only parts of the process that are not integrated are lesion segmentation and partial volume correction (PVC), which is applied as a post-processing step when reporting mean absorbed dose only. The workflow is currently in a beta testing stage. Following regulatory review and licensing agreements between The University of Michigan and MIM Software, a MC-based dosimetry workflow as described in this work will be made available for clinical use via MIM.

Segmentation. Organ segmentation is performed on the CT portion of the SPECT/CT at the reference timepoint, which is the timepoint where the CT was performed at a higher x-ray tube current and exposure (mAs) than at the other time points. For kidney and liver, we constructed convolutional neural networks (CNNs) based on RefineNet *(12),* which is based on the U-Net architecture with 3D convolution blocks to better leverage contextual information from all directions. The CNNs for kidney and liver were trained with 65 and 108 manually contoured data sets, respectively, with 5-fold cross validation. The previously evaluated Dice scores from these tests averaged 0.93 ± 0.04 and 0.97 ± 0.05 for the kidneys *(13)* and liver *(14)*, respectively. In the current study, we test the CNN for liver by visual assessment and for kidney by calculating geometric similarity measures (Dice and mean distance to agreement (MDA)) and comparing dose estimates relative to manual segmentations performed by an experienced technologist (JN), checked by a radiologist (JM). The Dice is a spatial overlap index and MDA is a spatial distance-based metric; both are widely used to assess the reliability of segmentations.

Auto-segmentation of lesions such as neuroendocrine tumor (NET) is not yet sufficiently developed/validated. Therefore, lesions (up to 5) were manually segmented by a radiologist on the patient's diagnostic quality baseline CT or MRI and then transferred to the reference SPECT/CT following co-registration. If mis-registration was evident, the location was manually adjusted based on SPECT uptake. Lesions clearly visible on the CT of the reference SPECT/CT were directly defined there. Only lesions that were well-defined and > 2 mL in volume were segmented to avoid large uncertainties. The healthy liver was defined as the liver minus any segmented lesions and therefore included any unsegmented lesions.

Local contour-intensity guided registration. To co-register sequential images, contour intensity-based SPECT-SPECT alignment was used. SPECT images are first rigidly registered to the reference SPECT as a gross alignment. The SPECT intensity information of voxels within each selected VOI plus a surrounding 7cm expansion zone is then used to perform multiple local rigid registrations between the images. These locally focused alignments are rigidly spliced together to generate a composite image aligned and tri-linearly interpolated to the grid of the reference SPECT. In this process, VOIs are automatically propagated from the reference to other timepoints. We tested this automated method previously in a subset of patients by comparing time-integrated activity values relative to rigidly transferring VOIs with visual fine-tuning and reported good agreement: 0.3% (95% CI: [-8.0%, 8.7%]) for kidneys and 1.9% (95% CI: [-17.8%, 21.7%]) for tumors *(15).*

Voxel dose-rate estimation: Details of the MC transport mechanics and physics data that is based on PENELOPE can be found in the original paper on DPM development *(9)* and our subsequent paper *(10)* on extension to RPT. For the current work, the 177Lu decay and spectral information came from NuDat 3.0 *(16)* and BetaShape 2.2 *(17).* The cutoff for photon and electron transport were set to 4 kev and 59 keV, respectively, below which energy was

locally deposited in the voxel. This is a reasonable considering that the range of a 59 keV electron in tissue (~ 0.07 mm) and the mean free path of a 4 keV photon in tissue (~ 0.1 mm) are much smaller than the voxel size of the dose-rate map indicated below. For a typical case (patient 12), we tested this assumption by running DPM with lower cutoffs (photons 2 kev, electrons 20 kev), which results in a 20% increase in run time, but the average difference in voxel dose-rate in a VOI was only 0.5%.

Dose-rate estimation by DPM MC and DVK convolution with soft tissue kernels pre-calculated using DPM were integrated as options within the dosimetry workflow to enable comparison. Exploiting the high speed of Fast Fourier Transform convolution, we chose to work with a large kernel (267x267x267, 3x3x3 mm^3) to ensure capturing of photon dose contribution from decays anywhere within the SPECT field-of-view. The mass density map for MC transport or for density scaling of the DVK results is generated via a CT-to-density calibration curve, that we determined specifically for our system. The co-registered quantitative SPECT image and the reference timepoint density map are the inputs to the dose-rate calculation, which is repeated at each timepoint. Images are re-sampled to 167x167x131 (3x3x3 mm^3) for DVK and to 128x128x100 (3.91x3.91x3.91 mm^3) for MC. For MC, based on testing of statistical uncertainty, 10^8 histories were simulated.

*Voxel-level curve fitting to generate dose maps***.** First, at each voxel, dose-rate, $\dot{D}$*,* as a function of time is fit by a mono-exponential and a 3-parameter bi-exponential of the form $\dot{D}(t) = C_1(e^{-\lambda_1 t} - e^{-\lambda_2 t})$ using standard least-squares optimization. Next, the best fit function is automatically selected independently for each voxel based on the Akaike information criterion as proposed by Sarrut et al *(18).* Other options such as trapezoidal approximation as well as a 4-parameter bi-exponential, suitable when more time points are available, were implemented,

but not used in the current study. Dose-rate functions are integrated analytically to generate the dose-map.

Mean absorbed dose and dose-volume histogram (DVH) metrics. The segmented contours are applied to the absorbed dose map to generate the mean absorbed dose and other DVH metrics. Absorbed dose estimates derived from SPECT images are degraded by partial volume errors associated with limited spatial resolution. For mean value partial volume correction (PVC), we determined recovery coefficients (RCs) defined as the ratio of SPECT measured activity to true activity *(19)* within CT-defined sphere VOIs of a multi-sphere phantom with sphere-to-background ratio of 6.2:1 and sphere sizes in the range 2 to 113 mL. The same SPECT/CT system and imaging/reconstruction parameters used in patient studies described below were used for this phantom experiment. The fit to our RC vs. volume data was: RC=1-1/(1+(v/3.673)^0.948), where v is the volume in mL. This function was used to determine volume-dependent RCs for lesions/organs and were applied (scaling by 1/RC) when reporting mean absorbed doses only.

Patient characteristics and SPECT/CT imaging

Patient studies. Data corresponds to 20 patients (Supplemental Table 1) with NETs who completed 4-timepoint SPECT/CT imaging after cycle 1 of standard 177Lu-DOTATATE PRRT performed at the University of Michigan Medical Center between August 2018 and July 2021. The research imaging was approved by the Institutional Review Board, and all patients provided written informed consent.

Quantitative SPECT/CT. The SPECT/CT system was a Siemens Intevo Bold equipped with a ME collimator and a 15 mm crystal. The 4 SPECT acquisitions (25sec/view) were performed at

day 0 (directly after completion of amino acid infusion), day 1, day 4 - 5, and day 5 - 8. Manufacture recommended settings of a 20% acquisition window at 208 keV with adjacent 10% scatter windows, 256 x 256 matrix and 120 views were used. Siemens xSPECT Quant software was used to directly generate reconstructed images in Bq/mL units. Here, a National Institute of Standards and Technology traceable ^{75}Se calibration source with a 3% uncertainty is used to perform a monthly site-specific check of the system sensitivity and associated fine tuning *(19).* SPECT reconstruction parameters were: 48 iterations (1 subset) of OSCG with resolution recovery, a 256 x 256 x 199 matrix (1.953 mm^3) and no post filtering *(19).* The non-contrast, free breathing CT at the reference timepoint was performed at 120 kVp and 80 mAs, and at 15 mAs at other timepoints. The CT reconstruction matrix was 512 x 512 x 130 (0.97 x 0.97 mm^2 x 3 mm).

Uncertainty analysis

In RPT dosimetry the traditional approach of propagating uncertainties associated with each step has identified VOI delineation as the largest source of variability *(20).* The analytical approach proposed for estimating this variability is not well suited when anatomical imaging is used for segmentation because factors other than spatial resolution (for example, use of contrast agents or mis-alignment with SPECT) can dominate. In addition, the spatial distribution of absorbed dose impacts the uncertainty estimate. Hence, we take an empirical approach of introducing realistic variabilities to the segmented VOIs and determining the corresponding variability in the mean absorbed dose when the perturbed contours are applied to the dose-map and the perturbed volumes used to determine RCs. We introduced volume expansions/contractions (organs:±5%, ±10%; lesions:±10%, ±20%) and 26 translations of 4 mm (~ SPECT voxel size) for each perturbation. These levels were selected based on an ongoing inter-operator variability study on lesion segmentation at our institution and on the kidney

volume differences for CNN vs. manual segmentation in the current study. For each VOI, the relative standard deviation (COV) in mean absorbed dose from all perturbations is reported as the uncertainty.

RESULTS

Results correspond to imaging data after cycle 1 with 18/20 patients treated with the 7.4 GBq administration and 2 treated with a reduced level of ~ 3.2 GBq (Supplemental Table 1).

Segmentation

Example segmentations are shown in Fig. 2. The median (range) of lesion volumes was 22 (2 - 1039) mL and 64/77 were in the liver (Supplemental Table 1). In most cases, the CNN-organ segmentation was accepted by the radiologist. When manual adjustment was needed, it took only 30 s to 3 m per organ. The fine tuning was mostly related to inclusion of bowel loops with liver and cysts with kidney (Fig. 2B). Comparing manual vs. CNN defined kidney, Dice scores were on average (range) 0.91 (0.77-0.94) without and 0.93 (0.91-0.94) with manual adjustment and differences in mean absorbed doses were 3% (0-21%) without and 2%(0-4%) with adjustment (Supplemental Table 2).

Contour intensity-based SPECT-SPECT alignment

Visual inspection demonstrated well-aligned images (Fig 3) across all time-points with only 3/164 VOIs needing manual adjustment. These 3 cases were two small lymph node lesions in the pelvis and a lesion in the mesentery where there was substantial movement between scans.

Voxel-level Curve fitting

For a typical patient, a map of the coefficient of determination (R^2) that measures the goodness of the fit at each voxel and example dose-rate curve fits are shown in Fig. 4. The R^2 values across all voxels in all patients are summarized in Fig 4B. Summary statistics for the effective half-life (T_{eff}) values of the main component of the exponential are presented in Supplemental Table 3.

Dosimetry

Example MC dose maps and DVHs are presented in Fig. 5A. Individual mean absorbed dose values following PVC are presented in Fig 6 and supplemental Figure 1 while summary statistics averaged across all patients are presented in Supplemental Table 3. The median (range) value of RCs applied were: lesions 0.85 (0.37-1.00); kidney 0.97 (0.95-0.98) and healthy liver 1.00 (1.00-1.00). There was a weak correlation between tumor volume and mean absorbed dose (R^2=0.052, p=0.046). DVH metrics are presented in supplemental Figure 2.

Uncertainty in mean absorbed dose estimates

With 10^8 MC histories, the contribution of statistical uncertainty is negligible. For example, the COV in VOI dose-rates (from the DPM uncertainty maps) was < 0.2% for the case shown in Fig 5A. The uncertainty associated with SPECT-SPECT misalignment can be approximated empirically by intentionally mis-aligning the four dose-rate maps and repeating the voxel-level auto fitting process. This process is not practical to perform due to the various combination of shifts, and we therefore limit this process to one illustrative example (patient #26). This case was chosen as the right kidney is in very close proximity to a tumor (Fig 5B) and therefore the impact of misalignment can be expected to be higher than typical, while the other structures are representative of typical patients. With shifts of +- 4 mm in all directions the estimated COV in mean absorbed dose due to mis-registration was: Tumor, 1.8 to 2.6%; L Kidney 1.6%; R Kidney,

9.8%; Healthy liver, 1.4%. In general, the 4 mm misalignment has a relatively low impact due to the poor resolution of SPECT.

The uncertainty that we report for mean absorbed doses is that associated with segmentation, which also impacts the volume-dependent RCs. The COV, represented as error bars in Supplemental Figure 1, was on average (range): left kidney, 5%(4-6%); right kidney, 7%(3-17%); spleen, 5%(4-9%); tumor >10mL, 8%(3-15%); tumor <10mL, 16%(8-37%). For organs and large tumor, the major contributor to this uncertainty was the contour variability itself while for small tumor it was the sensitivity of RCs to tumor volume uncertainty. For example, for tumor volumes < 10 mL, the COV in reported absorbed doses increased by a factor of 2 on average when applying RCs while for tumor > 100 cc this factor was only 1.1 on average.

Comparison between MC, DVK and OLINDA

Mean absorbed doses from MC, DVK and OLINDA 1.0 are compared in Fig. 6D-E and demonstrate good agreement. For the OLINDA calculation, the time-integrated activity values from the workflow were used. The difference between DVH metrics from MC and DVK are presented in Supplemental Figure 2.

Computational Cost

All processing was performed on a multi-core Mac Pro (3.2 GHz) desktop computer. For a typical patient (128x128x100 matrix), the total run time starting with four SPECT/CT images is ~ 25 minutes for the MC option and ~ 12 minutes for DVK, of which only ~ 2 min is manual time. This includes CNN-organ segmentation (~ 2 min), contour-based SPECT alignment (3 minutes), MC dose-rate estimation (4 min/per timepoint), and voxel-level dose-rate fitting and integration (3 min). Currently, DPM runs on a single processor but can be accelerated with parallelization.

DISCUSSION

A highly automated pipeline for all steps of RPT voxel dosimetry that needs minimal user interaction was constructed and tested for the ^{177}Lu PRRT application. The automated tools are suitable for other radionuclides and therapies with appropriate testing.

Although activity quantification was not evaluated in the current study, accuracy within 1.2% for large objects has been reported *(19)* with xSPECT Quant software using the same standardized calibration process and reconstruction parameters used in the current study. Furthermore, the sphere phantom-based RCs reported in that study and in our study are in good agreement. This simplistic PVC has limitations because RCs depend on factors other than the volume of the object, such as shape and activity distribution. For a 2-compartment kidney model with uniform activity in the renal cortex it has been shown that the RC for the cortex is substantially lower than that for a sphere of equal volume *(21).* Generating geometry/distribution specific RCs is beyond the scope of the current study and may not be practical due to wide patient-to-patient variations. Despite the limitations, the sphere-based RCs used in our study are widely used as a mean value correction to mitigate resolution effects in SPECT and PET quantification, including in dosimetry applications *(22,23).*

Ideally, when evaluating 3D dose-distributions PVC should be applied at the voxel-level, but this is a challenging and yet unresolved problem *(24).*Another challenge with voxel dosimetry is that iterative reconstruction with resolution modeling leads to edge artefacts that, depending on size of the object, can manifest itself as a visible dip or an overshoot at the center (*25).* This effect and partial volume effects leads to differences between SPECT-derived DVHs and histograms corresponding to the true activity distribution within an object *(26).* Despite the challenges, capturing non-uniformities in the underlying dose distribution to the extent possible is desired

when considering biological effect. There are alternative efforts using deep learning, including by our group *(27),* to mitigate the impact of poor spatial resolution and reconstruction artefacts on voxel dosimetry.

There was close agreement between dose estimates corresponding to manual vs. CNN-defined kidney, which is consistent with findings for a ^{177}Lu-PSMA therapy cohort *(28).* We achieved further improvement with quick manual adjustment of the CNN-segmentations. The Dice similarity metrics we achieved are comparable to what is reported for inter/intra-operator variability in manual organ segmentation *(29).* A possible refinement is to train the CNN to use both CT and SPECT information, which may be beneficial to identify cysts included in kidney contours (Fig. 2B) and to reduce effects of mis-registration between CT and SPECT images *(30).* Visually, the local contour intensity-based SPECT-SPECT registration led to improved alignments compared with global rigid registration (Fig. 3). The main advantage of SPECT intensity-based registration over using CT-based registration to align serial SPECT images is that it does not depend on SPECT-CT alignment, which is inconsistent due to respiratory motion and patient movement. The automated approach for selecting the optimal fit function led to high R^2 values and is especially beneficial for voxel-level fitting where selection by visual inspection is not feasible. The smooth voxel dose-rate vs. time data, high R^2 values and the good agreement in T_{eff} when comparing average values from voxel-level vs. organ-level fitting gives us confidence in the ability to perform accurate voxel-level fitting.

It is notable that the difference between the mean absorbed dose from OLINDA sphere model and MC was < 10% for 73/77 lesions (Fig 6E). This difference can, however, be larger for radionuclides with more significant photon yields than ^{177}Lu. The main advantage of MC over conventional DVK methods is in regions of heterogeneous tissue and at steep activity gradients, which is not fully resolved by simple density scaling. Recently, more sophisticated approaches

that utilize multiple kernels to cover media of varying density have been proposed to mitigate the limitations of using a single kernel *(31).* In the current study, lesions in heterogeneous tissue such as bone metastases were not considered due to their small size and/or the difficulty defining them. Accurate dosimetry of bone lesions and marrow requires modeling the fine structure of the spongiosa that can be coupled with MC transport *(32).*

RPT dosimetry is typically reported without uncertainty estimates due to the complexity and limitations of traditional error propagation. Although we did not attempt to assess all components of the uncertainty, we captured the main components: the variability associated with contour delineation and the volume-based RCs *(20).* Facilitated by voxel dose maps, our empirical approach of introducing perturbations is both practical to implement and is sensitive to each patient's underlying dose distribution, which is not the case with analytical approaches. For example, in patient 26 the relative uncertainty for the right kidney was higher than for the left due to proximity to a lesion in the liver (Fig 5B), which may also explain the higher values reported overall for right vs. left in our study and by others *(1,3).*

Across patients, the median T_{eff} that we observed for left (51.7h) and right (50.3 h) kidney agree closely with previous reports *(1,6)* for similar cohorts. The median value of absorbed dose to left (0.41 Gy/GBq) and right (0.43 Gy/GBq) kidney can be compared with past reported values for kidney of 0.61(1), 0.54 *(2)*, 0.47 *(33),* and 0.38 (average) *(34)* Gy/GBq. Differences can be attributed to differences in patient cohorts and quantification/dosimetry methods, as well as the renoprotective amino acid preparation. The median value of the tumor mean absorbed dose in our study (17.7 Gy or 2.7 Gy/GBq) is lower than the per cycle median values of 50 Gy *(23)* and 4.4 Gy/GBq (*2)* reported in two other studies. This is not surprising because instead of anatomical lesion segmentation used in our study, SPECT thresholding was used in the first study while the second used VOIs placed over the area of maximum uptake.

Although we do not report biological effective dose, it can be calculated using the dose-rate maps and fit parameters available from the workflow coupled with published values for the radiobiologic parameters as outlined by Baechler et al *(35)*. Furthermore, since most patients in our study are yet to finish their 4-cycle treatment, no attempt was made to evaluate dose-outcome relationships. PRRT clinical trials where dosimetry is used solely to avoid toxicity to critical organs have been reported *(1-3).* Performing both tumor and organ dosimetry enables consideration of both efficacy and toxicity when planning subsequent cycles. As a hypothetical example, using the cycle 1 mean dose estimates from our study, we determined the number of cycles each patient would need to receive 100 Gy to tumor and 23 Gy to kidney assuming the same Gy/per cycle. These thresholds, though somewhat arbitrary, were selected here based on a dose-response report *(23*) and the generally accepted 23 Gy limit for nephrotoxicity in external beam radiotherapy, although this limit is likely too low for PRRT. Most patients could receive more than the standard 4 cycles without exceeding 23 Gy to kidney, while most lesions need more than 4 cycles to achieve 100 Gy to tumor (Fig. 7). Access to this type of information in real time via efficient dosimetry calculations, will enable clinicians to make well-informed treatment decisions. Furthermore, highly patient specific dosimetry results are needed to establish the validity of simplified protocols

CONCLUSION

Multiple steps of the dosimetry chain, organ segmentation, co-registration, dose-rate estimation by MC, and curve fitting, were automated and integrated to construct an efficient pipeline for voxel dosimetry that was tested using imaging data from 177Lu-PRRT patients. Facilitated by this efficient workflow, highly patient specific dosimetry results including uncertainty estimates were reported. The speed, automation and accuracy that was achieved will facilitate implementation of real time dosimetry guided RPT in the clinical setting.

DISCLOSURES

This work was supported by grant R01CA240706, an Academic-Industry Partnership award from the NCI. Yuni Dewaraja is a consultant for MIM Software and David Mirando and Aaron Nelson are employees of MIM. No other potential conflicts of interest relevant to this article exist.

DATA AVAILABILITY

All images associated with patient #6 are shared at University of Michigan Deep Blue data repository: https://doi.org/10.7302/0n8e-rz46 and https://doi.org/10.7302/vhrh-qg23

KEY POINTS

Question Can accurate and highly patient specific dosimetry be performed in a clinical setting to guide RPT?

Pertinent findings An automated pipeline that includes CNNs for organ segmentation and a fast MC code for dose-estimation was constructed, tested and applied to report dosimetry in patients undergoing ^{177}Lu-PRRT. Excluding the time for lesion segmentation, voxel-level MC dose estimates were achieved in ~ 25 min.

Implications for patient care Patient specific real time dosimetry, such as performed in the current study after cycle 1, can be used to tailor subsequent cycles based on efficacy and toxicity considerations.

REFERENCES

1. Sundlöv A, Sjögreen-Gleisner K, Svensson J, et al. Individualised ^{177}Lu-DOTATATE treatment of neuroendocrine tumours based on kidney dosimetry. *Eur J Nucl Med Mol Imaging*. 2017;44(9):1480-1489.
2. Del Prete M, Buteau, FA, Arsenault F, et al. Personalized ^{177}Lu-octreotate peptide receptor radionuclide therapy of neuroendocrine tumours: initial results from the P-PRRT trial. *Eur J Nucl Med Mol Imaging.* 2019;46:728-742.3.
3. Sandstrom M, Garske-Roman U, Johansson S, et al. Kidney dosimetry during 177Lu-DOTATATE therapy in patients with neuroendocrine tumors: aspects on calculation and tolerance. Acta Oncol. 2018;57:516-521.
4. O'Donoghue JA. Implications of nonuniform tumor doses for radioimmunotherapy. *J Nucl Med.* 1999;40:1337-41.
5. Rydén T, Van Essen M, Marin I, Svensson J, Bernhardt P. Deep-Learning generation of synthetic intermediate projections improves ^{177}Lu SPECT images reconstructed with sparsely acquired projections. *J Nucl Med*. 2021;62:528-535.
6. Hanscheid H, Lapa C, Buck AK, Lassmann M, Werner RA. Dose mapping after endoradiotherapy with ^{177}Lu-DOTATATE/DOTATOC by a single measurement after 4 days. *J Nucl Med.* 2018;59:75-81.7.
7. Sahiner B, Pezeshk A, Hadjiiski LM, et al. Deep learning in medical imaging and radiation therapy. *Med Phys.* 2019;46:e1-e36. 8.
8. Huizing DMV, de Wit-van der Veen BJ, Verheij M, Stokkel MPM. Dosimetry methods and clinical applications in peptide receptor radionuclide therapy for neuroendocrine tumours: a literature review. *EJNMMI Res*. 2018;8:89. 9.

9. Sempau J, Wilderman SJ, Bielajew AF. DPM, a fast, accurate Monte Carlo code optimized for photon and electron radiotherapy treatment planning dose calculations. *Phys Med Biol.* 2000;45:2263-91. 10.
10. Wilderman SJ, Dewaraja YK. Method for fast CT/SPECT-Based 3D Monte Carlo absorbed dose computations in internal emitter therapy. *IEEE Trans Nucl Sci*. 2007;54:146-151.
11. Van B, Dewaraja YK, Niedbala J, et al. Reproducible absorbed dose measurements for validating dosimetry of short range therapeutic unsealed beta sources including Monte Carlo-based calculations. *J Nucl Med*. 2021;62(suppl 1):1577.
12. Lin G, Milan A, Shen C, Reid I. RefineNet: Multi-path refinement networks for high-resolution semantic segmentation. https://arxiv.org/abs/1611.06612. Accessed on February 16, 2022.
13. Cole N, Wan H, Niedbala J, et al. Impact of a 3D convolution neural network method on liver segmentation: An accuracy and time-savings evaluation. Annual meeting of the AAPM,2020.https://w3.aapm.org/meetings/2020AM/programInfo/programAbs.php?sid=8489&aid=51984. Accessed on February 16, 2022.
14. Lamba N, Wan H, Kruzer A, et al. Clinical utility of a 3D convolutional neural network kidney segmentation method for radionuclide dosimetry. *J Nuc Med*. 2019;60 (suppl 1):267.
15. Mirando D, Dewaraja YK, Cole N, Nelson AS. In pursuit of fully automated dosimetry: evaluation of an automatic VOI propagation algorithm using contour intensity-based SPECT alignments. *Eur J Nucl Med Mol Imaging.*2020;47(suppl 1):S236
16. National Nuclear Data Center. https://www.nndc.bnl.gov/nudat3/. Accessed on February 16, 2022.

17. Laboratorie National Henri Becquerel. http://www.lnhb.fr/rd-activities/spectrum-processing-software/. Accessed on February 16, 2022.
18. Sarrut D, Halty A, Badel JN, Ferrer L, Bardiès M. Voxel-based multimodel fitting method for modeling time activity curves in SPECT images. *Med Phys*. 2017;44:6280-6288.
19. Tran-Gia J, Lassmann M. Characterization of noise and resolution for quantitative ^{177}Lu SPECT/CT with xSPECT Quant. *J Nucl Med.* 2019;60:50-59.
20. Gear JI, Cox MG, Gustafsson J, et al. EANM practical guidance on uncertainty analysis for molecular radiotherapy absorbed dose calculations. *Eur J Nucl Med Mol Imaging*. 2018;45:2456-2474.
21. Tran-Gia J, Lassmann M. Optimizing Image Quantification for ^{177}Lu SPECT/CT based on a 3D printed 2-compartment kidney phantom. *J Nucl Med*. 2018;59:616-624.
22. Finocchiaro D, Berenato S, Grassi E, et al. Partial volume effect of SPECT images in PRRT with 177Lu labelled somatostatin analogues: A practical solution. *Phys Med.* 2019;57:153-159.
23. Ilan E, Sandström M, Wassberg C, et al. Dose response of pancreatic neuroendocrine tumors treated with peptide receptor radionuclide therapy using 177Lu-DOTATATE. *J Nucl Med*. 2015;56:177-82.24.
24. Tran-Gia J, Salas-Ramirez M, Lassmann M. What you see is not what you get: on the accuracy of voxel-based dosimetry in molecular radiotherapy. *J Nucl Med.* 2020;61:1178-1186.
25. Rahmim A, Qi J, Sossi V. Resolution modeling in PET imaging: theory, practice, benefits, and pitfalls. *Med Phys*. 2013;40:064301
26. Dewaraja YK, Frey EC, Sgouros G,et al. MIRD pamphlet No. 23: quantitative SPECT for patient-specific 3-dimensional dosimetry in internal radionuclide therapy. *J Nucl Med*. 2012;53:1310-1325.

27. Li Z, Fessler JA, Mikell JK, Wilderman SJ, Dewaraja YK. DblurDoseNet: A deep residual learning network for voxel radionuclide dosimetry compensating for single-photon emission computerized tomography imaging resolution. *Med Phys*. 2022;49:1216-1230.
28. Jackson P, Hardcastle N, Dawe N, et al. Deep learning renal segmentation for fully automated radiation dose estimation in unsealed source therapy. *Front Oncol*. 2018;8:215.28.
29. Daniel AJ, Buchanan CE, Allcock T, et al. Automated renal segmentation in healthy and chronic kidney disease subjects using a convolutional neural network. *Magn Reson Med*. 2021;86:1125-1136.
30. Khan J, Rydèn T, Van Essen M, Svensson J, Bernhardt P. Activity concentration estimation in automated kidney segmentation based on convolution neural network method for ^{177}lu–spect/ct kidney dosimetry, *Radiation Protection Dosimetry*. 2021;195:164-171.
31. Lee MS, Kim JH, Paeng JC,et al. Whole-body voxel-based personalized dosimetry: the multiple voxel S-value approach for heterogeneous media with nonuniform activity distributions. *J Nucl Med*. 2018 ;59:1133-1139.
32. Wilderman SJ, Roberson PL, Bolch WE, Dewaraja YK. Investigation of effect of variations in bone fraction and red marrow cellularity on bone marrow dosimetry in radio-immunotherapy. *Phys Med Biol*. 2013;58:4717-31.
33. Staanum PF, Frellsen AF, Olesen ML, et al. Practical kidney dosimetry in peptide receptor radionuclide therapy using [^{177}Lu]Lu-DOTATOC and [^{177}Lu]Lu-DOTATATE with focus on uncertainty estimates. *EJNMMI Phys.* 2021;8:78
34. Willowson KP, Eslick E, Ryu H, et al. Feasibility and accuracy of single timepoint imaging for renal dosimetry following ^{177}Lu-DOTATATE ('Lutate') therapy. *EJNMMI Phys*. 2018;5:33.

35. Baechler S, Hobbs RF, Prideaux AR, Wahl RL, Sgouros G. Extension of the biological effective dose to the MIRD schema and possible implications in radionuclide therapy dosimetry. *Med Phys*. 2008;35:1123-34.

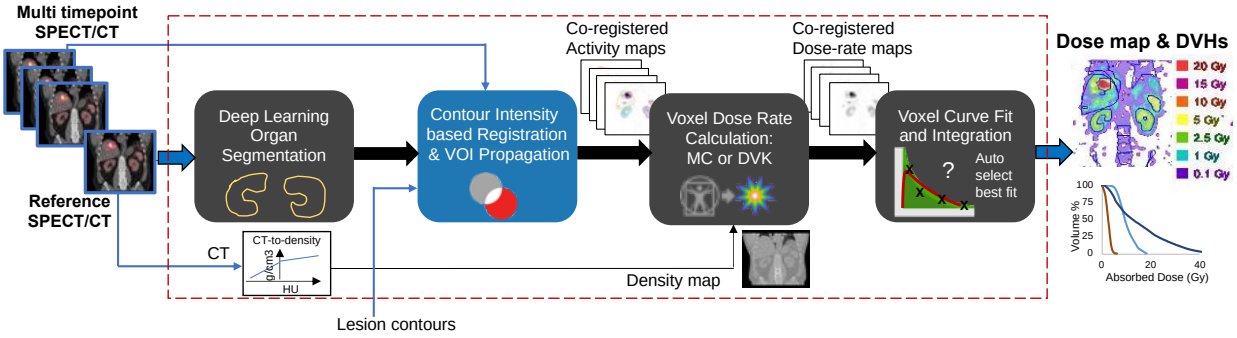


FIGURE 1. Overview of the automated dosimetry workflow. At the “reference” timepoint the CT of the SPECT/CT was performed at a higher mAs value than at other timepoints.

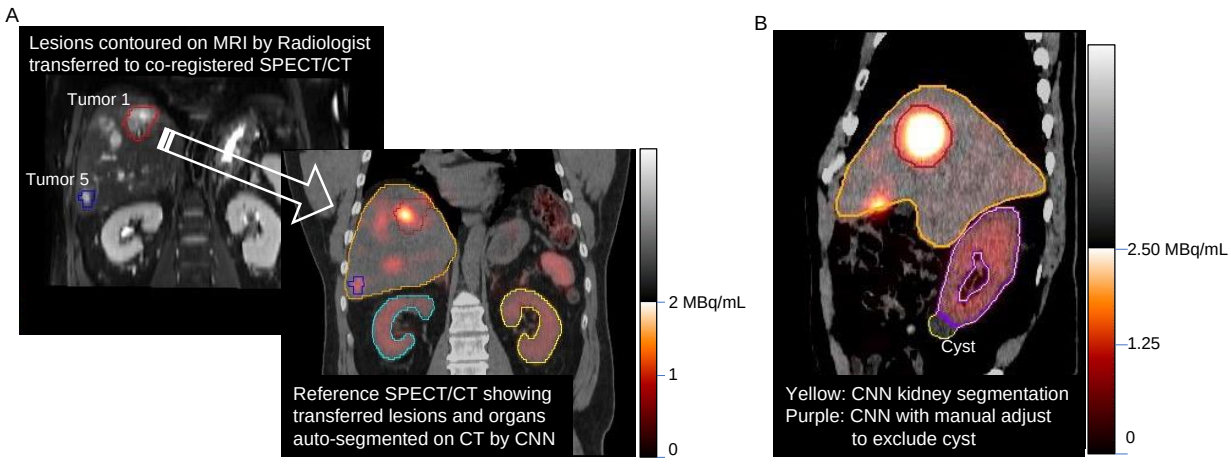


FIGURE 2. Example segmentations for (A) patient 12 and (B) patient 4. In (A), the CNN defined kidney and liver was accepted by the radiologist while in (B) manual adjustment of kidney was needed to avoid a cyst.

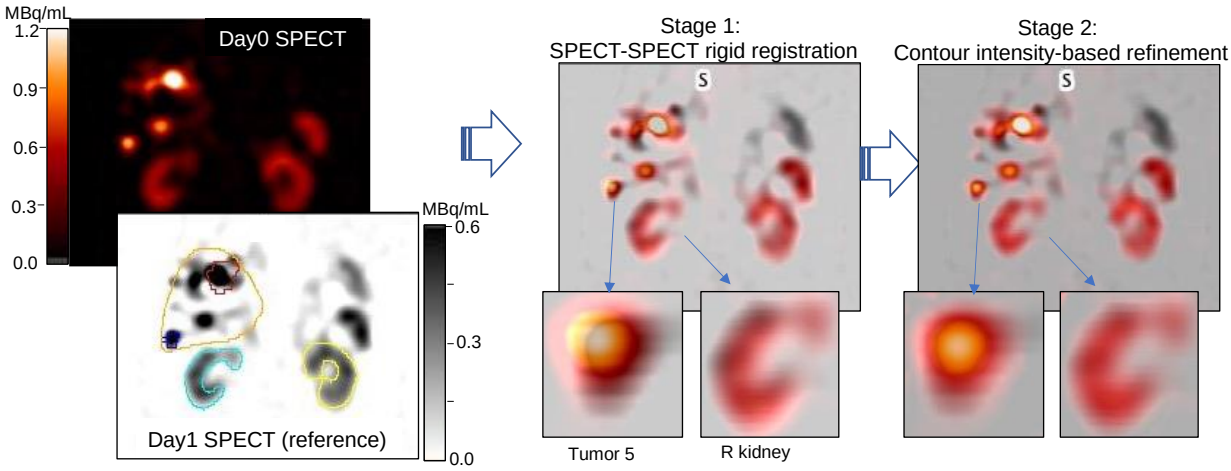


FIGURE 3. The two-stages of the automatic SPECT-SPECT registration demonstrated for patient 12 Day0 and Day1 images. Comparison of magnified inserts show subtle improvement in alignment with the contour intensity-based refinement.

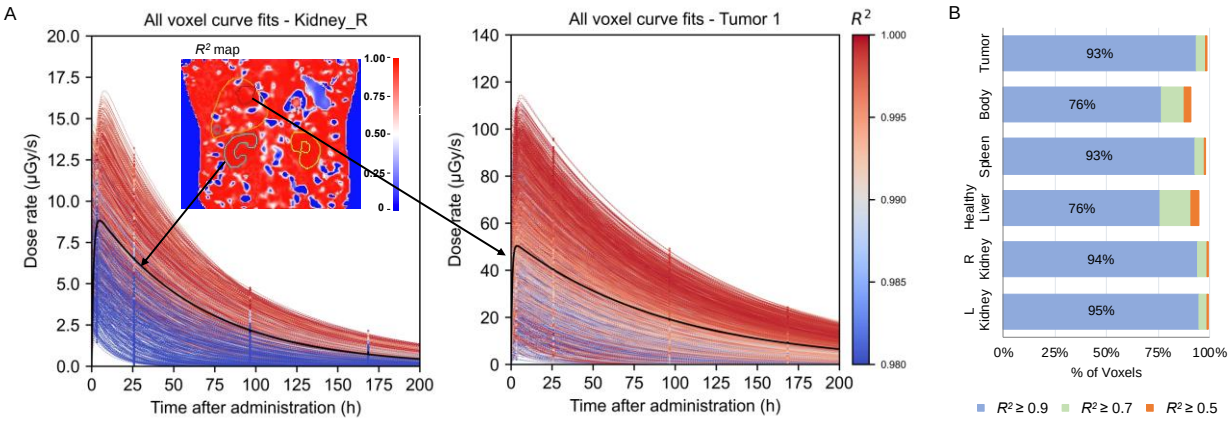
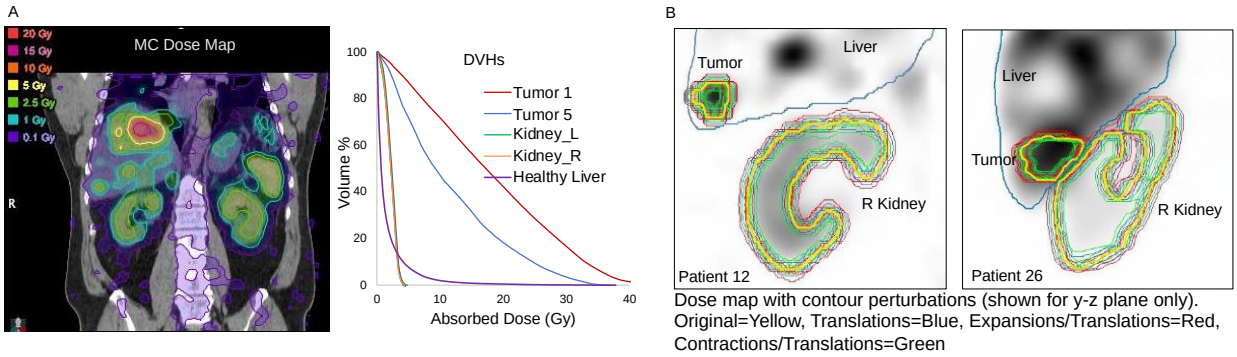


FIGURE 4. (A) Example dose-rate curve fits shown for all voxels in right kidney and tumor 1 of patient 12, color-coded by R^2 value of the fit. Note the selection of mono- and bi-exponential functions for different voxels. The organ-level fit-curve is shown in black. A coronal slice of the R^2 map is shown in the insert. (B) The percentage of voxels achieving $R^2 > 0.5$, 0.7 and 0.9 for fits across all patients.



Dose map with contour perturbations (shown for y-z plane only). Original=Yellow, Translations=Blue, Expansions/Translations=Red, Contractions/Translations=Green

FIGURE 5. (A) A coronal slice of the MC dose-map and dose volume-histograms for patient 12. (B) Examples of contour perturbations applied to dose maps of patient 12 and 26 for estimating uncertainty. The COV in right kidney absorbed dose was 4% for patient 12, but 17% for patient 26 due to proximity to a tumor and liver.

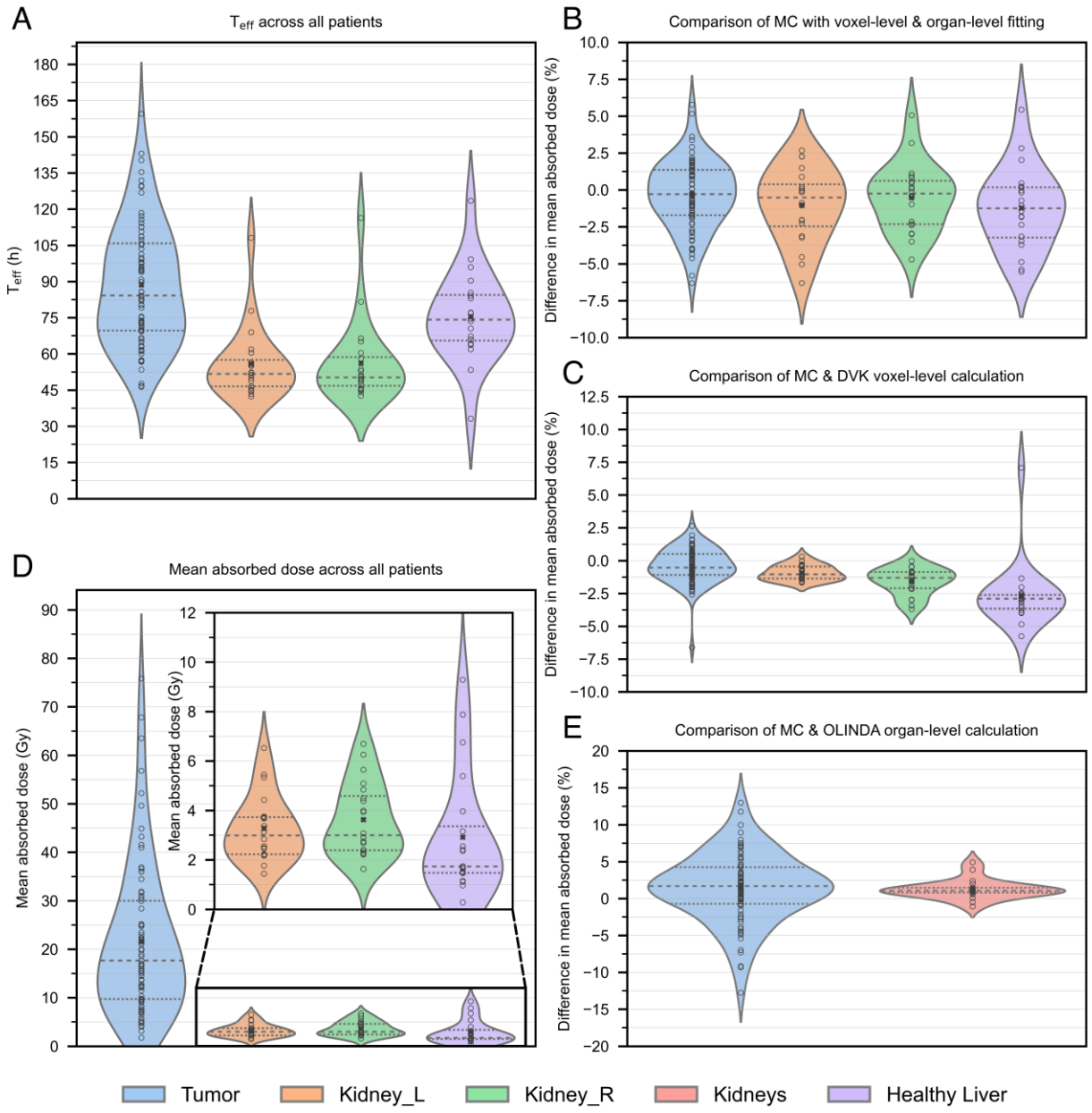


FIGURE 6. Violin plots of dosimetry results across all patients. A) Effective half-life. B) Mean absorbed dose with voxel-level MC calculation, C) Difference in MC dose estimates with voxel-level vs. organ-level dose-rate fitting, D) Difference between MC vs. DVK (with density scaling), E) Difference between MC vs. OLINDA (with mass scaling). Difference =100*(MC – XX)/MC. The circles indicate individual points, x indicates the average, and the dash lines indicate the 25%th and 75%th percentiles.

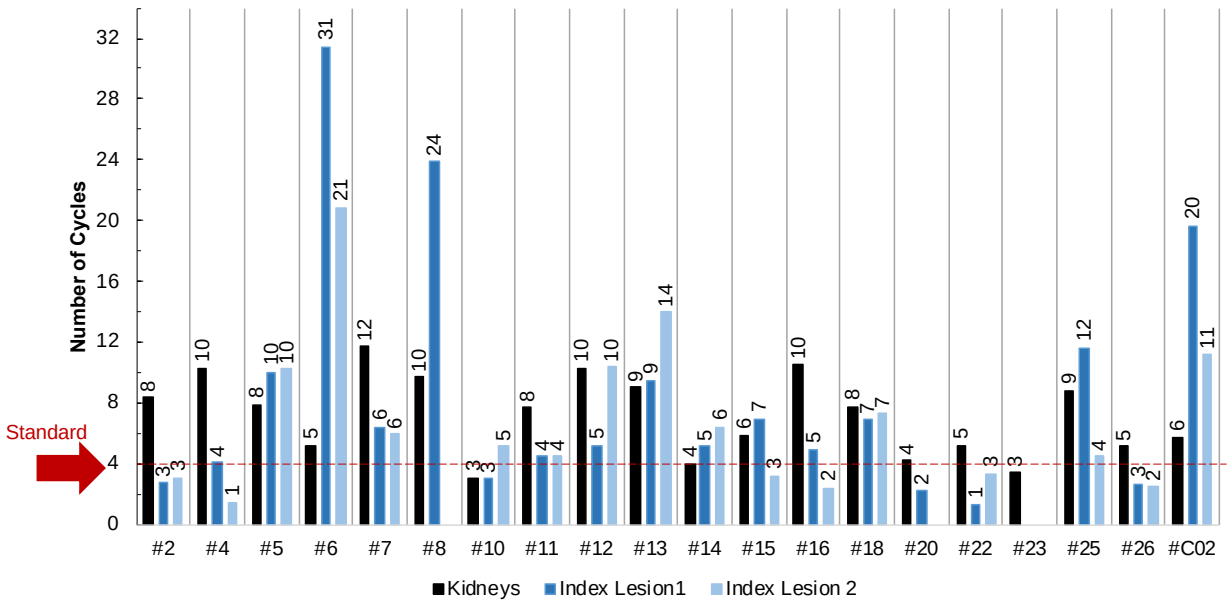
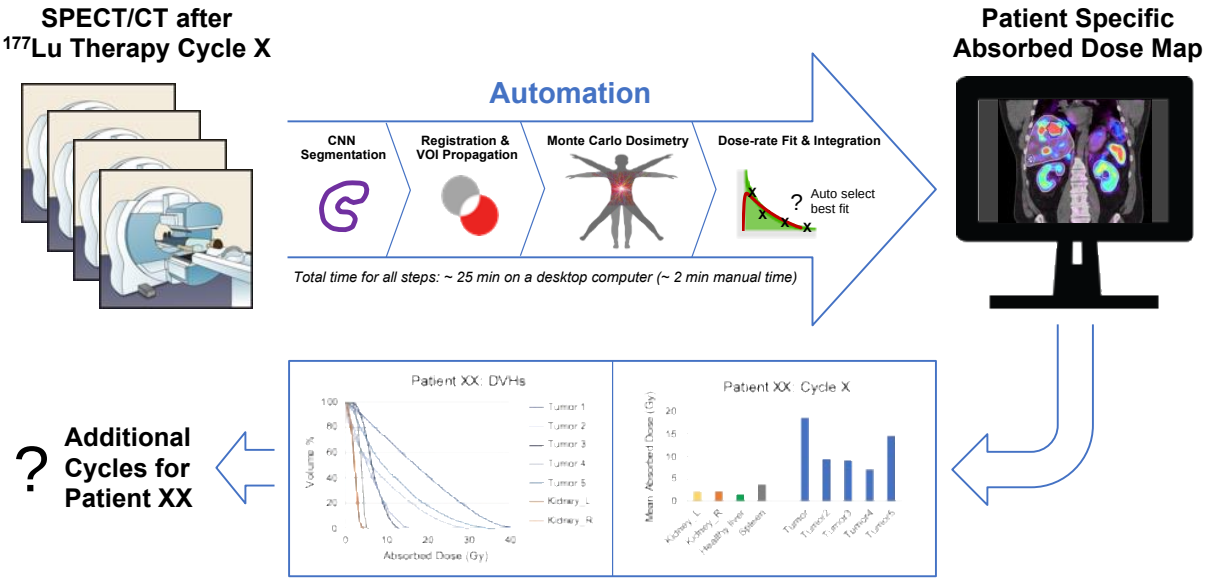


FIGURE 7. Variation in number of (7.4 GBq) cycles needed to deliver 23 Gy to kidney and 100 Gy to tumor.

Graphical Abstract



Supplemental Material

Platform for automated patient specific dosimetry combining deep learning segmentation and fast voxel-level tools including Monte Carlo: application in ^{177}Lu-PRRT Yuni K Dewaraja et al

Supplemental Table 1. *Patient baseline characteristics and cycle 1 mean absorbed dose estimate from the DPM Monte Carlo calculation with voxel-level dose-rate fitting. The standard deviation estimated from contour perturbations is in parenthesis. Healthy liver standard deviation is not included as not all liver lesions were segmented. Tumor volumes are indicated. Median (range) of organ volumes: L kidney, 162(78-275)mL; R kidney, 163(118-258)mL; Healthy liver, 1821(1166-3369)mL.*

								Mean Absorbed Dose (Gy)							
Patient Code	**Sex**	**Age (y)**	**Weight (kg)**	**Grade**	**Ki-67**	**eGFR^{+} (mL/min)**	**Admin Activity (GBq)**	**L Kidney**	**R Kidney**	**Healthy Liver**	**Tumor 1**	**Tumor 2**	**Tumor 3**	**Tumor 4**	**Tumor 5**
2	M	71	96	na	na	82	7.14	2.6 (0.2)	2.8 (0.1)	1.0	52.2(11.0) 2.9 mL*	36.1 (5.5) 4.3 mL*	31.7 (5.2) 3.5 mL	22.3 (5.6) 2.5 mL	
4	M	64	98	na	na	81	7.13	2.2 (0.1)	2.2 (0.1)	1.5	23.5 (2.4) 106.1 mL	67.8(12.0) 3.1 mL			
5	F	72	55	G2	11%	96	7.34	2.8 (0.1)	3.1 (0.3)	5.4	31.0 (3.1) 12.1 mL	13.6 (0.9) 55.4 mL	28.4 (2.5) 11.2 mL	9.6 (0.7) 58.2 mL	9.8 (0.7) 59 mL
6	F	62	52	na	na	43	7.33	3.7 (0.2)	4.8 (0.3)	1.5	4.2 (0.3) 11.7 mL	63.5(13.6) 3.0 mL	3.1 (0.2) 68.1 mL*	4.7 (0.6) 22.2 mL**	
7	F	39	129	G3	15%-20%	81	7.48	1.8 (0.1)	2.3 (0.3)	7.9	16.0 (0.8) 373 mL	17.1 (0.7) 207 mL	23.0 (1.4) 22.7 mL	19.0 (1.2) 13.7 mL	24.8 (2.0) 10.2 mL
8	M	56	90	na	na	109	7.07	2.2 (0.1)	2.4 (0.1)	0.3	4.1 (0.5) 569 mL***				
10	F	66	94	na	na	25	3.69^{++}	3.8 (0.2)	4.0 (0.3)	3.1	16.7 (1.2) 272 mL***	6.6 (0.4) 128 mL	9.7 (0.3) 296 mL	5.8 (0.5) 18.5 mL	12.4 (1.0) 90 mL
11	M	70	72	G2	3%	68	7.34	3.1 (0.2)	2.9 (0.2)	6.8	22.1 (1.3) 155 mL	21.1 (2.8) 12.6 mL	25.2 (1.3) 49.4 mL	25.0 (2.2) 22.8 mL	21.9 (1.2) 104 mL
12	M	57	90	G2	5%	91	7.20	2.2 (0.1)	2.3 (0.1)	1.6	18.5 (2.0) 142 mL	9.3 (1.3) 65.6 mL	9.2 (0.9) 10.9 mL	7.1 (1.1) 17.8 mL	14.5 (2.0) 15.5 mL
13	M	66	87	G2	10%-15%	91	7.31	2.5 (0.1)	2.7 (0.2)	4.0	7.0 (0.2) 29.8 mL	17.7 (1.6) 10.0 mL	7.8 (0.6) 28.3 mL	10.3 (0.7) 33.5 mL	12.1 (1.4) 5.6 mL*
14	M	56	113	G1	<2%	67	7.35	5.3 (0.3)	6.3 (0.6)	2.4	19.8 (1.2) 1039 mL	1.8 (0.2) 8.0 mL	21.5 (2.0) 22.7 mL	22.4 (2.6) 23.4 mL	15.4 (1.4) 99.7 mL*
15	M	72	79	G2	9 per 10 hpf	55	7.14	3.7 (0.2)	4.0 (0.1)	1.7	13.7 (0.9) 80.0 mL	30.0 (3.1) 13.4 mL	43.2(17.6) 2.1 mL		
16	M	74	86	G1	na	88	7.37	2.2 (0.1)	2.3 (0.1)	2.6	20.1 (1.7) 91.4 mL	16.5 (1.8) 8.3 mL	41.5 (4.0) 35.3 mL	36.0 (9.1) 2.4 mL	15.0 (1.3) 16.5 mL****
18	M	71	68	G2	10%	91	3.75^{++}	1.4 (0.1)	1.6 (0.1)	1.7	12.4 (1.1) 233 mL	7.0 (0.5) 256 mL	7.4 (0.3) 268 mL	11.5 (0.9) 208 mL	
20	M	70	97	G2	12%, 20%	66	7.38	5.5 (0.3)	5.7 (0.3)	1.5	44.9 (4.0) 28.8 mL				
22	M	67	74	G1	<2%	92	7.31	4.4 (0.2)	4.5 (0.3)	1.5	30.5 (2.5) 9.2 mL	15.7 (2.3) 3.9 mL	75.8 (7.3) 11.6 mL	21.2(2.1) 7.0 mL**	
23	F	74	69	G1	<2%	48	7.15	6.5 (0.3)	6.7 (0.2)	1.2					
25	M	66	77	G2	na	91	7.22	2.5 (0.1)	2.7 (0.1)	1.1	8.3 (0.7) 77.0 mL	21.96(2.0) 46.2 mL	56.8 (4.8) 7.5 mL*	31.9 (2.8) 10.9 mL*	
26	F	76	79	G2	5%-10%	55	6.81	3.4 (0.2)	5.1 (0.9)	9.3	34.5 (2.3) 24.0 mL	36.6 (2.1) 20.5 mL	41.0 (2.6) 10.3 ml	36.9 (3.3) 21.4 mL	49.6 (4.0) 9.9 mL
c02	F	76	89	G2	7%	49	7.30	3.7 (0.1)	4.4 (0.2)	2.4	9.1 (0.7) 7.3 mL	12.7 (1.6) 4.6 mL	9.7 (1.1) 3.9 mL	5.3 (0.6) 7.1 mL	5.0 (0.5) 71.5 mL****

$^{+}$Estimated glomerular filtration rate. Normal eGFR > 59 mL/min

$^{++}$ Reduced activity due to concern of renal or marrow toxicity

*Lymph node lesion; ** Body wall lesion; ***Pancreas lesion; **** Mesenteric lesion; All other lesions are in the liver

***Supplemental Table 2**. Performance of deep learning kidney segmentation (without and with fine tuning) relative to manual segmentation for the first 14 patients (27 kidney[*]). In general, fine tuning of CNN-generated contours did not result in large changes in similarity measures or absorbed dose, but in one case the Dice score improved from 0.77 to 0.91 and mean absorbed dose agreement improved from -21% to 4%.*

	Manual vs. Fully Automated CNN-segmentation					**Manual vs. CNN Segmentation with Manual Fine Tuning[*]**				
	Volume Absolute Difference	Mean Dose Absolute Difference	Dice	HD (mm)	MDA (mm)	Volume Absolute Difference	Mean Dose Absolute Difference	Dice	HD (mm)	MDA (mm)
L Kidney										
Mean	5%	2%	0.92	10.7	0.92	4%	1%	0.93	8.3	0.80
Median	4%	1%	0.93	8.5	0.78	3%	1%	0.93	8.2	0.76
Min	0%	0%	0.85	6.0	0.68	0%	0%	0.86	6.0	0.68
Max	18%	5%	0.94	36.0	2.04	17%	5%	0.94	12.2	1.19
R Kidney										
Mean	8%	3%	0.91	11.4	0.99	5%	2%	0.93	9.9	0.81
Median	6%	2%	0.93	9.2	0.84	6%	1%	0.93	8.8	0.81
Min	0%	0%	0.77	4.5	0.68	0%	0%	0.91	4.5	0.68
Max	27%	21%	0.94	24.4	2.05	11%	4%	0.94	24.4	0.99

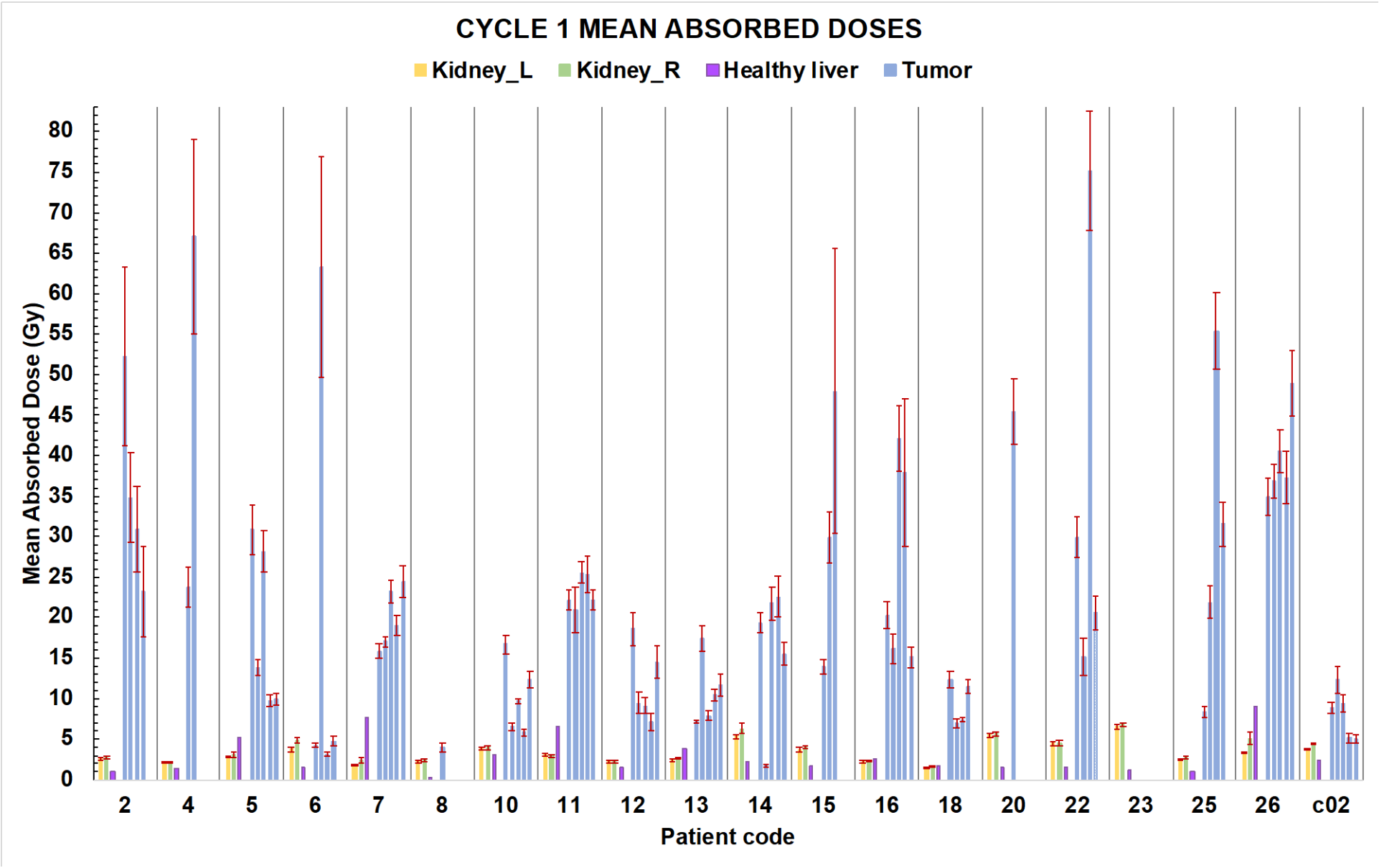
[*]CNN failed to locate right kidney in one case due to unusual location

***Supplemental Table 3.** Effective half-life and mean absorbed dose statistics across all patients from DPM MC dose estimation with voxel-level dose-rate fitting.*

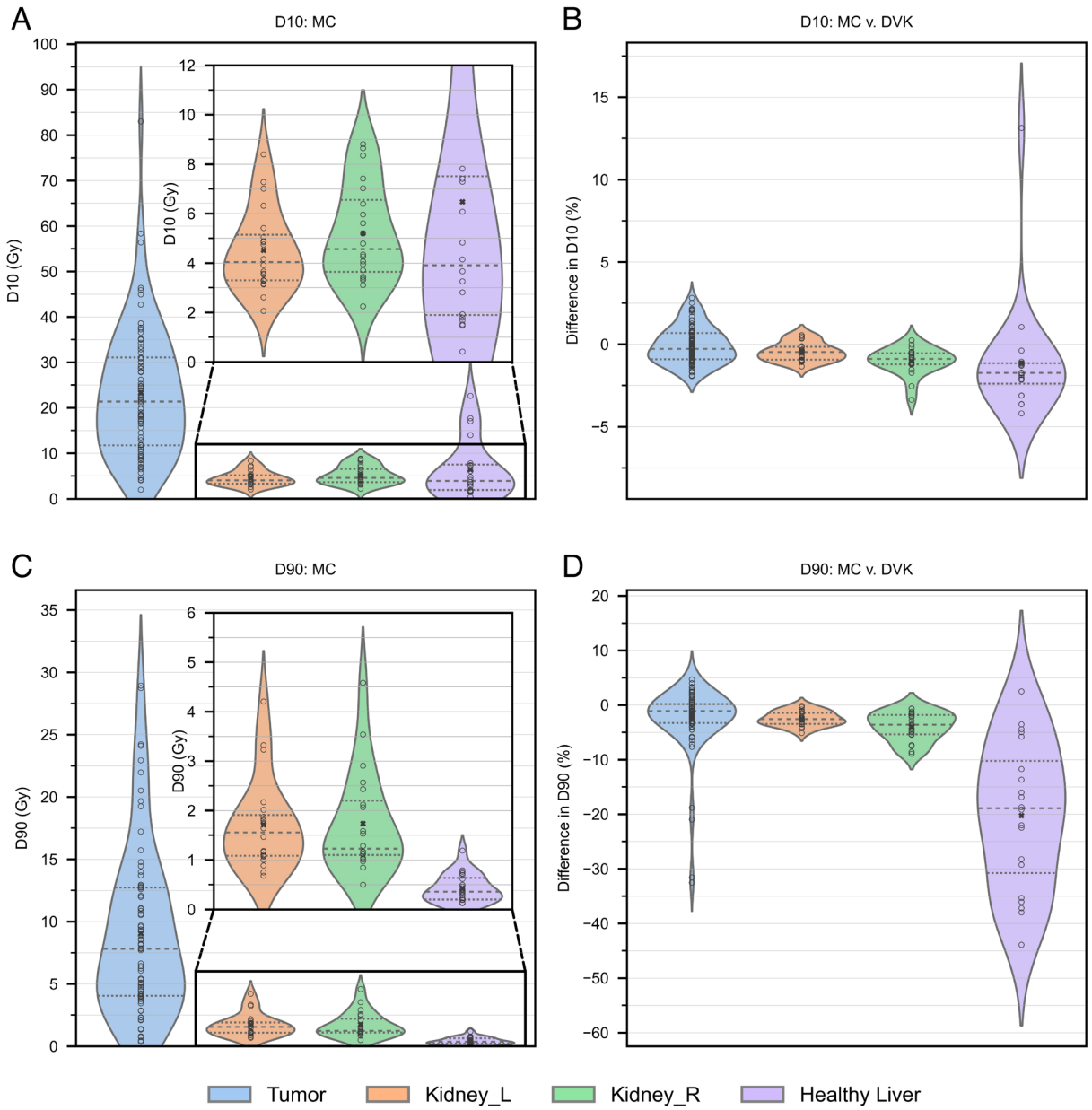
	Tumor	**L Kidney**	**R Kidney**	**Healthy Liver**
Sample size	77	20	20	20
Effective Half Life (h)[*]				
Average	84.4 (88.6)[**]	52.4 (55.9)	51.9 (56.2)	67.1 (75.5)
Median	80.4 (84.2)	49.1 (51.7)	48.6 (50.3)	64.1 (74.1)
Min	8.0 (46.3)	8.0 (42.4)	8.0 (42.7)	8.0 (33.1)
Max	159.5 (159.5)	159.5 (108.2)	159.5 (116.4)	159.5 (123.5)
STD	20.6 (25.3)	20.6 (15.2)	21.6 (17.0)	36.7 (18.8)
Mean Absorbed Dose				
Average (Gy) (Gy/GBq)	21.7 3.16	3.3 0.49	3.6 0.54	2.9 0.51
Median (Gy) (Gy/GBq)	17.7 2.72	3.0 0.41	3.0 0.43	1.7 0.41
Min. (Gy) (Gy/GBq)	1.8 0.24	1.4 0.24	1.6 0.31	0.3 0.27
Max. (Gy) (Gy/GBq)	75.8 10.37	6.5 1.02	6.7 1.07	9.3 1.04
STD (Gy) (Gy/GBq)	15.8 2.16	1.3 0.21	1.5 0.23	2.5 0.22

[*] Corresponding to the main component of the exponential fit

[**]*Effective half-life values in parenthesis correspond to organ level dose-rate fitting.*



Supplemental Figure 1. *Cycle 1 organ and lesion mean absorbed doses for all patients calculated using the MC voxel-level option. The error bars indicate 1 standard deviation estimated from the contour perturbation approach. Healthy liver standard deviation is not included as not all liver lesions were segmented. See supplemental Table 1 for administered activities.*



Supplemental Figure 2. *DVH statistics A) D10, B) D90 from DPM MC dose estimation. Difference in C) D10 and D) D90 estimated from MC vs. DVK convolution. Difference = 100*(MC-DVK)/MC. These metrics are reported without PVC, because RCs were used only as a mean value correction.*