UNIVERSITÉ DE GENÈVE FACULTÉ DE MÉDECINE

Section de médecine clinique
Département de radiologie et informatique médicale Professeur H. Zaidi

Strategies for radiation dose monitoring and optimization in diagnostic and therapeutic nuclear medicine procedures

THÈSE

présentée aux Facultés de médecine et des sciences de l'Université de Genève
pour obtenir le grade de Docteur ès sciences en sciences de la vie,
mention Physique du vivant

par

Azadeh AKHAVANALLAF

de

Mashhad, Iran

Thèse Nº 210

GENÈVE

Service de reprographie de l'Uni Mail

2023

LIFE SCIENCES PHD SCHOOL GENEVA

DOCTORAT ÈS SCIENCES EN SCIENCES DE LA VIE DES
FACULTÉS DE MÉDECINE ET DES SCIENCES
MENTION PHYSIQUE DU VIVANT

Thèse de Mme Azadeh Akhavanallaf

intitulée :

«Strategies for Radiation Dose Monitoring and Optimization in Diagnostic and Therapeutic Nuclear Medicine Procedures»

Les Facultés de médecine et des sciences, sur le préavis de Monsieur Habib ZAIDI, Professeur ordinaire et directeur de thèse (Département de Radiologie et Informatique), Monsieur Karsten KRUSE, Professeur ordinaire (Département de Biochimie), Monsieur Wesley BOLCH, Professeur (Université de Floride, Etats-Unis), Monsieur Manuel BARDIES, Professeur (INSERM,France), Monsieur Arman RAHMIM, Professeur (Université de la Colombie Britannique, Vancouver, Canada), Madame Yuni DEWARAJA, Professeure (Université du Michigan, Etats-Unis) autorisent l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 13 mars 2023

Thèse - 210 -

Le Doyen
Faculté de médecine

Le Doyen
Faculté des sciences

N.B. - La thèse doit porter la déclaration précédente et remplir les conditions énumérées dans les "Informations relatives aux thèses de doctorat à l'Université de Genève".

FACULTÉ DE MÉDECINE
FACULTÉ DES SCIENCES

UNIVERSITÉ DE GENÈVE

Summary

The use of ionizing radiation in medical imaging procedures, particularly in diagnostic radiology and nuclear medicine has significantly increased in the last decade, leading to major improvements in the diagnosis and treatment of various diseases. UNSCEAR indicated that over 3.6 billion x-ray examinations and 33 million nuclear medicine procedures are performed annually, worldwide. Although the use of ionizing radiation in medicine plays a pivotal role in healthcare, it is associated with risks of radiation-induced cancer. Therefore, its use is subject to standards of safety and stringent optimization procedures. Here, the term optimization refers to limiting the exposure of patients to the minimum necessary to achieve the required diagnostic or therapeutic objective, depending on the medical task at hand, the equipment being used, and patient-specific characteristics. To this end, the first step is the accurate assessment of radiation dose delivered to patients to assist the optimization of the given procedure.

The work presented in this dissertation and portfolio of published work aimed to establish an accurate and reliable methodology for monitoring and optimization of the radiation dose, initially from diagnostic imaging, but later in theranostic nuclear oncology. This research addressed two main questions: first, to establish a framework for habitus-specific and patient-specific dose monitoring and radiation dose reduction in hybrid PET/CT imaging; and second, to develop a practical dosimetry workflow to bring the full capacity of theragnostic dosimetry-guided planning to RadioPharmaceutical Therapy (RPT).

In the first phase of this thesis four works were carried out: i) a comprehensive library of computational phantom composed of 479 male/female phantoms were constructed representing anthropomorphic and anatomical diversities in adult population. ii) A framework for Monte Carlo (MC) based dose calculation from both internal (i.e. PET) and external (i.e. CT) exposure was developed and habitus-specific dose from a PET/CT examination was simulated. iii) The developed MC-based external dose simulator was benchmarked against experimental measurements and the dosimetric uncertainties associated with using different scan protocol parameters were investigated, iv) a methodology for construction of patient-specific computational model for dose estimation in radiological imaging was developed.

In the second phase, the application of deep learning in dose calculation and optimization was extended. Using the dosimetry toolkits and methods that have been developed in the first phase; we developed a novel deep learning-based algorithm for fast MC-based internal dosimetry. Patient-specific absorbed dose calculation using MC simulation is deemed the gold standard technique for internal dosimetry. However, this method is computationally expensive. Hence, we proposed a novel method to perform whole-body personalized voxel-level dosimetry taking into account the heterogeneity of activity distribution, non-uniformity of surrounding medium, and patient-specific anatomy using a physics-informed deep residual neural network. The proposed model was evaluated on the diagnostic ^{18}FDG-PET examinations showing a comparable accuracy with MC-based dose maps while being much faster in terms of execution time. This methodology is extendable to other diagnostic and therapeutic radionuclides. More specifically, we extended our model on ^{18}FDG to a betta- emitter therapeutic agent, ^{177}Lu-DOTATAE, using transfer learning that outperformed conventional methods, such as local energy deposition and MIRD formalism in high density regions (i.e. bone metastasis).

Furthermore, we applied the deep learning-based dose reconstruction methodology developed for internal dosimetry in nuclear medicine into high dose rate brachytherapy. To address the simplifications of TG-43 assumptions that ignore the dosimetric impact of medium heterogeneities, we proposed a neural network algorithm for reconstruction of MC-based dosimetry from CT images and radiation source characteristics. The results showed a comparable performance to the MC method while overcoming its computational burden and the inherent oversimplifications of TG-43 protocol.

Motivated by the dramatic increase of chest computed tomography exams following the global outbreak of COVID-19 in 2020, we designed an ultra-low-dose CT examination protocol for clinical diagnosis of COVID-19 patients using a deep neural network. In this work, we aimed to use deep learning algorithms on ultra-low-dose COVID-19 CT images to generate high-quality images with a comparable diagnostic accuracy to full-dose CT images. In our proposed protocol, the radiation dose in terms of CT dose index was reduced by up to 89% while the overall scoring to the predicted images assigned by radiologists showed an acceptance rate (score=4.72 ± 0.57) comparable to reference full-dose CT images (score=5).

Using the validated MC code that has already been developed for reconstruction of radiation dose associated with CT examinations, we developed a deep learning-based model to predict 3D voxel-level absorbed doses from anatomical density map and acquisition parameters. Through the generation of the dose map from a single source position, our model can generate accurate and personalized dose maps in few seconds for a wide range of acquisition parameters.

On the ground of previously developed personalized dosimetry methods, we further studied dosimetry in theragnostic in connection with targeted molecular radiotherapy. In recent years, the emergence of theragnosticas a single modality combining diagnostics and therapy, contributed to the current resurgence of interest in radiopharmaceutical therapy as a multidisciplinary endeavor. In pursuing personalized RPT based on patient-specific biology, tumor burden and dosimetry, a significant body of literature demonstrated the positive impact of dosimetry-guided dose planning on treatment efficacy.

Neuroendocrine tumors (NETs) with overexpressing somatostatin receptors provide the basis for peptide receptor radionuclide therapy (PRRT) through theragnostic pair of ^{68}Ga/^{177}Lu-DOTATATE. The main purpose of this study was to develop machine learning models to predict therapeutic tumor absorbed dose using pre-therapy ^{68}Ga-DOTATATE PET/CT and clinicopathological biomarkers. The patients included in this study underwent both pre-therapy ^{68}Ga-DOTATATE PET/CT and four time-points SPECT/CT at ~4, 24, 96 and 168 hours post ^{177}Lu-DOTATATE infusion. The preliminary results demonstrated the feasibility of using baseline PET images for estimating tumor absorbed dose prior to ^{177}Lu-PRRT to enable personalized treatment planning and patient stratification.

Résumé

L'utilisation des rayonnements ionisants dans les procédures d'imagerie médicale, notamment en radiologie et en médecine nucléaire, a considérablement augmenté au cours de la dernière décennie, ce qui a permis d'améliorer considérablement le diagnostic clinique et le traitement de diverses maladies. L'UNSCEAR indique que plus de 3,6 milliards d'examens radiologiques et 33 millions de procédures de médecine nucléaire sont réalisés chaque année dans le monde. Bien que l'utilisation des rayonnements ionisants en médecine joue un rôle essentiel dans le système de santé, elle est associée à des risques de cancer radio-induit. Par conséquent, son utilisation est soumise à des normes de sécurité et à des procédures d'optimisation rigoureuses. Le terme d'optimisation désigne ici la limitation de l'exposition des patients au minimum nécessaire pour atteindre l'objectif diagnostique ou thérapeutique requis, en fonction de la tâche médicale à accomplir, de l'équipement utilisé et des caractéristiques spécifiques du patient. À cette fin, la première étape est l'évaluation précise de la dose de rayonnement délivrée aux patients pour aider à l'optimisation de la procédure donnée.

Le travail présenté dans cette thèse et les travaux publiés visaient à établir une méthodologie précise et fiable pour le contrôle et l'optimisation de la dose de rayonnement, initialement en imagerie diagnostique, mais plus tard en médecine nucléaire thérapeutique ou théranostique. Cette recherche a abordée deux questions principales: premièrement, établir un cadre pour la surveillance de la dose spécifique à la morphologie et apparence du patient et la réduction de la dose de rayonnement en imagerie hybride TEP/TDM; et deuxièmement, développer un flux de travail pratique de dosimétrie pour apporter la pleine capacité de la planification guidée par la dosimétrie théranostique à la radiothérapie métabolique.

Dans la première phase de cette thèse, quatre travaux ont été réalisés : i) Une bibliothèque complète de fantômes computationnels composée de 479 fantômes masculins/féminins représentant les diversités anthropomorphiques et anatomiques de la population adulte a été construite. ii) Un cadre pour le calcul de la dose basé sur la méthode de Monte Carlo (MC) à partir de l'exposition interne (c'est-à-dire TEP) et externe (c'est-à-dire TDM) a été développé et la dose spécifique au patient lors d'un examen TEP/TDM a été simulée. iii) Le simulateur de dose externe basé sur la méthode de Monte Carlo a été évalué par rapport à des mesures expérimentales et et les incertitudes dosimétriques associées à l'utilisation de différents paramètres de protocoles d'imagerie ont été étudiées. iv) Enfin, une méthodologie pour la construction d'un modèle de calcul spécifique au patient pour l'estimation de la dose en imagerie radiologique a été développée.

Dans la deuxième phase de cette thèse, l'application de l'apprentissage profond dans le calcul et l'optimisation de la dose absorbée a été étendue. En utilisant les différents outils et méthodes de dosimétrie qui ont été développées dans la première phase, nous avons développé un nouvel algorithme basé sur l'apprentissage profond pour une dosimétrie interne rapide basée sur le modèle MC. Le calcul de la dose spécifique au patient par simulation MC est considéré comme la technique de référence en dosimétrie interne. Cependant, cette méthode est coûteuse en termes de temps de calcul. Nous avons donc proposé une nouvelle méthode pour effectuer une dosimétrie personnalisée voxelisée du corps entier en tenant compte de l'hétérogénéité de la distribution de l'activité, de la non-uniformité du milieu environnant et de l'anatomie spécifique du patient à l'aide d'un réseau neuronal résiduel profond. Le modèle proposé a été évalué sur les examens diagnostiques ^{18}FDG-PET montrant une précision comparable aux distributions de dose calculées par la méthode de MC tout en étant beaucoup plus rapide en termes de temps d'exécution. Cette méthodologie est extensible à d'autres radionucléides utilisés em diagnostic et en thérapie. Plus précisément, nous avons étendu notre modèle développé initialement pour le ^{18}FDG à un agent thérapeutique émetteur bêta, le ^{177}Lu-DOTATAE, en utilisant l'apprentissage par transfert qui a surpassé les méthodes conventionnelles, telles que le dépôt d'énergie local et le formalisme MIRD dans les régions à haute densité (c'est-à-dire les métastases osseuses).

De plus, nous avons appliqué à la curiethérapie à haut débit de dose la méthodologie de reconstruction de la dose basée sur l'apprentissage profond développée pour la dosimétrie interne en médecine nucléaire. Pour remédier aux simplifications des hypothèses TG-43 qui ignorent l'impact dosimétrique des hétérogénéités du milieu, nous avons proposé un algorithme de réseau neuronal pour la reconstruction de la dosimétrie basée sur le MC à partir d'images TDM et des caractéristiques de la source de rayonnement. Les résultats ont montré une performance

comparable à la méthode MC tout en surmontant sa charge de calcul et les simplifications excessives inhérentes au protocole TG-43.

Motivés par l'augmentation spectaculaire des examens tomodensitométriques thoraciques suite à l'épidémie mondiale de COVID-19 en 2020, nous avons conçu un protocole d'examen tomodensitométrique à ultra-faible dose pour le diagnostic clinique des patients COVID-19 en utilisant un réseau neuronal profond. Dans ce travail, nous avons voulu utiliser des algorithmes d'apprentissage profond sur des images TDM COVID-19 à ultra-faible dose pour générer des images de haute qualité avec une précision diagnostique comparable à celle des images TDM à pleine dose. Dans le protocole que nous avons proposé, la dose de rayonnement en termes d'indice de dose TDM a été réduite jusqu'à 89 %, tandis que la notation globale des images prédites attribuée par les radiologues a montré un taux d'acceptation (score = 4,72 ± 0,57) comparable à celui des images TDM pleine dose de référence (score = 5).

En utilisant le code MC validé qui a déjà été développé pour la reconstruction de la dose de radiation associée aux examens TDM, nous avons développé un modèle basé sur l'apprentissage profond pour prédire les doses absorbées au niveau du voxel 3D à partir de la carte de densité anatomique et des paramètres d'acquisition. Grâce à la génération de la distribution de dose à partir d'une seule position de la source, notre modèle peut générer des distributions de dose précises et personnalisées en quelques secondes pour une large gamme de paramètres d'acquisition.

Sur la base des méthodes de dosimétrie personnalisée précédemment développées, nous avons étudié la dosimétrie en théranostic en relation avec la radiothérapie moléculaire ciblée. Ces dernières années, l'émergence du théranostic en tant que modalité unique combinant diagnostic et thérapie a contribué à la résurgence actuelle de l'intérêt pour la radiothérapie moléculaire ciblée en tant qu'effort multidisciplinaire. Dans le cadre de la recherche d'une radiothérapie personnalisée basée sur la biologie, la charge tumorale et la dosimétrie spécifique au patient, un grand nombre de publications ont démontré l'impact positif de la planification de la dose guidée par la dosimétrie sur l'efficacité du traitement.

Les tumeurs neuroendocrines avec surexpression des récepteurs de la somatostatine constituent la base de la thérapie par radionucléides des récepteurs peptidiques (PRRT) grâce à la paire de radionucléides ^{68}Ga/^{177}Lu-DOTATATE. L'objectif principal de cette étude était de développer des modèles d'apprentissage automatique pour prédire la dose thérapeutique absorbée par la tumeur à l'aide de la TEP/TDM du ^{68}Ga-DOTATATE et de biomarqueurs clinicopathologiques avant la thérapie. Les patients inclus dans cette étude ont subi un TEP/TDM pré-thérapeutique au ^{68}Ga-DOTATATE et quatre points temporels TEMP/TDM à ~4, 24, 96 et 168 heures après la perfusion de ^{177}Lu-DOTATATE. Les résultats préliminaires ont démontré la faisabilité de l'utilisation d'images TEP de base pour estimer la dose absorbée par la tumeur avant le traitement par ^{177}Lu-PRRT afin de permettre une planification personnalisée du traitement et une stratification des patients.

To My Parents, My Husband, and My Twins

Papers included in this dissertation

This dissertation is based on a number of studies submitted and reported in the following peer-reviewed original articles. Published articles are reprinted by the kind permission of the referred publisher.

1. **Akhavanallaf A**, Fayad H, Salimi Y, Aly A, Kharita H, Al Naemi H and Zaidi H, An update on computational anthropomorphic anatomical models, Dig Health Vol. 8, pp 20552076221111941 (2022).
2. **Akhavanallaf A***, Xie T*, and Zaidi H, Development of a library of adult computational phantoms based on anthropometric indexes, IEEE Trans Rad Plasma Med Sci, Vol. 3, No. 1, pp 65-75 (2019).
3. Xie T* , **Akhavanallaf A***, and Zaidi H, Construction of patient-specific computational models for organ dose estimation in radiological imaging, Med Phys, Vol. 46, No.4, pp 1519-1522 (2019).
4. **Akhavanallaf A**, Xie T, and Zaidi H, Assessment of uncertainties associated with Monte Carlo-based personalized dosimetry in clinical CT examinations, Phys Med Biol, Vol. 65, No. 4, pp 045008 (2020).
5. Arabi H, **Akhavanallaf A**, Shiri I, Sanaat A, Zaidi H, The promise of artificial intelligence and deep learning in PET and SPECT imaging, Phys Med, Vol. 83, pp 122-137 (2021).
6. Shiri I*, **Akhavanallaf A***, Sanaat A, et al, Ultra-low dose chest CT imaging of Covid-19 patients using deep neural networks, Eur Radiol, 31(3), 1420–1431 (2021).
7. Salimi Y, **Akhavanallaf A**, Mansouri Y, Shiri I and Zaidi H, Real-time, acquisition parameter-free, voxelwise patient-specific Monte Carlo dose reconstruction in total-body CT scanning using deep neural networks, Eur Radiol (2022) *under revision.*
8. **Akhavanallaf A**, Shiri I, Arabi H, and Zaidi H, Whole-body voxel-based internal dosimetry using deep learning" Eur J Nucl Med Mol Imaging Vol. 48, No. 3, pp 670-682 (2021).
9. **Akhavanallaf A**, Mohammadi R, Shiri I, Salimi Y, Arabi H, and Zaidi H, Personalized brachytherapy dose reconstruction using deep learning, Comput Biol Med, Vol. 136, pp 104755 (2021).
10. **Akhavanallaf A,** Peterson A, Fitzpatrick K, Roseland M, Wong KK, El Naqa I, Zaidi H and Dewaraja Y. The predictive value of ^{68}Ga-PET and baseline biomarkers in ^{177}Lu-DOTATATE tumor dosimetry, Eur J Nucl Med Mol Imaging (2023) *under revision.*

*Authors contributed equally

List of additional papers

The following papers are not included in the thesis.

1. Riveira M, **Akhavanallaf A**, Bianchetto Wolf N, Salimi Y, Mainta I, Garibotto V, Lopez Medina AM and Zaidi H "Predictive Value of 99mTc-MAA-based Theragnostic Dosimetry in Personalized 90Y-SIRT planning of Liver Malignancies" Eur J Nucl Med Mol Imaging (2023) *submitted.*
2. Yixuan Jia, Zongyu Li, **Azadeh Akhavanallaf**, Jeffrey A. Fessler, Yuni K. Dewaraja, 90Y SPECT Scatter Estimation and Voxel Dosimetry in radioembolization Using a Unified Deep Learning Framework" EJNMMI Phys (2023) *submitted.*
3. Shiri I, Sadr AV, **Akhavanallaf A**, Salimi Y, Sanaat A, Amini M, Razeghi B, Saberi A, Arabi H, Ferdowsi S, Voloshynovskiy S, Gündüz D, Rahmim A and Zaidi H "Decentralized collaborative multi-institutional PET attenuation and scatter correction using federated deep learning" Eur J Nucl Med Mol Imaging (2023) *in press.*
4. Salimi Y, Shiri I, **Akhavanallaf A**, Mansouri Z, Arabi H and Zaidi H "Fully automated accurate patient positioning in computed tomography using anterior-posterior localizer images and a deep neural network: A dual-center study" Eur Radiol (2023) *in press.*
5. Asadi A, **Akhavanallaf A**, Hossein SA, Vosoughi N and Zaidi H "Development and validation of an optimal GATE model for proton pencil-beam scanning delivery" Z Med Phys (2023) *in press.*
6. Salimi Y, Shiri I, **Akhavanallaf A**, Mansouri Z, Sanaat A, Pakbin M, Hossein MG, Arabi H, and Zaidi H "Deep learning-based calculation of patient size and attenuation surrogates from localizer image: Toward personalized CT protocol optimization for thorax imaging" Eur J Radiol Vol. 157, pp 110602 (2022).
7. Asadi A, Hossein SA, **Akhavanallaf A**, Vosoughi N and Zaidi H "Comparative study of passive scattering and active scanning proton therapy techniques using Monte Carlo simulations" J Instrum Vol. 17, No. 9, pp P09008 (2022).
8. Sanaat A, **Akhavanallaf A**, Shiri I, Salimi Y, Arabi H and Zaidi H "Deep-TOF-PET: Deep learning-guided generation of time-of-flight (TOF) PET from non-TOF brain PET images in the image and projection domains" Hum Brain Mapp Vol. 43, No. 16, pp 5032-5043 (2022).
9. Shiri I, Mostafaei S, Avval AH, Salimi Y, Sanaat Y, **Akhavanallaf A**, Arabi H, Rahmim A and Zaidi H "High-dimensional multinomial multiclass severity scoring of COVID-19 pneumonia using CT radiomics features and machine learning algorithms" Sci Rep Vol. 12, No. 1, pp 14817 (2022).
10. Shiri I, Sadr AV, Amini M, Salimi Y, Sanaat A, **Akhavanallaf A**, Razeghi B, Ferdowsi S, Saberi A, Arabi H, Becker M, Voloshynovskiy S, Gündüz D, Rahmim A and Zaidi H, "Decentralized distributed multi-institutional PET image segmentation using a federated deep learning framework" Clin Nucl Med Vol. 47, No 7, pp 606-617 (2022).
11. Shiri I, Arabi H, Salimi Y, Sanaat A, **Akhavanallaf A**, Hajianfar G, Askari D, Moradi S, Mansouri Z, Pakbin M, Sandoughdaran S, Abdollahi H, Radmard AR, Rezaei-Kalantari K, Oghli MG, and Zaidi H"COLI-NET: Fully automated COVID-19 lung and infection pneumonia lesion detection and segmentation from chest CT images" Int J Imaging Syst Technol Vol. 32, No. 1, pp 12-25 (2022).
12. Salimi Y, Shiri I, **Akhavanallaf A**, Mansouri Z, Saberi A, Sanaat A, Pakbin M, Askari D, Sandoughdaran S, Sharifipour E, Arabi H, and Zaidi H "Deep learning-based fully automated Z-axis coverage range definition from scout scans to eliminate overscanning in chest CT scanning" Insights Imaging Vol. 12, No. 1, pp 162 (2021).

Contents

Chapter 1

Introduction

I. Motivation and background

Medical applications of ionizing radiation have substantially increased over the past three decades, leading to major improvements in the diagnosis and treatment of human diseases. While the patients gain recognized benefits from radiation-involved health technologies in medicine, they may also be exposed to unnecessary or unintended irradiations and be subject to potential health hazards in inappropriate applications. Therefore, it is essential to assess potential risks for a better management of safety and quality in the use of radiological medical equipment and protection of patients, health workers and members of the public. Since the beginning of radiation therapy, a lot of efforts have focused on the optimization of curative principles for maximization of absorbed dose to tumors while sparing healthy tissues. Although high radiation doses are known to be correlated with increased risks of cancer (deterministic effects), the effects of low doses, typically encountered in diagnostic medical imaging, are still a matter of debate (stochastic effects). There is some evidence indicating the association of increased cancer risks with diagnostic low-dose radiation exposure that raise concerns about potential carcinogenic harms associated with these examinations [1-4]. However, some research studies questioned the hypothesis behind the uncertain risk estimation for low dose exposure delivered by medical procedure (<100 mSv) [5]. Thus, it is considered prudent for public safety to assume that exposure to ionizing radiation, no matter how small it is, carries small risk of unwanted health effects, including cancer. In this context, the assessment of radiation dose to patients, as a metric for quantification of either deterministic or stochastic risks, plays a critical role in the development, clinical use, and optimization of radiation-involved procedures in both diagnostic and therapeutic applications. Here, the concept of optimization includes assessment of the risks and benefits of a procedure guiding the amendment of the protocols. To this end, accurate radiation dosimetry provides the basis for quantification of risk factors. While more accurate risk quantification leads to a better understanding of risk-benefit association depending on the medical task at hand, the equipment being used, and patient-specific characteristics. Therefore, the framework of current dosimetry methods plays a key role in the advancement of both safety and efficacy of medical exposures.

Precision medicine is a new paradigm aiming at improving healthcare while lowering the costs, thus offering great potential for patient-specific optimal treatment strategies [6]. In the era of precision medicine, it is time to shift from one-size-fits-all paradigm to personalized approach for dosimetry calculation that is essential for both diagnostic and therapeutic nuclear medicine procedures.

In recent years, the demand for radiopharmaceutical therapy (RPT) has grown rapidly in clinical oncology owing to its efficacy and specific theragnostic features. The introduction of theragnostics, as a single modality combining diagnosis and therapy in RPT, enables not only for therapy verification but also for personalized dosimetry-guided treatment planning [7]. Currently, RPT is administered intravenously or locoregionally, where planning is typically designed based on either a fixed dosage or patient's body weight/surface area, analogous to chemotherapy. This approach is very much in contrast with external beam radiotherapy where personalized planning is integrated in clinical routine. Considering advances in quantitative molecular imaging technologies, computational modeling of the human body, and sophisticated radiation transport techniques, it is time to shift from the standard fixed dose regimen to a personalized approach in RPT. There is mounting evidence that personalized dosimetry enables for more rigorous approach in treatment planning aiming at improving therapeutic outcome and for a better understanding of the dose–response association [8-11].

With respect to radiotracer imaging, i.e. single-photon emission computed tomography (SPECT) and positron emission tomography (PET), mostly implemented as hybrid imaging modalities integrated with computed tomography (CT), personalized dosimetry is recommended for optimizing clinical procedures while minimizing the risks of radiation-induced cancer. The International Commission on Radiological Protection (ICRP) suggested estimating the radiation dose delivered to patients from medical imaging procedures toward the optimization rule known as ALARA (As Low As Reasonably Achievable) in order to minimize the risks through the appropriate use of ionizing radiation [12]. Currently, CT scanning accounts for the primary source of medical radiation exposure to the population. Despite the technological innovations devised to optimize the radiation dose associated with CT. Yet, "*CT is still not a low-dose imaging modality*" [13]. Hence, personalized dosimetry is of critical importance for effective analysis of the risk-benefits of medical imaging and the design of radiation dose optimization strategies according to ALARA principle.

In the context of personalized dose calculation, three main components are required:

1. The definition of radiation sources (external or internal);
2. Patient-specific computerized representations of the human body, referred to as computational models.
3. Radiation transport algorithm and scoring of energy depositions.

The definition of radiation source considers the characteristics of the emitted radiation, including the type of particle (e.g. photons, electrons, positrons, alpha particles, etc), particle energy and the geometrical distribution of the source. Computational models are digitized representations of the human anatomy that have been introduced to represent the spatial distribution of the different tissues in the body. These phantoms were developed originally for applications in radiation protection and medical imaging instrumentation and protocols optimization. Computational phantoms have been extended from simple water-filled slabs and spheres to patient-specific anthropomorphic models with a realistic detailed anatomy and material compositions. The principle of absorbed dose calculation involves solving the Boltzmann transport equation which describes the interaction of radiation particles travelling through a medium. There are two main categories of approaches for solving this equation: deterministic and stochastic algorithms. The Monte Carlo (MC) technique has been introduced as a stochastic solution for the Boltzmann transport equation through probabilistic simulation of particle interactions with matter. The main advantage of the MC method compared to the deterministic approach is its capability to simulate complex problems allowing the definition of either intricate geometries (i.e. complex heterogeneous medium) or source models. Direct MC simulations are considered the gold standard for implementation of a reliable dosimetry framework. However, this technique is computationaly intensive and requires significant expertise in computer programming [14].

II. Scope and outline of the research

According to the structure of this dissertation, illustrated in Figure 1, we conducted multiple studies in two distinct parts: the first part focused on advanced methods for dosimetry in diagnostic nuclear medicine imaging centering mainly around hybrid PET/CT systems, whereas the second part focused on radiation dosimetry in theragnostic applications.

PART 1: diagnostic nuclear medicine dosimetry

In current clinical radiopharmaceutical dosimetry, patient absorbed dose monitoring is commonly based on simplified models, such as the Medical Internal Radiation Dose Committee (MIRD) formalism [15]. The traditional MIRD technique is based on organ-level dosimetry using time-integrated activities and radionuclide-specific S-values, which represents the mean absorbed dose to a target organ per radioactive decay in a source organ. These quantitative parameters are modeled based on a reference computational phantom. This approach assumes a uniform activity distribution within each organ and ignores individual anatomical characteristics. OLINDA/EXM® software, released at 2004, was the first computer code based on organ-level MIRD formalism for radiopharmaceutical absorbed dose calculation (OLINDA/EXM stands for Organ Level INternal Dose Assessment/EXponential Modeling). It calculates radiation doses to different organs of the body from systemically administered radiopharmaceuticals and performs regression analysis on user-supplied biokinetic data to support such calculations for nuclear medicine pharmaceutical [16].

To cope with inter-subject variability of anatomical features and heterogeneity of radiotracer distribution within the source volume, voxel-based dosimetry techniques have been developed, using either dose point kernels [17, 18] or voxel S-values (VSV) [15] approaches. Dose point kernel is a deterministic approach which calculates the radial absorbed dose distribution around an isotropic point source in a homogeneous water medium [19, 20]. The voxel-level MIRD schema is defined as a 3D voxel matrix representing the mean absorbed dose to a target voxel per unit activity in the central source voxel embedded in an infinite homogeneous medium using MC simulations. However, voxel-based dose calculation should in principle take into account non-uniform activity distribution of the radiotracer, the heterogeneity of the medium density consisting of different tissue compositions, e.g., lung, soft tissue, and bone, in radiation transport calculation

is ignored. Therefore, MC-based radiation transport algorithms that are capable to simulate either a heterogenous activity distribution or a complex inhomogenous medium is considered as the gold standard, in particular for dosimetry in heterogenous medium, such as lung and bone lesions and marrow dosimetry and also for imaging radiotracers with longer ranges.

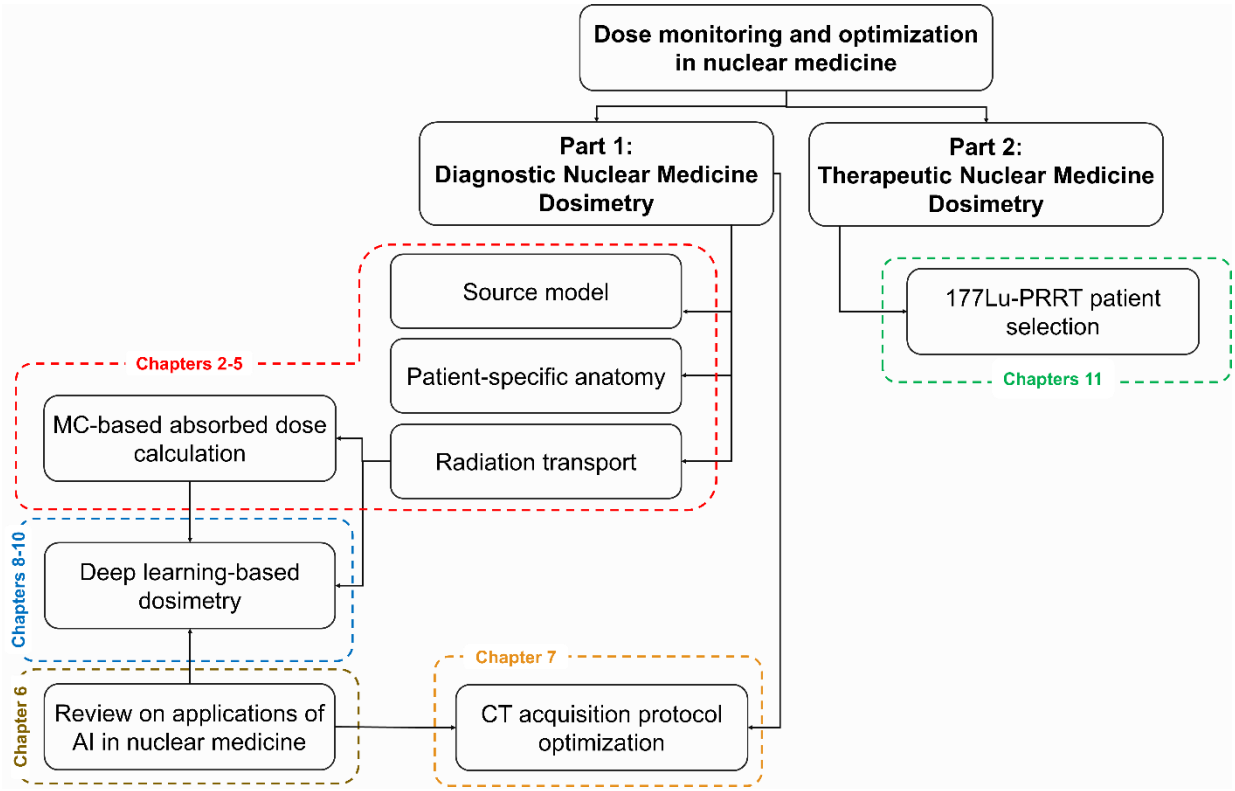


Figure 1. Outline of the Thesis.

Computed tomography (CT) is a widely deployed imaging modality in clinical setting. Its natural growth and widespread adoption, inevitably became a public health concern raising awareness about patients' exposure to ionizing radiation (UNSCEAR [21] reported ~ 3·6 billion diagnostic x-ray examinations, annualy). Patient radiation dose associated with CT examinations is typically reported using exposure indices, such as CT dose index (CTDI) and dose-length product (DLP). These metrics represent the radiation output of a specific examination that are useful for quality assurance but do not provide any information on tissue absorbed dose. The estimation of absorbed doses can be performed using various methodologies. The most straightforward approach uses conversion factors specific to the scanning protocols, such as size-specific dose estimate (SSDE). An alternative option is to use dedicated software tools, such as ImpactDose[2] and Radimetrics [22], that have been developed based on MC calculations using a general x-ray tube as source model and reference computational phantoms representing population-based patient anatomy. Both above mentioned software packages proved to have some discrepancies with organ absorbed doses calculated through patient-specific MC simulations [23-25]. MC calculations using patient-specific computational models is deemed to be the most accurate approach in CT dosimetry while its downsides, including computational burden, and required expertise in computer programming, limit its adoption in clinical setting.

In the context of personalized dosimetry, patient-specific computational model is the key component enabling accurate estimation of organ-level dose. However, detailed segmentation of patient images to construct personalized models was not feasible for clinical routine applications till recently, where the emergence of deep learning in the computer vision domain revolutionized the area of patient-specific computational modelling. In this regard, to cope with inter-subject variability of anatomical features, a potential

[2] https://impactdose.software.informer.com/

alternative for person-specific organ absorbed dose estimation was introduced through developing a library of computational models where habitus-specific phantoms could serve as alternative models covering various anthropometric and anatomical characteristics of patients [26]. Several habitus-dependent phantom series have been developed to perform patient-specific absorbed dose estimation by matching anthropometric characteristics of patients, such as gender, age, height, and body weight [27, 28]. The construction of patient-specific models from regional CT images is another alternative for patient-specific dosimetry, which was adopted in a number of studies and also within Radimetrics™ commercial dose tracking software (Bayer HealthCare, Berlin, Germany) by mapping the segmented model of patient CT images to a template anatomy through a simple protocol-based registration of CT scan localizer to predefined anatomical landmarks on Cristy & Eckerman stylized phantom [29-31].

Since Monte Carlo calculations using patient-specific models are commonly considered as gold standard for organ absorbed dose estimation from diagnostic imaging procedures, the implementation of an easy to use and reliable framework enabling to estimate patient-specific organ absorbed doses for individual patients in clinical setting is highly desirable. This was the focus of Part 1 of this dissertation.

Focusing on public health aspects related to the risks of the use of radiation in medicine, this part considered hybrid PET/CT scanner as example to establish a framework for personalized dosimetry in molecular imaging. Patients undergoing PET/CT scanning is exposure to two different radiation sources: the internal radionuclide source and the external X-ray source. This part includes components linked to patient representation, radiation dose calculation, radiation dose optimization and application of artificial intelligence (AI) in fast MC-based dosimetry. The studies conducted in Part 1 were divided in two sub-categories:

A. Develop and validate Monte Carlo simulation code system for personalized dosimetry of a clinical PET/CT examination.

1. Developing a comprehensive library of anthropomorphic computational models representing the main anatomical characteristics of the majority of patients at different ages (both genders), and different anthropomorphic characteristics.
2. Developing a MC simulation code system for clinical PET/CT scanners to incorporate the geometry of the CT component and the biokinetic data of fluorodeoxyglucose (^{18}F-FDG) that is extendable to other radiotracers.
3. Developing a unified methodology for constructing patient-pecific computational models from CT regional images based on deformable registration algorithms for organ absorbed dose estimation in radiological imaging.
4. Assessment of uncertainties associated with MC-based personalized dosimetry in clinical CT examinations, in comparison with experimental measurements.

B. Applications of artificial intelligence, in particular deep learning algorithms, into MC-based personalized dosimetry. The tools, techniques and data developed in part A, established the foundations for application of AI in radiation dosimetry.

1. Implementation of deep learning for optimization of acquisition protocols in CT scanning to reduce patient exposure while recovering image quality in the chest region.
2. Implementation of deep learning in external dosimetry for real-time, acquisition parameter-free, patient-specific MC dose reconstruction in CT examinations.
3. Implementation of a fast MC-based dose reconstruction in whole-body internal dosimetry using a physics-informed deep learning algorithm.
4. Extension of deep learning-based fast MC dose simulation in personalized brachytherapy dose reconstruction.

PART 2: Therapeutic nuclear medicine dosimetry

Radiopharmaceutical therapy involves a biochemical pathway to deliver cytotoxic levels of radiation to a specific cancer cell type through targeting agents while minimizing damage to normal cells. In recent years, the progress

in radiopharmacology with the emergence of theragnostics, combining diagnostic and therapy, contributed to the current resurgence of interests in RPT as a multidisciplinary endeavor [32]. The promising clinical outcome of RPT for the treatment of both local cancers and metastatic malignancies is being approved as an effective treatment technique with a good safety profile, while economically and logistically viable [33]. Currently, RPT planning is typically based on either a fixed dosage or patient's body weight/surface area, analogous to chemotherapy. This approach is very much in contrast with the new EU directive 2013/59/EURATOM Article 56, indicating that "*For all medical exposure of patients for radiotherapeutic purposes, exposures of target volumes shall be individually planned and their delivery appropriately verified taking into account that doses to non-target volumes and tissues shall be as low as reasonably achievable and consistent with the intended radiotherapeutic purpose of the exposure*". In Chapter II of the above directive, Definitions (Article 5, Definition 81), it is further stated that "'*radiotherapeutic' means pertaining to radiotherapy, including nuclear medicine for therapeutic purposes*" [34, 35].

Pursuing the concept of precision medicine in RPT, we developed a theragnostic dosimetry workflows for ^{177}Lu-labeled radiopharmaceutical were developed based on patient-specific biology, molecular pathology, and dosimetry. Lu-DOTATATE has been recently approved by the FDA for the treatment of somatostatin receptor-positive neuroendocrine tumors (NETs) [36]. While the currently recommended dosage is a fixed administered regimen (7.4 GBq/fraction), there is mounting evidence that dosimetry-guided adaptation of administered activity may be beneficial to optimize therapy outcome. The preliminary results indicate wide margins for increasing the administered activity with tolerable toxicity [9, 37].

In Part 2 we focused on the capacity of theragnostic dosimetry-based planning to RPT setting by developing a practical dosimetry workflow for ^{177}lu-Labeled peptide receptor radionuclide therapy (PRRT); and furthere to develop a decision-support model for ^{177}Lu-DOTATATE RPT for patient stratification based on quantitative biomarkers from pre-therapy imaging and pathology to ensure optimal efficacy and economy.

References

[1] M. Hauptmann *et al.*, "Brain Cancer after Radiation Exposure from Ct Examinations of Children and Young Adults: Results from the Epi-Ct Cohort Study,"*Lancet Oncol,* vol. 24, no. 1, pp. 45-53, 2023.

[2] J.-Y. Hong, K. Han, J.-H. Jung, and J. S. Kim, "Association of Exposure to Diagnostic Low-Dose Ionizing Radiation with Risk of Cancer among Youths in South Korea,"*JAMA Network Open,* vol. 2, no. 9, pp. e1910584-e1910584, 2019.

[3] K. H. Lee *et al.*, "Risk of Hematologic Malignant Neoplasms from Abdominopelvic Computed Tomographic Radiation in Patients Who Underwent Appendectomy,"*JAMA Surgery,* vol. 156, no. 4, pp. 343-351, 2021.

[4] N. Zewde, F. Ria, and M. M. Rehani, "Organ Doses and Cancer Risk Assessment in Patients Exposed to High Doses from Recurrent Ct Exams,"*Eur J Radiol,* vol. 149, p. 110224, 2022.

[5] W. R. Hendee and M. K. O'Connor, "Radiation Risks of Medical Imaging: Separating Fact from Fantasy,"*Radiology,* vol. 264, no. 2, pp. 312-21, 2012.

[6] M. Baumann *et al.*, "Radiation Oncology in the Era of Precision Medicine,"*Nat Rev Cancer,* vol. 16, no. 4, pp. 234-49, 2016.

[7] S. A. Graves and R. F. Hobbs, "Dosimetry for Optimized, Personalized Radiopharmaceutical Therapy,"*Seminars in Radiation Oncology,* vol. 31, no. 1, pp. 37-44, 2021.

[8] G. Sgouros *et al.*, "Tumor Response to Radiopharmaceutical Therapies: The Knowns and the Unknowns,"*J Nucl Med,* vol. 62, no. Suppl 3, pp. 12s-22s, 2021.

[9] M. Del Prete *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,* vol. 46, no. 3, pp. 728-742, 2019.

[10] A. Sundlöv *et al.*, "Individualised (177)Lu-Dotatate Treatment of Neuroendocrine Tumours Based on Kidney Dosimetry,"*Eur J Nucl Med Mol Imaging,* vol. 44, no. 9, pp. 1480-1489, 2017.

[11] E. Garin *et al.*, "Personalised Versus Standard Dosimetry Approach of Selective Internal Radiation Therapy in Patients with Locally Advanced Hepatocellular Carcinoma (Dosisphere-01): A Randomised, Multicentre, Open-Label Phase 2 Trial,"*Lancet Gastroenterol Hepatol,* vol. 6, no. 1, pp. 17-29, 2021.

[12] H. G. Menzel, C. Clement, and P. DeLuca, "Icrp Publication 110. Realistic Reference Phantoms: An Icrp/Icru Joint Effort. A Report of Adult Reference Computational Phantoms,"*Ann ICRP,* vol. 39, no. 2, pp. 1-164, 2009.

[13] M. M. Rehani, T. P. Szczykutowicz, and H. Zaidi, "Ct Is Still Not a Low-Dose Imaging Modality,"*Med Phys,* vol. 47, no. 2, pp. 293-296, 2020.

[14] H. Zaidi, "Relevance of Accurate Monte Carlo Modeling in Nuclear Medical Imaging,"*Med Phys,* vol. 26, no. 4, pp. 574-608, 1999.

[15] W. E. Bolch *et al.*, "Mird Pamphlet No. 17: The Dosimetry of Nonuniform Activity Distributions--Radionuclide S Values at the Voxel Level. Medical Internal Radiation Dose Committee,"*J Nucl Med,* vol. 40, no. 1, pp. 11s-36s, 1999.

[16] M. G. Stabin, R. B. Sparks, and E. Crowe, "Olinda/Exm: The Second-Generation Personal Computer Software for Internal Dose Assessment in Nuclear Medicine,"*J Nucl Med,* vol. 46, no. 6, pp. 1023-7, 2005.

[17] M. J. Berger, "Distribution of Absorbed Dose around Point Sources of Electrons and Beta Particles in Water and Other Media," National Bureau of Standards, Washington, DC, 1971.

[18] M. Bardiès and L. Strigari, "Dose Point-Kernels for Radionuclide Dosimetry," in *Monte Carlo Calculations in Nuclear Medicine (Second Edition)*, H. Zaidi Ed.: Institute of Physics Publishing, Bristol, 2022.

[19] K. S. Kolbert *et al.*, "Implementation and Evaluation of Patient-Specific Three-Dimensional Internal Dosimetry,"*J Nucl Med,* vol. 38, no. 2, pp. 301-8, 1997.

[20] H. B. Giap, D. J. Macey, J. E. Bayouth, and A. L. Boyer, "Validation of a Dose-Point Kernel Convolution Technique for Internal Dosimetry,"*Phys Med Biol,* vol. 40, no. 3, pp. 365-81, 1995.

[21] U. N. S. C. o. t. E. o. A. Radiation, *Sources and Effects of Ionizing Radiation, United Nations Scientific Committee on the Effects of Atomic Radiation (Unscear) 2008 Report, Volume I: Report to the General Assembly, with Scientific Annexes a and B-Sources*. United Nations, 2010.

[22] Bayer HealthCare. "Radimetrics Enterprise Platform: Dose Management Solution." Bayer HealthCare. http://www.radiologysolutions.bayer.com/products/ct/dosemanagement/rep/. (accessed 2016).

[23] N. Guberina *et al.*, "Verification of Organ Doses Calculated by a Dose Monitoring Software Tool Based on Monte Carlo Simulation in Thoracic Ct Protocols,"*Acta Radiol,* vol. 59, no. 3, pp. 322-326, 2018.

[24] R. Al-Senan, K. Brown, M. Erdman, and S. King, "The Uncertainty of Thyroid Dose Estimate in Chest Ct,"*Biomed Phys Eng Express,* vol. 6, no. 6, 2020.

[25] M. Tahiri, M. Mkimel, Y. Benameur, R. El Baydaoui, M. R. Mesradi, and O. El Rhazouani, "Organ Dose Estimation for Adult Chest Ct Examination Using Gate Monte Carlo Simulation,"*Physics of Particles and Nuclei Letters,* vol. 18, no. 4, pp. 502-509, 2021.

[26] B. M. Moore, S. L. Brady, A. E. Mirro, and R. A. Kaufman, "Size-Specific Dose Estimate (Ssde) Provides a Simple Method to Calculate Organ Dose for Pediatric Ct Examinations,"*Med Phys,* vol. 41, no. 7, p. 071917, 2014.

[27] Y. H. Na, B. Zhang, J. Zhang, P. F. Caracappa, and X. G. Xu, "Deformable Adult Human Phantoms for Radiation Protection Dosimetry: Anthropometric Data Representing Size Distributions of Adult Worker Populations and Software Algorithms,"*Phys Med Biol,* vol. 55, no. 13, pp. 3789-811, 2010.

[28] P. B. Johnson, A. Geyer, D. Borrego, K. Ficarrotta, K. Johnson, and W. E. Bolch, "The Impact of Anthropometric Patient-Phantom Matching on Organ Dose: A Hybrid Phantom Study for Fluoroscopy Guided Interventions,"*Med Phys,* vol. 38, no. 2, pp. 1008-17, 2011.

[29] P. Sahbaee, W. P. Segars, and E. Samei, "Patient-Based Estimation of Organ Dose for a Population of 58 Adult Patients across 13 Protocol Categories,"*Med Phys,* vol. 41, no. 7, p. 072104, 2014.

[30] W. A. Kalender, N. Saltybaeva, D. Kolditz, M. Hupfer, M. Beister, and B. Schmidt, "Generating and Using Patient-Specific Whole-Body Models for Organ Dose Estimates in Ct with Increased Accuracy: Feasibility and Validation,"*Phys Med,* vol. 30, no. 8, pp. 925-33, 2014.

[31] M. Cristy and K. Eckerman, "Specific Absorbed Fractions of Energy at Various Ages from Internal Photon Sources: 1, Methods," Oak Ridge National Lab., TN (USA), 1987.

[32] J. H. Turner, "Recent Advances in Theranostics and Challenges for the Future,"*Br J Radiol,* vol. 91, no. 1091, p. 20170893, 2018.

[33] G. Sgouros, L. Bodei, M. R. McDevitt, and J. R. Nedrow, "Radiopharmaceutical Therapy in Cancer: Clinical Advances and Challenges,"*Nat Rev Drug Discov,* vol. 19, no. 9, pp. 589-608, 2020.

[34] M. Ljungberg and K. Sjögreen Gleisner, "Personalized Dosimetry for Radionuclide Therapy Using Molecular Imaging Tools,"*Biomedicines,* vol. 4, no. 4, 2016.

[35] C. Stokke *et al.*, "Dosimetry-Based Treatment Planning for Molecular Radiotherapy: A Summary of the 2017 Report from the Internal Dosimetry Task Force,"*EJNMMI Phys,* vol. 4, no. 1, p. 27, 2017.

[36] U. Hennrich and K. Kopka, "Lutathera(®): The First Fda- and Ema-Approved Radiopharmaceutical for Peptide Receptor Radionuclide Therapy,"*Pharmaceuticals (Basel),* vol. 12, no. 3, 2019.

[37] A. Sundlöv *et al.*, "Phase Ii Trial Demonstrates the Efficacy and Safety of Individualized, Dosimetry-Based (177)Lu-Dotatate Treatment of Net Patients,"*Eur J Nucl Med Mol Imaging,* vol. 49, no. 11, pp. 3830-3840, 2022.

Chapter 2

An update on computational anthropomorphic anatomical models

Azadeh Akhavanallaf, Hadi Fayad, Yazdan Salimi, Antar Aly, Hassan Kharita, Hoda Al Naemi and Habib Zaidi

Contribution: Study conception and design, literature review, analysis and interpretation of results, manuscript preparation.

Dig Health Vol. 8, pp 20552076221111941 (2022)

Abstract

The prevalent availability of high-performance computing coupled with validated computerized simulation platforms as open-source packages have motivated progress in the development of realistic anthropomorphic computational models of the human anatomy. The main application of these advanced tools focused on imaging physics and computational internal/external radiation dosimetry research. This paper provides an updated review of state-of-the-art developments and recent advances in the design of sophisticated computational models of the human anatomy with a particular focus on their use in radiation dosimetry calculations. The consolidation of flexible and realistic computational models with biological data and accurate radiation transport modeling tools enables the capability to produce dosimetric data reflecting actual setup in clinical setting. These simulation methodologies and results are helpful resources for the medical physics and medical imaging communities and are expected to impact the fields of medical imaging and dosimetry calculations profoundly.

I. Introduction

Human anatomical models have been developed to represent the spatial distribution of different tissues in the body (human anatomy) [1, 2]. These models were mainly constructed to provide a non-invasive and inexpensive way to test various diagnostic imaging and interventional/therapeutic procedures [3], such as dosimetric calculations for ionizing/non-ionizing radiation exposure, optimizing medical imaging facilities, and personalized medicine. An important category of anatomical models is physical phantoms composed of solid materials with properties equivalent to human tissues, e.g., homogenized cylindrical water phantoms used for the calibration of radiation detectors and diagnostic imaging systems [4, 5]. However, these phantoms are usually expensive while reflecting a crude approximation of the human body. In addition, using physical phantoms for the calibration of advanced systems can be very costly and time-consuming. As a result, computational phantoms representing a mathematical model of the human anatomy in a digital format were developed originally for applications in radiation protection and medical imaging optimization. Recently, the ultimate objective of constructing human computational models as the ancestor of the digital twins (i.e. computational objects employed in medicine or other fields as surrogate or replica of the human body to certain process, e.g., to ionizing radiation) is the personalization of medical procedures within the paradigm of precision medicine [6]. Starting in the 1960s, the development of the computational models evolved through many generations, and in the 1980s, further efforts were made in this domain. In this regard, the Visible Human Project led to the creation of the first complete anatomical model for dose calculation purposes [7]. The first generation of computational phantoms suffered from a variety of limitations, including lack of anatomical realism, the non-inclusion of tissue characteristics, calculation speed, as well as their incompatibility with available analytical or Monte Carlo simulation codes. More importantly, these computational phantoms have not been designed for subject-specific modeling and ignored inter-subject anatomical variability [8].

Advances in high-performance computing stimulated the development and usage in research of realistic computational anthropomorphic models. To date, more than 200 computational phantoms have been reported in the literature [2]. Examples of widely used anatomical models are the NURBS-based XCAT phantom series [9] and the Virtual Population based on triangle mesh [10, 11]. Advances in medical imaging modalities and computational algorithms allow fast and construction of personalized computational models through automated segmentation techniques and enable the incorporation of physiological motion into anatomical models.

II. The fundamentals of computational models design

Multiple factors should be considered during the construction of a realistic anthropomorphic anatomical model [12]. These include anatomy (tissues, organs and regions), tissue properties, computational efficiency as well as compatibility with analytical or Monte Carlo simulation codes.

As a result, the first step in anthropomorphic anatomical model construction consists of defining geometrical surfaces and tissue properties. This can be done either by using constructive solid geometry (CSG) or boundary representation (BREP) approaches [13, 14]. In CSG, objects are created using primitives, such as cylinders, ellipsoids, spheres, … etc. A number of examples can be found in the literature that fall under this category ranging from whole organ representation [8] to voxel-based representations [15]. Although whole organ representation approaches have the advantage that they are computationally efficient and compatible with existing Monte Carlo radiation transport simulation codes, they suffer from the lack of anatomical realism. Conversely, voxel-based representation has the advantage of reflecting anatomical realism that can be integrated into simulation codes [15]. However, the geometric fidelity is dependent on voxel size, and the simulations are computationally inefficient, especially for organ shapes readjustment. In BREP modeling, tissues can be characterized using boundary surfaces, such as non-uniform rational B-splines (NURBS) or polygon mesh surfaces. As in the case of voxel-based CSG, the data can be extracted from CT images by contouring organ surfaces followed by modeling to end up with smooth and continuous boundaries. The BREP representation better reflects anatomical realism compared to CSG, given that it can model complex anatomical features using an extended set of operation tools.

Although BREP models provide improved realism compared to previous modeling techniques, the corresponding models are still static. Therefore, a number of additional parameters have to be included to mimic the reality. There are many reasons behind this as summarized by Neufeld et al. [12]:

- Slow changes in anatomy during treatment in radiotherapy,
- Inter-subject anatomical variability in anthropomorphic parameters, such as height, weight, age, BMI, etc.
- Need to have personalized models reflecting specific patient‘s anatomy/physiology.
- Voluntary or involuntary motion of organs, such as respiratory/cardiac motion or bowel movement that may affect quantitative analysis or radiation therapy planning.

Much worthwhile research efforts have been carried out in previous studies to handle some of the above-referenced limitations. The Visible Korean male phantom [16] is a landmark example where morphing techniques have been developed in order to modify the volume and shape of static phantoms. The employed methods include physics-based approaches [12], image registration techniques [17, 18] and geometrical approaches [11, 19, 20]. Finally, organ motion modeling techniques were developed to consider patients' involuntary respiratory motion. This led to a new generation of 4D computational models (3D space + time) that became practical tools for simulation in medical imaging as well as in radiotherapy treatment for oncological applications. Examples include respiratory motion simulation using rigid or elastic transformations, as adopted in the popular 4-D XCAT phantom [21], and deformable voxelized phantoms [22] using more sophisticated techniques, such as finite element algorithms [23].

From mathematical to voxel-based to boundary representation models

Computational human phantoms have been developed to realistically model patients' anatomy and physiology, considering the geometry and structures of organs/tissues, material composition, temporal changes, such as respiratory/cardiac motions, fluid dynamics such as blood flow or contrast perfusion, …etc [24]. Computational phantoms have been extended from simple water-filled slabs and spheres to anthropomorphic models with a realistic representation of the anatomy and material composition. Computational models are typically classified into three main categories; a) stylized phantoms, b) voxel phantoms and c) boundary representation phantoms (Figure 1).

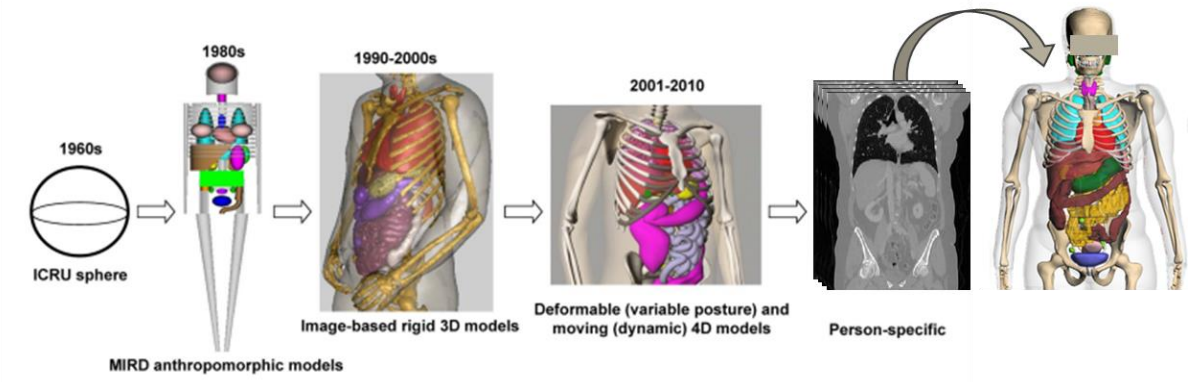


Figure 1. The evolution of computational phantoms from simple macro-bodies to detailed personalized models. Adapted with permission from [25].

The first generation of computational phantoms was constructed for radiation protection purposes in the 1960s. It was primarily composed of simple macrobodies, easily described based on quadratic equations, such as cuboids, cylinders, spheres, ellipsoids, … etc [26]. The first anthropomorphic stylized phantom developed by Fisher and Snyder comprised only three regions (skeleton, lungs and remainder tissues) [27]. Nine years later, they built an improved version of their phantom composed of main organs defined by simple geometric primitives [28]. Along with the technical developments of stylized phantoms, diversities in the target population according to age (from newborn to adult), gender (male/ female) and pregnancy gestation (fetus models) [29-31] were modeled. For a long time, mathematical models served as the de facto standard in radiation protection and dose management. Many upgraded versions of these phantoms have been constructed, such as Adam and Eva [32], precise head and brain models [33], bone and marrow [34, 35], gastrointestinal tract [36], … etc. Furthermore, to cover anatomical diversities of patient/worker populations, a library of stylized phantoms with different statures has been devised [37]. 4D stylized phantoms representing organ motion were developed based on surface equations, such as superquadratics [38] and non-uniform rational B-splines (NURBS) [39]. Mathematical models have the

advantages of easy manipulation of shape and size adjustment or motion simulation [40]. However, this design lacks anatomical realism as the model represents only a crude approximation of organs' shape and position. In addition, the definition of heterogenous tissue composition in macrobodies is not possible.

Through the advent of tomographic medical imaging modalities, such as CT and MRI, the visualization of the human anatomy in three-dimensions was made feasible. Medical images consist of small elements called pixels in 2D and voxels in 3D representing tissue information. To construct a 3D computational phantom, a label is assigned to each voxel according to the anatomical region (i.e., liver, brain, bone, etc) and tissue characteristics (material composition and density) obtained from medical images (i.e., CT or MRI). The segmentation of organs and tissues from medical images is traditionally performed manually, a labor-intensive and time-consuming process. Although voxel phantoms provided significant anatomical realism compared to stylized models, they suffer from limitations attributed to the finite voxel resolution of structural images (in the order of millimeters) and the inherent nature of voxel element geometry (uneven steps). In fact, tomographic images are not capable of representing fine structures in micrometer dimensions, such as the skin, eye lens and epithelial tissue in the digestive tract. As a result, the anatomical fidelity of the developed model depends on voxel size, and most existing voxel models involve some level of assumptions about the anatomical structures. Furthermore, CT images that are mostly used as reference structural images do not generally represent soft-tissue contrast and typically cover only part of the body (not total-body images).

A number of reference anthropomorphic voxel phantoms have been developed mostly based on CT images. These reference models were first developed for the adult male model and later extended to the adult female, pediatric and pregnant phantoms. In the late 1980s, Zankl et al. constructed voxel computational models using CT images of healthy patients that eventually ended up in 12-phantoms family representing different ages, gender and size [41-43]. In 1994, a head-torso voxel model, referred to as the VoxelMan was developed from CT images to support imaging physics research in nuclear medicine [44]. In 2000, the VIP-Man phantom was developed by Xu et al. as the first model constructed based on color photographic images of a cadaver [45]. In 2002, the dose Calculation task group of the International Commission on Radiological Protection (ICRP) Committee launched a project focusing on the development of a set of standard voxel phantoms to be released to the public as the ICRP Reference phantoms (adult male and female) [46]. Bolch et al. created a series of pediatric reference phantoms from newborn to 15-years old teenager [47]. As the most recent developed pediatric family phantoms, ICRP publication 143 describes the development of a series of 10 computational models composed of male and female newborn, 1 year, 5 years, 10 years, and 15 years old phantoms [48].

Computational models based on boundary representation techniques were introduced as a new computational model generation taking advantage of both mathematical and voxel-based models. BREP phantoms are able to represent realistic anatomy of the human body while can benefiting from the advantages of mathematical phantoms in modeling the deformations. Surface-based models, such as non-uniform rational B-splines and polygon mesh models, are subcategories of BREP modeling. These advanced surface models are capable of realistically representing the anatomical structures while enabling the simulation of anatomical deformation (posture and involuntary organ motion) by providing a rich set of mathematical operation tools. In surface-based designs, the transformations can be simply applied to the surfaces or vertex points to morph the object.

In the BREP or hybrid approach, voxel data are combined with stylized modeling techniques to design a computational phantom that benefits the advantages of both voxel models in representing the anatomical realism and stylized phantoms in providing the flexibility for anatomical variations [49]. A series of reference phantoms (adult and pediatric) in hybrid format has been developed by the University of Florida [50-52]. A series of hybrid pregnant female phantoms at the end of three gestational periods has been developed by Xu et al. [53] whereas a family of Iranian BREP phantoms (adult male/female and pregnant reference phantoms) has been developed at Ferdowsi University of Mashhad [54-56] as illustrated in Figure 2. The Virtual Family, a series of surface-based computational phantoms, has been developed based on high-resolution MR images [10, 11]. A polygon surface phantom at Hanyang University in Korea extended from the reference voxel model of VKH-Man was also designed [57]. Recently, mesh-type ICRP reference adult phantoms, which account for surface-based counterparts of the voxel-type ICRP reference phantoms, have also been developed [58, 59].

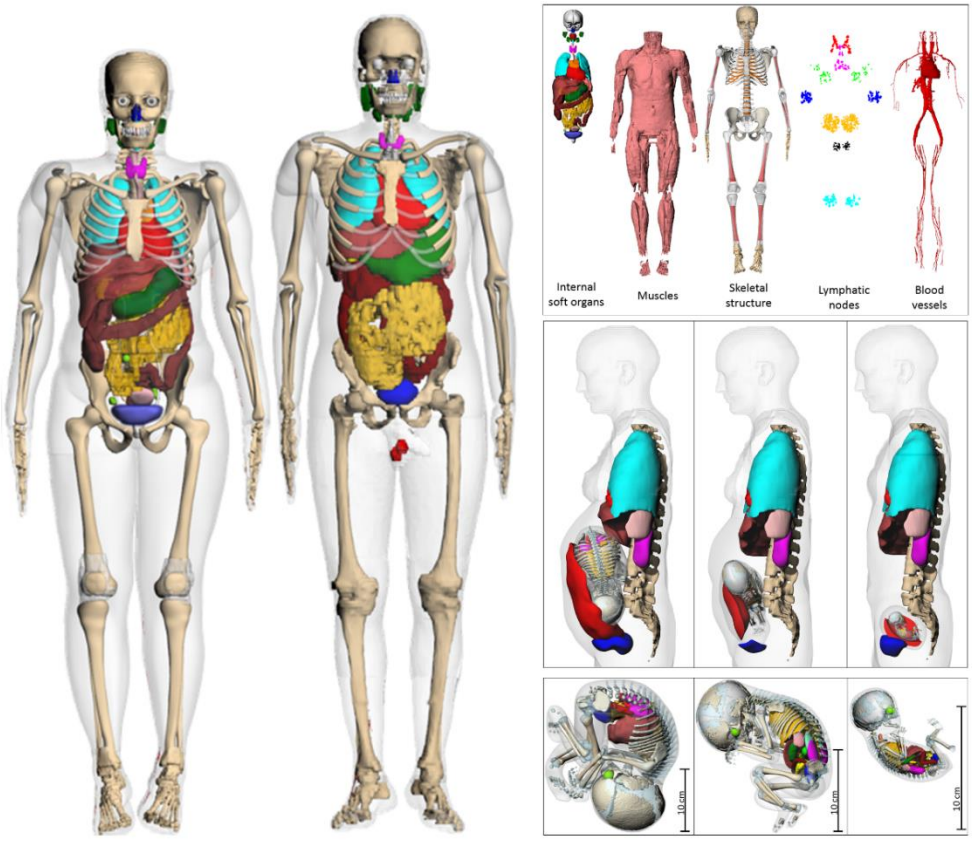


Figure 2. The reference BREP Iranian computational phantoms. Left panel: male/female adult reference computational phantoms. Right-top panel: the segmented structural/anatomical details and, Right-bottom panel: the reference pregnant phantom with fetus model at three gestation periods. Courtesy of Dr. Miri and Dr. Rafat, Ferdowsi University of Mashhad.

III. Extensions of reference phantoms

Motion and posture simulation

Reference computational phantoms are constructed using tomographic images of a single subject, thus lacking inter-subject anatomical variability. In addition, reference models have been traditionally developed as static models where the physiological dynamics of the human body behavior is ignored. To address these limitations, a library of anatomically variable computational phantoms and time-varying 4-D reference phantoms have been developed.

In digital models, physiological motion is typically captured from gated imaging, where the data acquisition is synchronized with a physiological signal. This information is used to simulate motion through time-varying transformations of the body structures. In BREP designed phantoms, the topological transformations are applied to surface control points. The 4-D NCAT phantom, an extension from the earlier mathematical MCAT phantom by Segars et al. [39, 60], representing cardiac and respiratory motions, was the first NURBS-based torso model. In an updated version, the 4-D XCAT phantom family was extended to include a series of 47 phantoms representing cardiac and respiratory motions of different patients [61]. A number of studies reported on the extension of 4D XCAT phantoms. For instance, Ghaly et al. [62] developed a population of 4D phantoms by deforming the 3D XCAT reference model. In addition, Konik et al. [63] simulated non-rigid respiratory and voluntary body motion based on the XCAT model. The 4-D VIP-Man phantom developed based on polygonal mesh was employed for external beam treatment planning in the lung region [64]. For CSG design, respiratory motion transformations are applied to individual voxels by linear interpolation of the deformation vector fields to generate a series of high-temporal-resolution voxel phantoms [22].

Morphing and changing the posture of reference phantoms is a useful technique to mimic real-world scenarios in the radiation protection domain [2]. Since building a new posture-specific phantom is challenging and time-consuming, morphing techniques have been developed enabling the deformation of volumetric and topological

features of organs and structures from existing reference computational models. Strategies for morphing the anatomy encompass simple heuristic methods for scaling and transpositions of organs to complex non-rigid registration techniques [19, 65]. These strategies were exploited to extend a series of computational phantoms with different anatomical characteristics, such as height, weight, BMI, …etc, from a reference computational model. This strategy will be further elaborated in the section below "Libraries of computational human phantoms". To simulate physiological motion of anatomical structures, posture-specific phantoms have been devised. The postures were adjusted based on the information derived from a body motion capture system to realistically simulate sequence of body movements [20, 66]. Han et al. [67] developed walking phantoms suitable for radiation dosimetry in external photon exposure scenarios. Another study by Su et al. [68] reported on sitting phantoms designed for internal radiation dosimetry studies. Recently, mesh type ICRP phantoms were deformed to multiple non-standing postures using a posture-change method based on a rigid shape-deformation algorithm and motion-capture technology [69] to measure the radiation dose in specific situations.

Libraries of computational human phantoms

Reference phantoms are constructed according to the anatomical characteristics of a single subject considering anthropomorphic data of the average population, therefore lacking inter-subject anatomical variability (Figure 3). The diversity of anthropometric parameters between individuals raises the demand for building patient-specific computational phantoms. Although personalized phantoms are deemed to represent the ideal digital twins, there are some limitations associated with the construction of individualized phantoms. This includes the lack of high-resolution tomographic images for specific patients and the time-consuming procedures for organ/tissue segmentation. In this regard, habitus-specific phantom series created based upon the deformation of a reference phantom assembling different anatomical variables for population-based assessments have been introduced. Deformation algorithms have been typically developed based on morphing the tissues considering hyper-elastic soft-tissue and stiff joints. Some interactive tools enabling topological morphing and interpolation of tissues, such as tissue growth (analogous to thermal expansion), to construct a habitus-variable computational population from a reference model have been developed. A number of studies reported on size-adjustable phantoms representing the variability of anatomical and anthropomorphic parameters, such as body size, organ volume/shape, …etc. Johnson et al. [70] extended the UF hybrid adult phantoms to 25 habitus-specific computational phantoms. Na et al. [19] reported on the construction of a library of adult phantoms (weight-specific) extended from the RPI reference models using an automated deformation algorithm implemented on polygon mesh surfaces. In addition, a number of obese phantoms and a set of age-dependent Chinese computational models in mesh format have been developed based on the RPI reference phantoms to examine the effect of obesity on CT dosimetry [71, 72]. Broggio et al. [73] constructed 25 adult phantoms to cover the diversity of heights and weights in the adult male population. Lloyd et al. [74] developed a non-rigid deformation algorithm to extend the population of the Virtual Family phantoms using biomechanical finite element methods. Geyer et al. [75] extended the UF reference phantom family to height/weight-specific phantoms. A Korean library of voxel phantoms has been developed to represent different body shapes and sizes [76]. More recently, Akhavanallaf et al. [65] developed an automated algorithm to construct a comprehensive library of phantoms extended from voxel-based ICRP reference phantoms. Choi et al. [77] extended a body-size dependent family of adult phantoms based on mesh-type ICRP reference phantoms. Hoseinian et al. [78-80] created a comprehensive series of BREP whole-body phantoms covering statistical diversities of the Iranian population (Figure 3). Beside the development of total-body phantom families, Erickson et al [81] established a database of realistic virtual breast models based on breast computed tomography images.

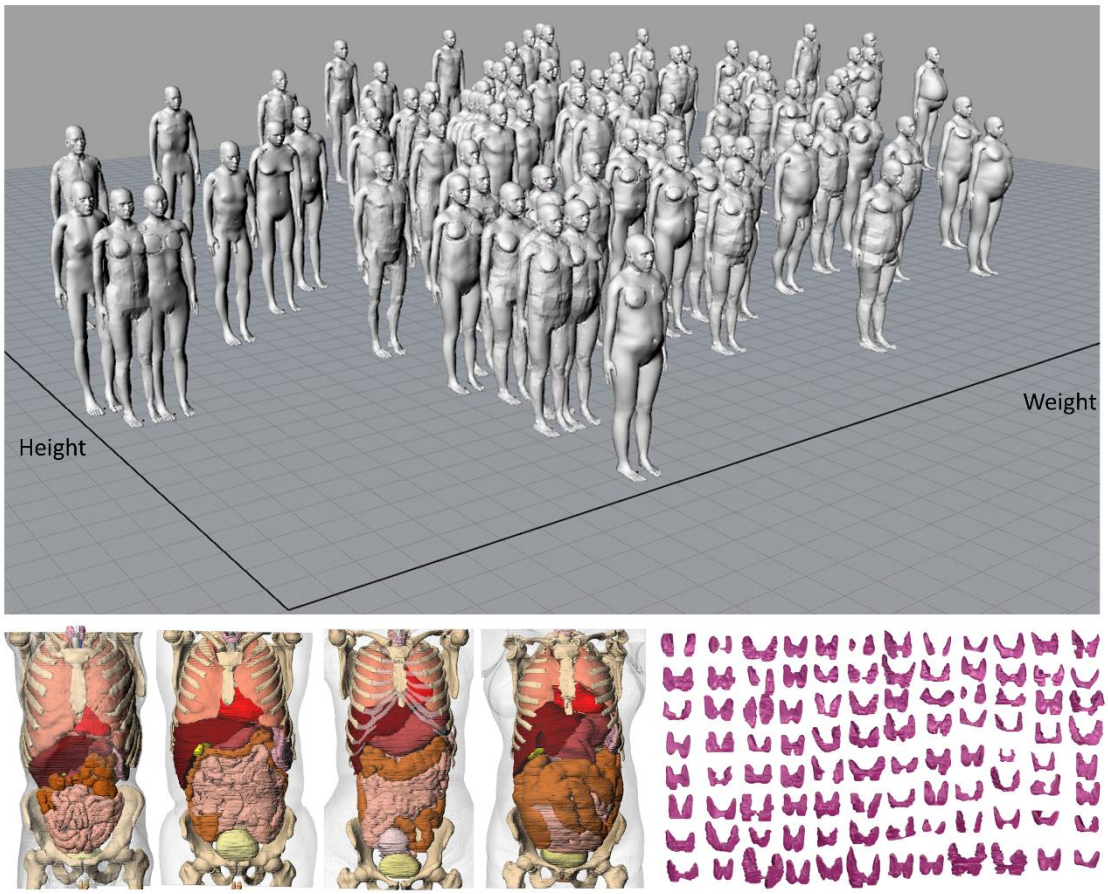


Figure 3. Series of adult computational phantoms (males and females) developed based on CT images of healthy Iranian population. The distribution of anthropomorphic indices, height and weight (top panel), along with the structural details of the developed computational models (left-bottom panel) are shown. As an example, the anatomical deviations of the thyroid gland in this population is illustrated (right-bottom panel). Courtesy of Dr. Miri and Dr. Rafat, Ferdowsi University of Iran.

IV. Advances in computational models

Recent advances in computational phantoms design focused on two main aspects: first, realistic representation of patient-specific anatomy; and second, upgrade of reference phantoms by adding small (micrometer-scale) or complex anatomical structures. The anatomical fidelity of the developed computational phantoms depends strongly on the voxel resolution of the reference tomographic images. Current imaging technologies are not capable of representing complex or fine structures, such as bone marrow, eye lens, alimentary tract structures, …etc, in micrometer-scale. In this context, Yeom et al. from Hanyang University developed mesh type ICRP phantoms extended from the reference ICRP voxel models through preserving the original anatomical structures. Complex structures of the gastrointestinal system have been improved and the fine structures of the alimentary and respiratory tracts and lung airways were added (Figure 4) [82, 83]. In a recent study, mesh-type skeletal systems for pediatric population and detailed eye models for children and adolescents of the ICRP reference phantoms were developed in a fine-structure [84, 85]. Abadi et al. [86] elaborated on the lung architecture of the XCAT series by modeling the airways and pulmonary vasculature. MIDA is a detailed head and neck model (over 160 structures) constructed from a series of high-resolution multimodal MR sequences [87]. In parallel, advanced functionalized anatomical models have been developed [12].

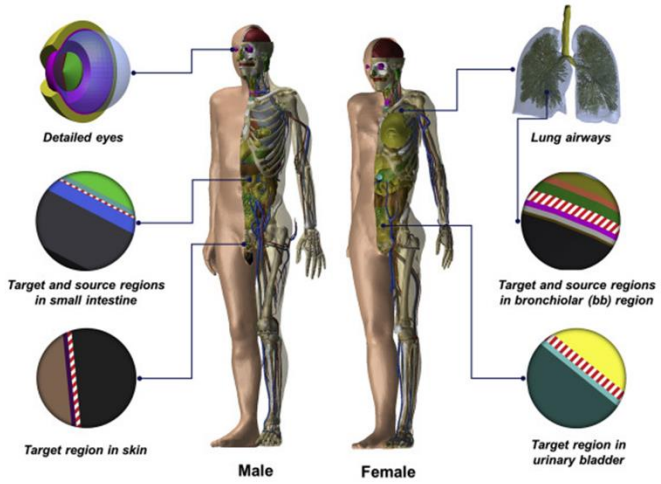


Figure 4. Adult male and female mesh-type ICRP reference computational phantoms. Micron-scale radiosensitive regions of major organs and tissues are visualized on the left and right sides of the phantoms. Reprinted with permission from [88].

Toward the 4th generation patient-specific digital twins, semi-automatic segmentation techniques based on texture patterns or manual delineation along with deformable registration algorithms have been traditionally employed. In state-of-the-art approaches, this process was labor-intensive and time-consuming which limited the construction of patient-specific computational models [89]. Recently Carter et al. proposed to use deformable registration techniques to create individualized phantoms to better support patient-specific dosimetry [90]. Thanks to recent advances in artificial intelligence algorithms, fully automated segmentation of medical images became feasible. In this field, machine learning and deep learning techniques proved to serve as useful techniques to generate patient-specific phantoms. Deep learning algorithms demonstrated their capabilities in image segmentation [91-95] and image registration [91, 96] that can be integrated into the workflow for the construction of patient-specific computational phantoms for diagnostic and radiotherapy risk assessment purposes [97, 98]. In 2019, Xie et al. [99] constructed patient-specific pregnant phantoms by means of deep learning anchor organ segmentation and used them as input for Monte Carlo organ dose calculations. As illustrated in Figure 5, the generated patient-specific phantoms were utilized to estimate fetal exposure from abdominal CT examinations.

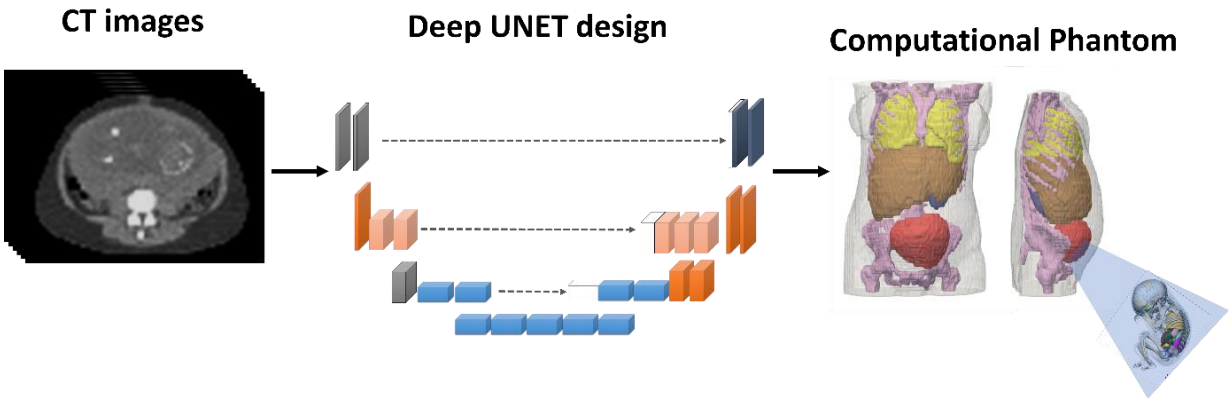


Figure 5. Illustration of the deep learning pipeline used to automatically generate pregnant computational phantoms.

Peng et al. [92] used deep learning techniques to automatically segment CT images and combined it with accelerated Monte Carlo simulations to calculate patient-specific radiation dose to make the implementation of the proposed approach in clinical setting feasible. They claimed that the proposed patient-specific phantom constructed based on automatic segmentation is prone to much less error compared with the selection of a

computational phantom from available libraries. Recently Fu et al. [91] developed a unified pipeline to create personalized computational models from radiological images. The proposed pipeline is composed of three main steps: first, some anchor organs were segmented from CT images using deep learning algorithms; second, the best-matching reference phantom was selected from a template phantom library using parametrized template matching approach; and third, a deformable registration between CT images merged with anchor organ masks and the selected reference phantom was carried out. They registered patient images to one of the phantoms selected from the XCAT library (Figure 6).

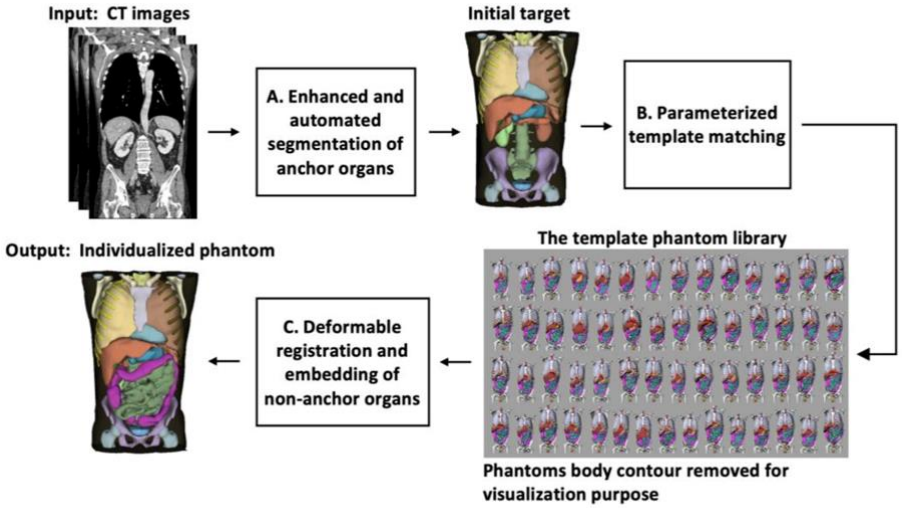


Figure 6. The pipeline for automated construction of personalized computational phantoms. Reprinted with permission from [91].

V. Summary and future perspectives

Advances in high-performance computing and the capabilities offered by deep learning-based algorithms has triggered important developments toward the 4th generation of human computational models called digital twins which represent the biological and physical characteristics of the human body from gene information to anthropomorphic parameters. Recent advances in deep learning-assisted medical image analysis and processing successfully pushed the borders toward real-time patient-specific computational models. Considering the time-consuming process of organs labeling and generation of ground truth in the supervised approach, novel unsupervised models, such as variational auto-encoders or generative adversarial networks, seem promising for application in this area [100].

Acknowledgements

We acknowledge Dr. Hashem Miri and Dr. Lale Rafat at Ferdowsi University of Iran for providing illustrations of their computational models. The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Qatar national research fund (grant number NPRP10-0126-170263) and the Swiss National Science Foundation under grant SNRF 320030_176052.

References

[1] M. Caon, "Voxel-based computational models of real human anatomy: a review,"*Radiat Environ Biophys*, vol. 42, no. 4, pp. 229-35, 2004.

[2] W. Kainz *et al.*, "Advances in Computational Human Phantoms and Their Applications in Biomedical Engineering - A Topical Review,"*IEEE Trans Radiat Plasma Med Sci*, vol. 3, no. 1, pp. 1-23, 2019.

[3] H. Zaidi and X. G. Xu, "Computational anthropomorphic models of the human anatomy: the path to realistic Monte Carlo modeling in radiological sciences,"*Annu Rev Biomed Eng*, vol. 9, pp. 471-500, 2007.

[4] M. Tajik, M. M. Akhlaqi, and S. Gholami, "Advances in anthropomorphic thorax phantoms for radiotherapy: a review,"*Biomed Phys Eng Express*, p. *in press*, 2021.

[5] V. Filippou and C. Tsoumpas, "Recent advances on the development of phantoms using 3D printing for imaging with CT, MRI, PET, SPECT, and ultrasound,"*Med Phys*, vol. 45, no. 9, pp. e740-760, 2018.

[6] E. O. Popa, M. van Hilten, E. Oosterkamp, and M. J. Bogaardt, "The use of digital twins in healthcare: socio-ethical benefits and socio-ethical risks,"*Life Sci Soc Policy*, vol. 17, no. 1, p. 6, 2021.

[7] V. Spitzer, M. J. Ackerman, A. L. Scherzinger, and D. Whitlock, "The visible human male: a technical report,"*J Am Med Inform Assoc*, vol. 3, no. 2, pp. 118-30, 1996.

[8] X. G. Xu, "An exponential growth of computational phantom research in radiation protection, imaging, and radiotherapy: a review of the fifty-year history,"*Phys Med Biol*, vol. 59, no. 18, pp. R233-302, 2014.

[9] W. P. Segars, B. M. W. Tsui, C. Jing, Y. Fang-Fang, G. S. K. Fung, and E. Samei, "Application of the 4-D XCAT Phantoms in Biomedical Imaging and Beyond,"*IEEE Trans Med Imaging*, vol. 37, no. 3, pp. 680-692, 2018.

[10] A. Christ *et al.*, "The Virtual Family--development of surface-based anatomical models of two adults and two children for dosimetric simulations,"*Phys Med Biol*, vol. 55, no. 2, pp. N23-38, 2010.

[11] M. C. Gosselin *et al.*, "Development of a new generation of high-resolution anatomical models for medical device evaluation: the Virtual Population 3.0,"*Phys Med Biol*, vol. 59, no. 18, pp. 5287-303, 2014.

[12] E. Neufeld, B. Lloyd, B. Schneider, W. Kainz, and N. Kuster, "Functionalized Anatomical Models for Computational Life Sciences,"*Front Physiol*, vol. 9, p. 1594, 2018.

[13] M. Leyton, *A generative theory of shape*. Springer, 2003.

[14] I. Stroud, *Boundary representation modelling techniques*. Springer Science & Business Media, 2006.

[15] W. Schneider, T. Bortfeld, and W. Schlegel, "Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions,"*Phys Med Biol*, vol. 45, no. 2, pp. 459-78, 2000.

[16] J. S. Park, M. S. Chung, S. B. Hwang, Y. S. Lee, D. H. Har, and H. S. Park, "Visible Korean human: improved serially sectioned images of the entire body,"*IEEE Trans Med Imaging*, vol. 24, no. 3, pp. 352-60, 2005.

[17] W. P. Segars *et al.*, "The development of a population of 4D pediatric XCAT phantoms for imaging research and optimization,"*Med Phys*, vol. 42, no. 8, pp. 4719-26, 2015.

[18] A. Akbarzadeh *et al.*, "Evaluation of whole-body MR to CT deformable image registration,"*J Appl Clin Med Phys*, vol. 14, no. 4, p. 4163, 2013.

[19] Y. H. Na, B. Zhang, J. Zhang, P. F. Caracappa, and X. G. Xu, "Deformable adult human phantoms for radiation protection dosimetry: anthropometric data representing size distributions of adult worker populations and software algorithms,"*Phys Med Biol*, vol. 55, no. 13, pp. 3789-811, 2010.

[20] T. C. Fonseca, R. Bogaerts, J. Hunt, and F. Vanhavere, "A methodology to develop computational phantoms with adjustable posture for WBC calibration,"*Phys Med Biol*, vol. 59, no. 22, pp. 6811-25, 2014.

[21] W. P. Segars, G. Sturgeon, S. Mendonca, J. Grimes, and B. M. Tsui, "4D XCAT phantom for multimodality imaging research.,"*Med Phys*, vol. 37, no. 9, pp. 4902-4915, 2010.

[22] M. C. Han *et al.*, "Continuously Deforming 4D Voxel Phantom for Realistic Representation of Respiratory Motion in Monte Carlo Dose Calculation,"*IEEE Transactions on Nuclear Science*, vol. 63, no. 6, pp. 2918-2924, 2016.

[23] W. P. Segars, A. I. Veress, G. M. Sturgeon, and E. Samei, "Incorporation of the Living Heart Model into the 4D XCAT Phantom for Cardiac Imaging Research,"*IEEE Trans Radiat Plasma Med Sci*, vol. 3, no. 1, pp. 54-60, 2019.

[24] E. Abadi *et al.*, "Virtual clinical trials in medical imaging: a review,"*J Med Imaging (Bellingham)*, vol. 7, no. 4, p. 042805, 2020.

[25] H. Zaidi and B. M. W. Tsui, "Review of Computational Anthropomorphic Anatomical and Physiological Models,"*Proceedings of the IEEE*, vol. 97, no. 12, pp. 1938-1953, 2009.

[26] J. Fisher and W. Snyder, "Variation of dose delivered by137 Cs as a function of body size from infancy to adulthood,"*ORNL-4007*, pp. 221-228, 1966.

[27] W. S. Snyder, H. L. Fisher, Jr., M. R. Ford, and G. G. Warner, "Estimates of absorbed fractions for monoenergetic photon sources uniformly distributed in various organs of a heterogeneous phantom,"*J Nucl Med*, pp. Suppl 3:7-52, 1969.

[28] W. Snyder, M. Ford, and G. Warner, "Estimates of specific absorbed fractions for photon sources uniformly distributed in various organs of a heterogeneous phantom 1978 Oak Ridge National Laboratory NM/MIRD Pamphlet No. 5,"*Oak Ridge National Laboratory NM/MIRD Pamphlet*, no. 5, 1978.

[29] M. Cristy, "Mathematical phantoms representing children of various ages for use in estimates of internal dose," Oak Ridge National Lab., TN (USA), 1980.

[30] M. Stabin *et al.*, "Mathematical models and specific absorbed fractions of photon energy in the nonpregnant adult female and at the end of each trimester of pregnancy," Oak Ridge National Lab., TN (United States), 1995.

[31] M. Cristy and K. Eckerman, "Specific Absorbed fractions of energy at various ages from internal photon sources: 1, Methods," Oak Ridge National Lab., TN (USA), 1987.

[32] Kramer R, Zankl M, Williams G, and D. G., *The Male (Adam) and Female (Eva) Adult Mathematical Phantoms*. GSF, 1986.

[33] L. G. Bouchet, W. E. Bolch, D. A. Weber, H. L. Atkins, and J. W. Poston, "Mird pamphlet no. 15: Radionuclide S values in a revised dosimetrie model of the adult head and brain,"*Journal of Nuclear Medicine*, vol. 40, no. 3, pp. 62S-101S, 1999.

[34] K. F. Eckerman and M. G. Stabin, "Electron absorbed fractions and dose conversion factors for marrow and bone by skeletal regions,"*Health Phys*, vol. 78, no. 2, pp. 199-214, 2000.

[35] L. G. Bouchet, W. E. Bolch, R. W. Howell, and D. V. Rao, "S values for radionuclides localized within the skeleton,"*J Nucl Med*, vol. 41, no. 1, pp. 189-212, 2000.

[36] J. W. Poston, Jr., K. A. Kodimer, W. E. Bolch, and J. W. Poston, Sr., "A revised model for the calculation of absorbed energy in the gastrointestinal tract,"*Health Phys*, vol. 71, no. 3, pp. 307-14, 1996.

[37] I. Clairand, L. G. Bouchet, M. Ricard, M. Durigon, M. Di Paola, and B. Aubert, "Improvement of internal dose calculations using mathematical models of different adult heights,"*Phys Med Biol*, vol. 45, no. 10, pp. 2771-85, 2000.

[38] J. Peter, D. Gilland, R. Jaszczak, and R. Coleman, "Four-dimensional superquadric-based cardiac phantom for Monte Carlo simulation of radiological imaging systems,"*IEEE Trans Nucl Sci* vol. 46, no. 6, pp. 2211-2217, 1999.

[39] W. P. Segars, *Development and application of the new dynamic NURBS-based cardiac-torso (NCAT) phantom*. The University of North Carolina at Chapel Hill, 2001.

[40] P. H. Pretorius, W. Xia, M. A. King, B. M. Tsui, T. S. Pan, and B. J. Villegas, "Evaluation of right and left ventricular volume and ejection fraction using a mathematical cardiac torso phantom,"*J Nucl Med*, vol. 38, no. 10, pp. 1528-35, 1997.

[41] G. Williams, M. Zankl, W. Abmayr, R. Veit, and G. Drexler, "The calculations of dose from external photon exposures using reference and realistic human phantoms and Monte Carlo methods,"*Physics in Medicine and Biology*, vol. 31, no. 4, pp. 449-452, 1986.

[42] M. Zankl *et al.*, "The construction of computer tomographic phantoms and their application in radiology and radiation protection,"*Radiat Environ Biophys*, vol. 27, no. 2, pp. 153-64, 1988.

[43] M. Zankl, J. Becker, U. A. Fill, N. Petoussi-Henss, and K. F. Eckerman, "GSF male and female adult voxel models representing ICRP reference man-The present status," in *The Monte Carlo Method: Versatility Unbounded in a Dynamic Computing World Proc. of the Monte Carlo 2005 Topical Meeting on CD-ROM (Chattanooga, TN, 17–21 April 2005) (LaGrange, Park IL: American Nuclear Society)*, 2005.

[44] I. G. Zubal, C. R. Harrell, E. O. Smith, Z. Rattner, G. Gindi, and P. B. Hoffer, "Computerized three-dimensional segmented human anatomy,"*Med Phys*, vol. 21, no. 2, pp. 299-302, 1994.

[45] X. G. Xu, T. C. Chao, and A. Bozkurt, "VIP-Man: an image-based whole-body adult male model constructed from color photographs of the Visible Human Project for multi-particle Monte Carlo calculations,"*Health Phys*, vol. 78, no. 5, pp. 476-86, 2000.

[46] H. G. Menzel, C. Clement, and P. DeLuca, "ICRP. Adult reference computational phantoms. ICRP Publication 110,"*Ann ICRP*, vol. 39, no. 2, pp. 1-164, 2009.

[47] C. Lee, J. L. Williams, C. Lee, and W. E. Bolch, "The UF series of tomographic computational phantoms of pediatric patients,"*Med Phys*, vol. 32, no. 12, pp. 3537-48, 2005.

[48] W. E. Bolch *et al.*, "ICRP Publication 143: Paediatric Reference Computational Phantoms,"*Ann ICRP*, vol. 49, no. 1, pp. 5-297, 2020.

[49] W. Bolch, C. Lee, M. Wayson, and P. Johnson, "Hybrid computational phantoms for medical dose reconstruction,"*Radiat Environ Biophys*, vol. 49, no. 2, pp. 155-68, 2010.

[50] C. Lee, C. Lee, J. L. Williams, and W. E. Bolch, "Whole-body voxel phantoms of paediatric patients--UF Series B,"*Phys Med Biol*, vol. 51, no. 18, pp. 4649-61, 2006.

[51] C. Lee, D. Lodwick, D. Hasenauer, J. L. Williams, C. Lee, and W. E. Bolch, "Hybrid computational phantoms of the male and female newborn patient: NURBS-based whole-body models,"*Phys Med Biol,* vol. 52, no. 12, pp. 3309-33, 2007.

[52] C. Lee, D. Lodwick, J. L. Williams, and W. E. Bolch, "Hybrid computational phantoms of the 15-year male and female adolescent: applications to CT organ dosimetry for patients of variable morphometry,"*Med Phys,* vol. 35, no. 6, pp. 2366-82, 2008.

[53] X. G. Xu, V. Taranenko, J. Zhang, and C. Shi, "A boundary-representation method for designing whole-body radiation dosimetry models: pregnant females at the ends of three gestational periods--RPI-P3, -P6 and -P9,"*Phys Med Biol,* vol. 52, no. 23, pp. 7023-44, 2007.

[54] E. Hoseinian-Azghadi, L. Rafat-Motavalli, and H. Miri-Hakimabad, "Development of a 9-months pregnant hybrid phantom and its internal dosimetry for thyroid agents,"*J Radiat Res,* vol. 55, no. 4, pp. 730-47, 2014.

[55] L. Rafat-Motavalli, H. Miri-Hakimabad, and E. Hoseinian-Azghadi, "Hybrid pregnant reference phantom series based on adult female ICRP reference phantom,"*Radiat. Phys. Chem.,* vol. 144, pp. 386-395, 2018.

[56] L. Rafat Motavali, N. Rafat Motevalli, and S. H. Miri Hakimabad, "The first series of Iranian BREP phantoms," in *7th International Workshop on Computational Human Phantoms*, 2019.

[57] C. H. Kim, J. H. Jeong, W. E. Bolch, K. W. Cho, and S. B. Hwang, "A polygon-surface reference Korean male phantom (PSRK-Man) and its direct implementation in Geant4 Monte Carlo simulation,"*Phys Med Biol,* vol. 56, no. 10, pp. 3137-61, 2011.

[58] C. H. Kim *et al.,* "New mesh-type phantoms and their dosimetric applications, including emergencies,"*Ann ICRP,* vol. 47, no. 3-4, pp. 45-62, 2018.

[59] C. H. Kim *et al.,* "ICRP Publication 145: Adult Mesh-Type Reference Computational Phantoms,"*Annals of the ICRP,* vol. 49, no. 3, pp. 13-201, 2020.

[60] K. LaCroix, "'Evaluation of an attenuation compensation method with respect to defect detection in Tc-99m-sestamibi myocardial SPECT," in *Dept. of Biomedical Engineering*: Univ. of North Carolina, 1997, p. 161.

[61] W. P. Segars *et al.,* "Population of anatomically variable 4D XCAT adult phantoms for imaging research and optimization,"*Med Phys,* vol. 40, no. 4, p. 043701, 2013.

[62] M. Ghaly, Y. Du, G. S. Fung, B. M. Tsui, J. M. Links, and E. Frey, "Design of a digital phantom population for myocardial perfusion SPECT imaging research,"*Phys Med Biol,* vol. 59, no. 12, pp. 2935-53, 2014.

[63] A. Könik *et al.,* "Digital anthropomorphic phantoms of non-rigid human respiratory and voluntary body motion for investigating motion correction in emission imaging,"*Phys Med Biol,* vol. 59, no. 14, pp. 3669-82, 2014.

[64] L. Hegenbart, Y. H. Na, J. Y. Zhang, M. Urban, and X. G. Xu, "A Monte Carlo study of lung counting efficiency for female workers of different breast sizes using deformable phantoms,"*Phys Med Biol,* vol. 53, no. 19, pp. 5527-38, 2008.

[65] A. Akhavanallaf, T. Xie, and H. Zaidi, "Development of a Library of Adult Computational Phantoms Based on Anthropometric Indexes,"*IEEE Transactions on Radiation and Plasma Medical Sciences,* vol. 3, no. 1, pp. 65-75, 2019.

[66] J. A. Vazquez, A. Ding, T. Haley, P. F. Caracappa, and X. G. Xu, "A dose-reconstruction study of the 1997 Sarov criticality accident using animated dosimetry techniques,"*Health Phys,* vol. 106, no. 5, pp. 571-82, 2014.

[67] B. Han, J. Zhang, Y. H. Na, P. F. Caracappa, and X. G. Xu, "Modelling and Monte Carlo organ dose calculations for workers walking on ground contaminated with Cs-137 and Co-60 gamma sources,"*Radiat Prot Dosimetry,* vol. 141, no. 3, pp. 299-304, 2010.

[68] L. Su, B. Han, and X. G. Xu, "Calculated organ equivalent doses for individuals in a sitting posture above a contaminated ground and a PET imaging room,"*Radiat Prot Dosimetry,* vol. 148, no. 4, pp. 439-43, 2012.

[69] Y. S. Yeom *et al.,* "Posture-dependent dose coefficients of mesh-type ICRP reference computational phantoms for photon external exposures,"*Phys Med Biol,* vol. 64, no. 7, p. 075018, 2019.

[70] P. B. Johnson, S. R. Whalen, M. Wayson, B. Juneja, C. Lee, and W. E. Bolch, "Hybrid patient-dependent phantoms covering statistical distributions of body morphometry in the US adult and pediatric population,"*Proceedings of the IEEE,* vol. 97, no. 12, pp. 2060-2075, 2009.

[71] A. Ding, M. M. Mille, T. Liu, P. F. Caracappa, and X. G. Xu, "Extension of RPI-adult male and female computational phantoms to obese patients and a Monte Carlo study of the effect on CT imaging dose,"*Phys Med Biol,* vol. 57, no. 9, pp. 2441-59, 2012.

[72] Y. Pi, T. Liu, and X. G. Xu, "Development of A Set of Mesh-Based and Age-Dependent Chinese Phantoms and Application For CT Dose Calculations,"*Radiat Prot Dosimetry,* vol. 179, no. 4, pp. 370-382, 2018.

[73] D. Broggio *et al.*, "Construction of an extended library of adult male 3D models: rationale and results,"*Physics in Medicine & Biology*, vol. 56, no. 23, p. 7659, 2011.
[74] B. Lloyd, E. Cherubini, S. Farcito, E. Neufeld, C. Baumgartner, and N. Kuster, "Covering population variability: morphing of computation anatomical models," in *International Workshop on Simulation and Synthesis in Medical Imaging*, 2016: Springer, pp. 13-22.
[75] A. M. Geyer, S. O'Reilly, C. Lee, D. J. Long, and W. E. Bolch, "The UF/NCI family of hybrid computational phantoms representing the current US population of male and female children, adolescents, and adults--application to CT dosimetry,"*Phys Med Biol*, vol. 59, no. 18, pp. 5225-42, 2014.
[76] J. S. Kim, W. H. Ha, J. H. Jeong, K. W. Cho, and J. K. Lee, "Use of photographic images to construct voxel phantoms for use in whole-body counting,"*Radiat Prot Dosimetry*, vol. 138, no. 2, pp. 119-22, 2010.
[77] C. Choi *et al.*, "Body-size-dependent phantom library constructed from ICRP mesh-type reference computational phantoms,"*Phys Med Biol*, vol. 65, no. 12, p. 125014, 2020.
[78] E. Hoseinian Azghadi, S. H. Miri Hakimabad, and L. Rafat Motavali, "Population of whole-body statistical adult phantoms and assessing the uncertainty of organ doses in hyperthyroid treatment with 131I," in *5th International Workshop on Computational Human Phantoms*, 2015.
[79] M. Karami, H. Miri-Hakimabad, E. Hoseinian-Azghadi, and N. Mohammadi, "A method for assessing subject-specific counting efficiency of whole-body monitoring systems for radioiodine measurements,"*Radiat. Meas.*, vol. 137, p. 106430, 2020.
[80] L. Rafat Motavali and S. H. Miri Hakimabad, "Virtual calibration of whole-body counter using a library of statistical phantoms," in *7th International Workshop on Computational Human Phantoms*, 2019.
[81] D. W. Erickson *et al.*, "Population of 224 realistic human subject-based computational breast phantoms,"*Med Phys*, vol. 43, no. 1, p. 23, 2016.
[82] H. S. Kim *et al.*, "Inclusion of thin target and source regions in alimentary and respiratory tract systems of mesh-type ICRP adult reference phantoms,"*Phys Med Biol*, vol. 62, no. 6, pp. 2132-2152, 2017.
[83] Y. S. Yeom *et al.*, "New small-intestine modeling method for surface-based computational human phantoms,"*J Radiol Prot*, vol. 36, no. 2, pp. 230-45, 2016.
[84] C. Choi *et al.*, "Development of skeletal systems for ICRP pediatric mesh-type reference computational phantoms,"*J Radiol Prot*, vol. 41, no. 2, 2021.
[85] H. Han *et al.*, "Development of detailed pediatric eye models for lens dose calculations,"*J Radiol Prot*, vol. 41, no. 2, 2021.
[86] E. Abadi, W. P. Segars, G. M. Sturgeon, J. E. Roos, C. E. Ravin, and E. Samei, "Modeling Lung Architecture in the XCAT Series of Phantoms: Physiologically Based Airways, Arteries and Veins,"*IEEE Trans Med Imaging*, vol. 37, no. 3, pp. 693-702, 2018.
[87] M. I. Iacono *et al.*, "MIDA: A Multimodal Imaging-Based Detailed Anatomical Model of the Human Head and Neck,"*PLoS One*, vol. 10, no. 4, p. e0124126, 2015.
[88] Y. S. Yeom *et al.*, "Dose coefficients of mesh-type ICRP reference computational phantoms for external exposures of neutrons, protons, and helium ions,"*Nucl. Eng. Technol.*, vol. 52, no. 7, pp. 1545-1556, 2020.
[89] T. Xie, A. Akhavanallaf, and H. Zaidi, "Construction of patient-specific computational models for organ dose estimation in radiological imaging,"*Med Phys*, vol. 46, no. 5, pp. 2403-2411, 2019.
[90] L. M. Carter, J. Camilo Ocampo Ramos, W. E. Bolch, J. S. Lewis, and A. L. Kesner, "Technical Note: Patient-morphed mesh-type phantoms to support personalized nuclear medicine dosimetry - a proof of concept study,"*Med Phys*, vol. 48, no. 4, pp. 2018-2026, 2021.
[91] W. Fu *et al.*, "iPhantom: A Framework for Automated Creation of Individualized Computational Phantoms and Its Application to CT Organ Dosimetry,"*IEEE J Biomed Health Inform*, vol. 25, no. 8, pp. 3061-3072, 2021.
[92] Z. Peng *et al.*, "A method of rapid quantification of patient-specific organ doses for CT using deep-learning-based multi-organ segmentation and GPU-accelerated Monte Carlo dose computing,"*Med Phys*, vol. 47, no. 6, pp. 2526-2536, 2020.
[93] E. Gibson *et al.*, "Automatic Multi-Organ Segmentation on Abdominal CT With Dense V-Networks,"*IEEE Trans Med Imaging*, vol. 37, no. 8, pp. 1822-1834, 2018.
[94] M. Nazari, L. D. Jimenez-Franco, M. Schroeder, A. Kluge, M. Bronzel, and S. Kimiaei, "Automated and robust organ segmentation for 3D-based internal dose calculation,"*EJNMMI Res*, vol. 11, no. 1, p. 53, 2021.
[95] H. Arabi, A. AkhavanAllaf, A. Sanaat, I. Shiri, and H. Zaidi, "The promise of artificial intelligence and deep learning in PET and SPECT imaging,"*Phys Med*, vol. 83, pp. 122-137, 2021.
[96] Y. Fu, Y. Lei, T. Wang, W. J. Curran, T. Liu, and X. Yang, "Deep learning in medical image registration: a review,"*Phys Med Biol*, vol. 65, no. 20, p. 20TR01, 2020.

[97] J. Chen, Y. Li, Y. Du, and E. C. Frey, "Generating anthropomorphic phantoms using fully unsupervised deformable image registration with convolutional neural networks,"*Med Phys,* vol. 47, no. 12, pp. 6366-6380, 2020.
[98] D. F. Bauer *et al.*, "Generation of annotated multimodal ground truth datasets for abdominal medical image registration,"*Int J Comput Assist Radiol Surg,* vol. 16, no. 8, pp. 1277-1285, 2021.
[99] T. Xie and H. Zaidi, "Estimation of the radiation dose in pregnancy: an automated patient-specific model using convolutional neural networks,"*Eur Radiol,* vol. 29, no. 12, pp. 6805-6815, 2019.
[100] Y. Zhang, S. Miao, T. Mansi, and R. Liao, "Unsupervised X-ray image segmentation with task driven generative adversarial networks,"*Medical image analysis,* vol. 62, p. 101664, 2020.

Chapter 3

Development of a library of adult computational phantoms based on anthropometric indexes

Azadeh Akhavanallaf* , Tianwu Xie*, and Habib Zaidi

Contribution: Study conception and design, methodological development, computer programming, analysis and interpretation of results, manuscript preparation.

IEEE Trans Rad Plasma Med Sci, Vol. 3, No. 1, pp 65-75 (2019)

* Authors contributed equally

Abstract

Computational phantom libraries have been developed over the years to enhance the accuracy of Monte Carlobased radiation dose calculations from radiological procedures. In this paper, we report on the development of an adult computational anthropomorphic phantom library covering different body morphometries among the 20–80 years old population. The anatomical diversities of different populations are modeled based on anthropometric parameters extracted from the National Health and Nutrition Examination Survey database, including standing height, total weight, and body mass index. Organ masses were modified to match the corresponding data. The ICRP reference male and female models were selected as anchor phantoms. A computer code was developed for adjusting standing height and percentage of fat free mass of anchor phantoms by 3-D scaling. The waist circumference and total body mass were further adjusted. The diversity of organ masses due to anthropometric differences deviates from the mean values by about 3%–21%, while this deviation exceeds 50% for genital organs. Thereafter, organ-level absorbed doses from both internal and external radiation exposure conditions were estimated. A total of 479 phantoms corresponding to seven age groups were constructed for both genders, thus fulfilling the criteria for representing a diverse adult population with different anthropomorphic and anatomical characteristics.

I. Introduction

The utilization of radiation-based diagnostic imaging systems is progressively increasing worldwide, raising concerns about the potential hazards of radiation exposure. Computational phantoms were developed to accurately model radiation interaction within the human body using Monte Carlo (MC)-based radiation transport software packages targeting a number of applications, including radiation dose calculations and imaging physics research [1]. The first generation of computational phantoms were defined by simple surface equations and initially developed in the 1960s [2]. In the late 1980s, following the advent of tomographic medical imaging technologies, such as CT and MRI, voxel-based phantoms were developed to represent the anatomical features of the human body. Voxel-based phantoms were rapidly adopted in Monte Carlo simulations owing to their ability to model anatomical details and have been continued to be developed over the years taking advantage from advances in high resolution imaging. The third generation of computational phantoms using boundary representation (BREP) techniques emerged in the form of Non-Uniform Rational B-Splines (NURBS) or polygon mesh surfaces. They offer better flexibility in terms of modeling deformation, motion and change in posture [3-5].

From a radiation protection perspective, it is essential to determine and quantify the variability of radiation dose with respect to variations in anthropometry and anatomy. To this end, anthropomorphic phantoms coupled with Monte Carlo methods play an important role in radiation dosimetry calculations. Reference anthropomorphic models were developed based on the average population [4, 6-8] but limited to fixed anthropometric and anatomical parameters. However, the diversity of anthropometric parameters between reference models and individuals may introduce significant uncertainties, thus motivating and raise the demand of personalized computational phantoms. Although, person-specific phantoms added up to an ideal model, some limitations such as lack of high resolution tomographic images for each person and a time consuming segmentation process drew the attention of researchers to a more efficient approach of computational phantoms in medical dosimetry. Habitus-specific phantoms introduced as a size adjustable type of phantoms which were constructed based on deformation in a reference phantom. These phantoms are neither too individualized like subject-specific phantoms nor population-averaged as are reference phantoms [3, 9].

A number of studies focused on size-adjustable phantoms to account for variability in body size, organ masses and other parameters, such as body fat percentage and subcutaneous fat distribution in dosimetry applications. Johnson et al. [9] built patient-dependent phantom series containing 25 models based on the University of Florida (UF) hybrid adult male (UFHADM) phantom [10] using anthropometric parameters extracted from NHANES III (1988–1994). The phantoms were remodeled considering two classes of target parameters: primary parameters (body height and weight) and secondary parameters (waist and thigh circumferences). The internal organ masses of Johnson's models deviated from reference values due to the 3D scaling during the deformation process.

In another study, an automated algorithm was developed to generate an adult phantom library using polygon mesh surfaces where the Rensselaer Polytechnic Institute (RPI) adult male and adult female models [4] were extended into a library representing morphometric diversities in the US population for the 19-year old males and females. The organ masses were assumed to follow a Gaussian normal distribution [11] according to the mean and standard deviation compiled from various sources [5].

A study was performed by Broggio et al. [12] to construct a library of 25 adult males using NURBS surfaces constructed based on full body optical models. The anthropometric parameters were extracted from Civilian American and European Surface Anthropometry Resource (CAESAR) with 109 identified organs scaled by ICRP reference data and height-dependent linear formula [13, 14]. Cassola et al. [15] produced a library of 18 phantoms from FASH and MASH mesh-based anchor phantoms using 3D modeling software where the organ masses were scaled as a function of height [16].

The first library of 4D phantoms comprising 58 NURBS models was developed by Segars et al. [17]. This library was extended based on reference XCAT phantoms [18] using real anatomy of chest-abdomen-pelvis CT data of normal patients having different Body Mass Indices (BMIs). A library of 84 adult phantoms based on the Chinese Reference Adult Male polygon Surface (CRAM_S) was constructed by Chen et al. [19] where the organ models were adjusted to match reported Chinese reference data. An extended library of UF family containing 351 computational phantoms has also been developed based on morphometric data from NHANES (1999-2006) [20].

In previous works, attention was paid to reflect more realistically the distribution of physical properties while the anatomical diversities were not appropriately considered. Different approaches have been adopted to determine organ masses representative of the considered population for the design of computational phantoms. For instance, organ masses can be set to reference values or deviate from reference masses during 3D scaling. Alternatively, organ masses were linearly correlated to only phantom statures or simply sampled using a normal distribution. A practical approximation to realistically model the anatomical variability among individuals consists in implementing multiple correlations of each organ mass with different external anthropometric parameters into the phantom series.

In this work, a software tool was developed to automatically remodel anchor reference phantoms to match target morphometric and anatomical characteristics. The developed library of voxel-based models is capable of representing internal details unlike surface models that cannot represent an inhomogeneous density distribution of organs or tissues [21]. The diversity of 13 organ masses depending on different morphometric parameters was considered. ICRP male and female voxel phantoms [22] were selected as reference models. A total of 230 male and 249 female adult voxel-based phantoms were constructed considering the diversity of anthropometric parameters from NHANES (2011-2014) [23] and variability in internal organ masses between individuals. As an application, Monte Carlo-based dosimetry calculations using the developed phantom library were performed for internal and external radiation exposure conditions.

II. Material and Methods

ICRP adult reference phantoms

The ICRP phantoms were constructed through modification of the voxel models (Golem and Laura) of a 38-year-old male (176 cm, 70 kg) and a 43-year-old female individual (167 cm, 59 kg), whose body height and mass resembled the physical characteristics of the reference male and reference female phantoms. In total, 140 organs with 53 tissue types were segmented and the organ masses of both models adjusted to the ICRP data of the adult reference male and female [22, 24]. The main characteristic of these phantoms are summarized in Table 1.

Table 1. Main characteristics of the ICRP adult male and female reference computational phantoms.

Reference Phantom Properties	Male	Female
Height (cm)	176	163
Mass (kg)	73.0	60.0
Number of tissue voxels	1,946,375	3,886,020
Slice thickness (voxel height, mm)	8.0	4.84
Voxel in-plane resolution (mm)	2.137	1.775
Voxel volume (mm^3)	36.54	15.25
Number of columns	254	299
Number of rows	127	137
Number of slices	222	348

Anthropometric data

The field of anthropometry encompasses a variety of human body measurements, such as weight, height, circumferences and lengths to represent the physical characteristics of a population. According to the trends of obesity among individuals, an updated database is required to represent a realistic body morphometry distribution. As shown in Figure 1, the results of anthropometric reference data from 1999 to 2014 reflect the obesity prevalence in the US population [25]. In the present study, morphometric parameters in seven age groups from 20 to 80 years old were obtained from the recently published NHANES (2011-2014) database. Height-weight grids obtained from the combination of height and weight percentile data (10%, 15%, 25%, 50%, 75%, 85% and 90%) and BMI range, indicating the ratio of weight to squared height, were carefully selected to limit unrealistic physical properties.

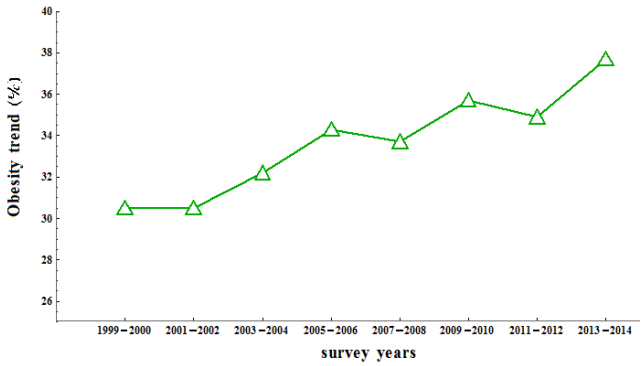


Figure 1. Trends in obesity prevalence among adults aged 20 and over, United States, 1999–2000 through 2013–2014.

Fat free mass and waist circumference target values

Since the extraction of a precise model for calculation of fat percentage and related fat free mass (FFM) percentage correlated with anthropometric parameters is not straightforward, the average values of different models reported in the literature were derived. The results show that FFM percent increases by increasing the height and decreases by increasing BMI [26-28]. Although both BMI and waist circumference measure the level of obesity, waist circumference may be more important because it is more sensitive to the distribution of body fat than BMI. To construct a more realistic body shape, waist circumference data extracted from NHANES (2011–2014) was combined with a linear model of waist circumference proportional with BMI using demographic data obtained from NHANES (2003) [29, 30].

Correlation between organ masses and anthropometric parameters

One of the most important factors influencing radiation dosimetry calculations is the mass of organs, which substantially vary between different subjects owing to specific anatomical characteristics. In the absence of person-specific imaging data, the estimation of organ masses relies on their correlation with external physical parameters. To this end, organ mass data were extracted from anthropometric parameters including age, standing height, body weight, and BMI of an individual. A survey of published articles reporting organ masses in correlation with morphometric parameters is given in Table 2. In this work, 13 organ masses including brain, heart, right and left lungs, liver, spleen, thyroid, right and left kidneys, pancreas and three genital organs for each gender were derived from autopsies and diagnostic measurements [13, 31-42]. Information on organ masses from different scientific publications has been culled in a software to extract the masses of 13 organs based on specific anthropometric parameters.

Methodology for automated model deformation

In this work, an automated algorithm was developed to remodel the reference phantom into various anthropometric and anatomical data. Computer software written in MATHEMATICA 7 (Wolfram Research Inc, Champaign, IL, USA) coupled with MATLAB 8.1 scripts (The MathWorks Inc., Natick, MA, USA) was used to implement the whole deformation process schematically displayed in Figure 2. In the reference phantoms, the blood vessels were inwardly replaced in 2D to avoid losing the vessels located in the residual tissue region during the adjustment of the fat mass for thin phantoms. The first step to reach the target anthropometric parameters consists in adjusting the height. Once the height is exactly matched, deforming the whole frame of phantoms using FFM percent was done by rescaling the phantom in 2D. To adjust organ masses, the deformation process was applied on each organ to fine tune the volume, considering the conservation of organ's center of mass position. In addition to the 13 organs scaled in association with anthropometric parameters, 4 other organs including gall bladder, stomach, thymus and urinary bladder were scaled to ICRP reference masses. The scaled organs were embedded into the body where the centroid position of each organ was kept constant through the phantom. To avoid the overlap of adjacent organs, a priority was defined for organs to be embedded according to their volume as well as their sensitivity to radiation. In the next step, for adjusting total body mass, firstly the waist circumference was considered as the estimation of belly fat. To this end, the outer contour of the phantom was detected, then adding

or removing fat layer of the trunk was iterated until waist circumference was achieved. Waist circumference was measured at the uppermost lateral border of the hip crest (ilium) by measuring the perimeter of the outer body contour. Secondly, by tuning fat mass in the legs and arms, the target total body mass excluding the skin mass was fixed to within 3% of reference values. Lastly, the whole body phantom was checked for some intersections or holes and smoothness of outer contour, and the skin layer with a thickness equal to in-plain resolution was coated. After completion of the deformation process, each deformed phantom was visually analyzed in 2D and 3D to check the anatomical structures as well as whole body habitus.

Table 2. Summary of the main publications reporting on organ mass estimation based on correlations with anthropometric parameters.

Reference	Studied Population	Organs /Fat free mass/ fat mass /waist circumference	Anthropometric parameters				Gender	Statistical analysis	Comments
			height	weight	age	BMI			
Mandal et al. [31]	300	Brain, heart, lungs, liver, spleen, thyroid, kidneys, Uterus/prostate			✓	✓	M/F	Linear correlation	Higher organ masses in higher BMI (both genders) and younger ages (liver, spleen, kidney)
Sheikhazadi et al. [32]	1222	Brain, heart, lungs, liver, spleen, thyroid, kidneys, pancreas, testis, Uterus/prostate	✓		✓	✓	M/F	Linear correlation	organ masses increase with body height and BMI values
Grandmaison et al. [13]	684	Heart, lungs, liver, spleen, thyroid, kidneys, pancreas	✓		✓	✓	M/F	Linear correlation	Organ masses decrease with age except for heart, thyroid and prostate
Molina & DiMaio [34]	232	Brain, lungs, liver, spleen, kidneys				✓	M	Statistical categorization	Reference range
Molina & DiMaio [33]	102	Brain, lungs, liver, spleen, kidneys				✓	F	Statistical categorization	Reference range
Molina & DiMaio [36]	232	Heart	✓	✓		✓	M	Strong Linear correlation	Reference range/ heart mass increases linearly with increase in body weight
Molina & DiMaio [35]	102	Heart	✓	✓		✓	F	Strong Linear correlation	
He et al. [37]	111	Brain, liver, spleen, kidneys	✓	✓	✓		M/F	Multi-correlation formula	MRI study/ smaller organ masses at higher ages, except heart
Heymsfield et al. [38]	411	Brain, liver	✓		✓		M/F	Exponential approximation	Liver scaled with height (power≈2), male brain (power≈0.83)
Chouker et al. [43]	728	Liver		✓	✓	✓	M/F	Multi-correlation formula	Liver mass increases with weight, decreases with age older than 50-60 y
Kelsey et al. [44]	59994	Ovary			✓		F	Statistical categorization / polynomial	MRI study/ 69% of the variation in ovarian volume is due to age
Perven et al. [39]	140	Ovary			✓		F	Statistical categorization	Ovarian volume shrinking with age in adults
Kelsey et al. [40]	1418	Uterus			✓		F	Statistical categorization/polynomial	Age: 0-40 y/ 84% of the variation in uterus volumes is due to age
Xia et al. [42]	1301	Prostate			✓		M	Statistical categorization	Prostate volume growth with age
Zhang et al. [41]	1000	Prostate			✓		M	Polynomial formula	Prostate volume growth with age
Heymsfield et al. [28]	13183	FFM	✓		✓		M/F	Exponential approximation	FFM scale to height with power ~2
Meeuwsen et al. [26]	23627	FM				✓	M/F	curvilinear relationship	FM percent rose compared to an increase in BMI
Schutz et al. [27]	5635	FM, FFM				✓	M/F	curvilinear relationship	
Bozeman et al. [30]		WC				✓	M/F	linear correlation	NHANES (2003)

M (male), F (female), FFM (fat free mass), FM (fat mass), WC (waist circumference)

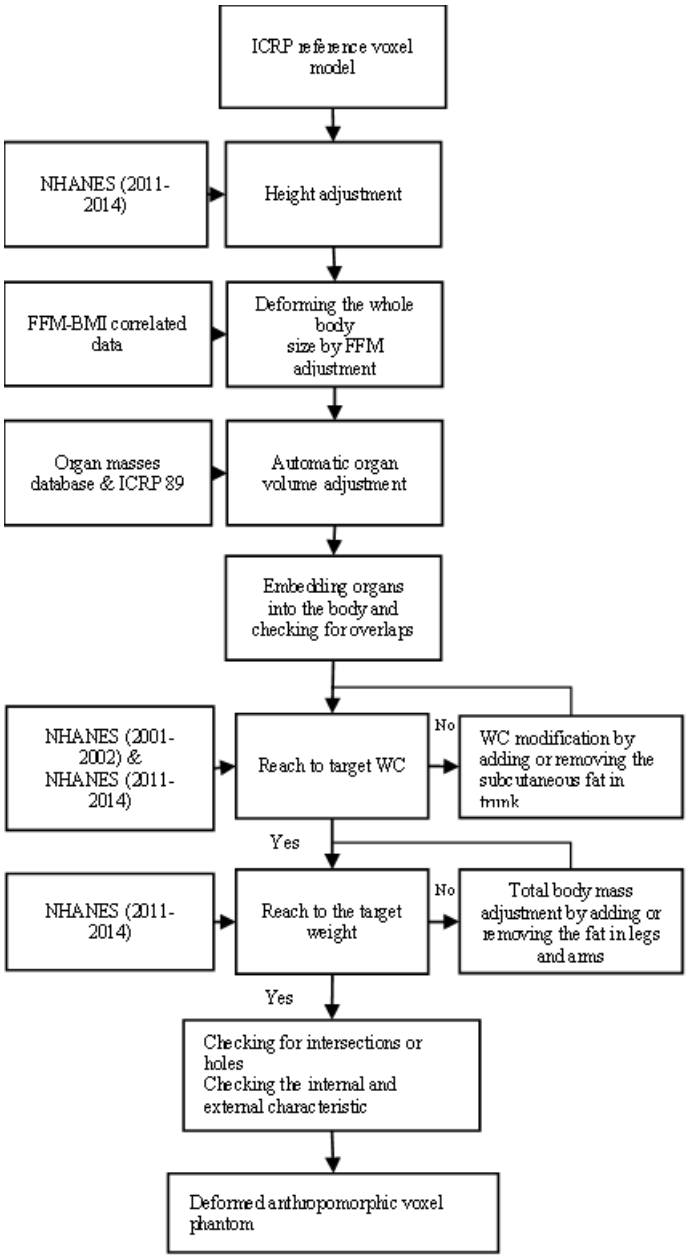


Figure 2. Flowchart of the deformation process.

Monte Carlo calculations-based organ dose assessment

1) Internal dose estimation

The absorbed doses to three morphometrically different computational phantoms from ^{18}F-FDG as PET tracer were estimated through Monte Carlo simulations using the N-Particle eXtended (MCNPX) code. ^{18}F positron-emitting source with an average energy of 0.2498 MeV was simulated in 6 source regions according to ^{18}F-FDG biokinetic data reported in ICRP 106 [45]. A total number of 10^7 primary particles were generated to reach less than 1% statistical uncertainty in most cases [46, 47].

In the MIRD formalism, the radiation absorbed dose from any source organ r_S delivered to target tissue r_T is given by Equation (1) [48]:

$$D(r_T, T_D) = \sum_{r_s} \int_0^{T_D} A(r_s, t) S(r_T \leftarrow r_s) dt = \sum_{r_s} \tilde{A}(r_s, t)\, S(r_T \leftarrow r_s) \quad (1)$$

where $\tilde{A}(r_s, t)$ is the time integrated activity in the source organ during the dose-integration period T_D, $S(r_T \leftarrow r_s)$ is the S-value defining the equivalent dose rate in the target organ per unit activity in the source organ. Using MCNPX tally card ∗F8, S-values were estimated per particle. Time-integrated activity in the source organs were obtained from ICRP 106 [45] and the administered activity of ^{18}F-FDG was similar to PET/CT acquisition protocols used in our department (3.5 MBq/kg with a maximum of 350 MBq for patients heavier than 100 kg). The effective dose was calculated based on ICRP definition:

$$E = \sum W_T \sum W_R D_{T,R} \qquad (2)$$

where E denotes the effective dose, W_R is the radiation weighting factor, $D_{T,R}$ is the absorbed dose in tissue or organ T, and W_T is the tissue weighting factor.

2) External dose estimation

To benchmark the CT radiation dose calculation using the developed library, the dose report of a female patient with high BMI who underwent a CT examination in our department under an IRB approved protocol was selected for comparison of the results with the corresponding phantom in the developed series. The characteristics of the patient and phantom are tabulated in Table 3.

Table 3. Morphometric characteristics of the patient and Female40y-25h-75w phantom (40-50 years age group, 25th percentile height and 75TH percentile weight).

Characteristics	Patient	Female$_{40y\text{-}25h\text{-}75w}$
Sex	Female	Female
Age	50 y	40 -50 y
Weight	88 Kg	85.3 Kg
Height	160 cm	158.1 cm
BMI	34.37	34.12

Examination details were extracted from the DICOM header. The study was performed on the Discovery CT 750 HD scanner (GE Healthcare, Waukesha, WI) with a scan range covering the thorax and abdomen using the following acquisition parameters: a table speed of 55 mm/rot, 0.7 sec revolution time, 1.37 pitch factor, and 40 mm total collimation width. A tube voltage of 120 kVp with tube current modulation (varying between 296 and 495 mA) were applied. CT dose was obtained using Radimetrics Enterprise Platform™, a dose monitoring software tool using Monte Carlo simulations (Bayer HealthCare) [49]. Radimetrics calculates patient-specific absorbed dose by adjusting the CT images of a patient with Cristy & Eckerman stylized computational phantom [50] considering sex, age and size of body (diameter). To estimate the effective dose and absorbed dose in target organs of Female$_{40y\text{-}25h\text{-}75w}$, the CT acquisition parameters, a model of the 750 HD CT scanner and exposed phantom geometry were used as input to MCNPX [51, 52].

III. Results

Anthropomorphic parameters

The percentile data including 10%, 15%, 25%, 50%, 75%, 85% and 90% of height and weight, extracted from the recent version of NHANES (2011-2014) database, were combined to provide 49 height-weight grids for each age group. To restrict unrealistic body morphometries, BMI percentile data varying from 19.8 to 40.7 kg/m^2 for adult females and from 20.5 to 36.9 kg/m^2 for adult males was assigned to the height-weight bins. Once BMI data were applied on grids, a total of 249 grids for females and 230 grids for males felt within acceptable BMI range. Figure 3 display the height-weight grids of this library in seven age groups for females and males, respectively. Different categories of weight status include underweight individuals with BMI below 18.5, healthy people with BMI within the range 18.5–24.9, overweight, obese and morbidly obese within BMI ranges 25.0–29.9, 29.9–39.9 and exceeding 40.0, respectively [20]. By combining weights and heights in this library, 33.7% of female models fall into the normal BMI category, 23.2% in overweight, 40.9% in obese and 2% in morbidly obese. For male models, 27.8% of phantoms fall in healthy BMI class, 34.3% in overweight and 37.8% in obese category.

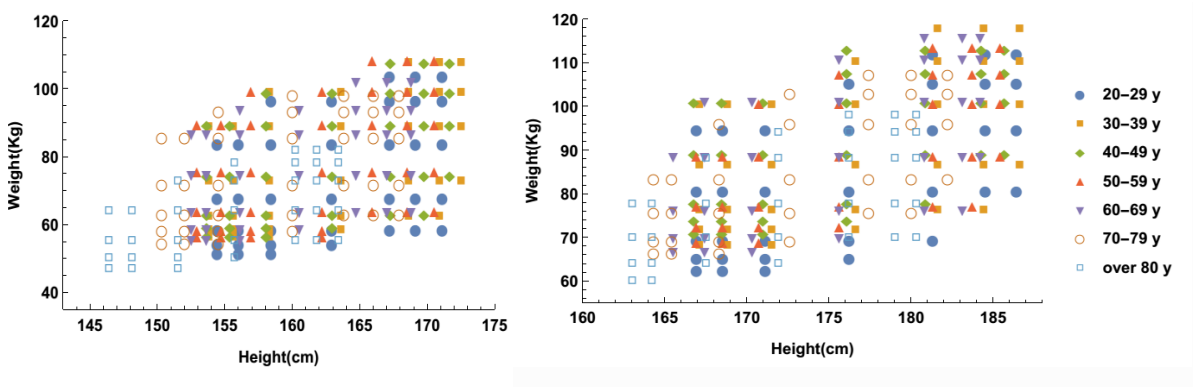


Figure 3. Targeted grid for female (left)/ male (right) phantoms in seven age groups

The FFM percent and waist circumference were considered in this library as secondary parameters. The BMI and waist circumference parameters are widely used in the characterization of obesity. As depicted in Figure 4, waist circumference increases with increasing BMI values, while the FFM is inversely proportional to BMI.

The diversity of organ masses as a function of four morphometric variables is illustrated in Table 4. These data are sampled based on anthropometric parameters of the phantoms belonging to the current library. The mean values of organ masses calculated by multiple correlation considering anthropometric variables show a deviation from ICRP reference organ masses, but still in the same order of magnitude. The difference between calculated organ masses and ICRP reference data falls within the range 0.4% - 51%.

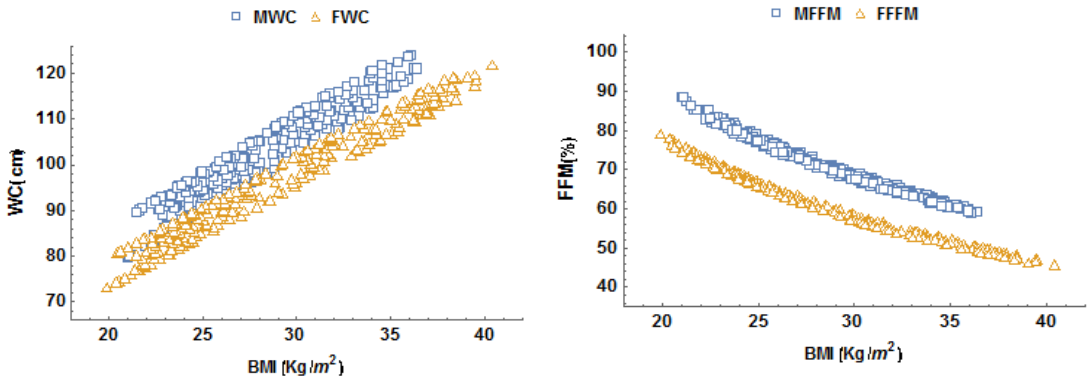


Figure 4. plots of waist circumference (left) versus bmi for males (MWC) and females (FWC) and ffm percent (right) versus bmi for males (MFFM) and females (FFFM).

Table 4. Organ masses correlated with anthropometric parameters for female and male phantoms.

Organs	Female organ mass (g)			Male organ mass (g)		
	Mean ± SD	ICRP	Range	Mean ± SD	ICRP	Range
Brain	1226.9±20.64	1300	1184.46-1265.37	1380.9±20.57	1450	1336.53-1419.57
Lung	491.83±9.2	475	471.32-502.27	621.24±17.17	600	587.15-649.48
Kidney	126.85±2.2	137	122.4-129.46	159.31±6.57	160	143.63-174.73
Spleen	149.52±12.13	130	127.32-180.99	197.01±22.04	150	153.60-250.99
Pancreas	99.672±3.618	120	92.87-105.97	123.73±3.59	140	114.87-127.55
Liver	1406.1±90.04	1400	1201.13-1601.02	1719.2±112.8	1800	1448.91-1942.58
Heart	595.95±9.618	620	571.42-605.24	879.71±14.80	840	847.05-902.42
Thyroid	17.455±0.8127	17	16.24-18.54	23.911±0.72	20	22.16-25.3
Ovary	3.64±1.988	5.5	1.872-6.76	-	-	-
Uterus	69.11±5.787	80	61.54-79.16	-	-	-
Testes	-	-	-	21.312±0.177	17.5	20.96-21.51
Prostate	-	-	-	32.289±7.33	17	21.50-42.15

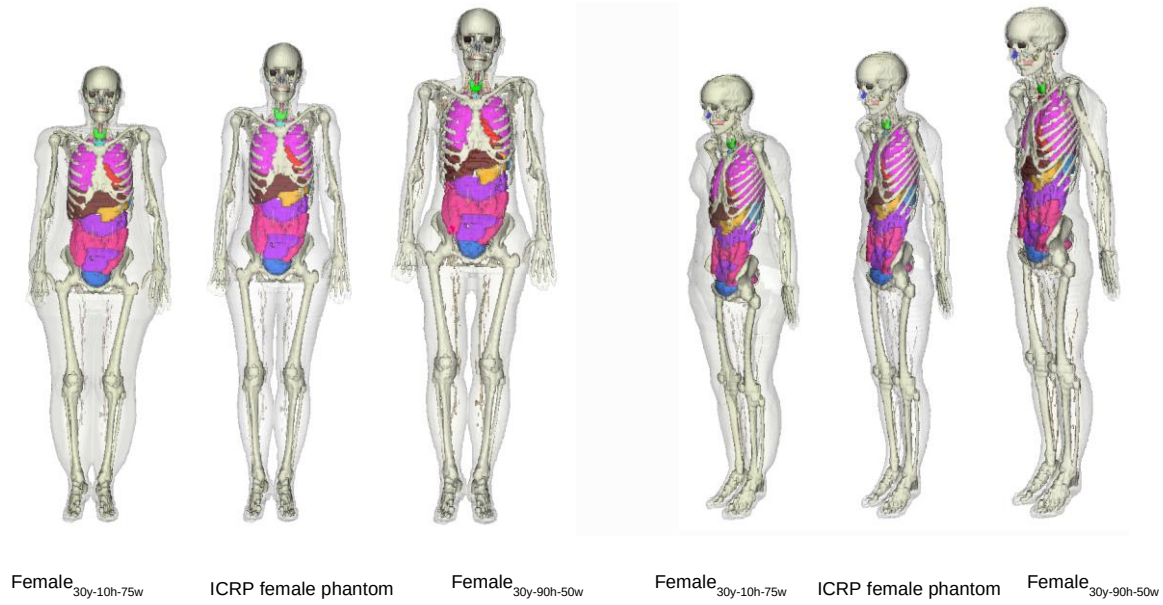


Figure 5. Frontal and rotated views of the adult female phantoms at 30-40 year age group at 10th percentile standing height and 75th percentile of weight (Female$_{30y\text{-}10h\text{-}75w}$), ICRP voxel adult female phantom, 90th percentile of height and 50th percentile of weight (Female$_{30y\text{-}90h\text{-}50w}$).

Deformed phantoms

To demonstrate the variability of the morphometry of the developed phantoms, frontal and rotated views of the ICRP female reference model compared with two female phantoms at 30-40 years age group at 10th percentile standing height and 75th percentile of body weight (Female$_{30y\text{-}10h\text{-}75w}$) and 90th percentile of standing height and 50th percentile of body weight (Female$_{30y\text{-}90h\text{-}50w}$) are depicted in Figure 5. The male ICRP reference model is compared in Figure 6 to two male phantoms at 20-30 years age group, at 10th percentile standing height and 75th percentile of body weight (Male$_{20y\text{-}10h\text{-}75w}$) and 90th percentile of standing height and 50th percentile of body weight (Male$_{20y\text{-}90h\text{-}50w}$).

We considered the diversity of internal organ masses between individuals to go one step closer to person-specific phantoms as standard models.

In this work, an initial database of 13 internal organ masses was culled using the surveyed dependence of organ masses and anthropometric parameters shown in Table 2. In Figure 7 (left), transaxial slices of two morphometrically different female phantoms with the same height but at different age groups and weight displaying obvious differences for the liver are shown.

The total body masses of the developed series were adjusted to target percentile values within 2% and waist circumference within 3% of target values. Evaluation of phantom anatomies was performed by scaling 17 organ masses correlated with morphometric parameters to agree within ±5% of target values. However, some cases have shown a larger deviation from the target data because of interpolation errors for small organs and the overlap correction of adjacent organs. The deformation process was implemented on a PC with Intel® Xeon® Processor of 2.4 GHz. The computational time required for deforming the internal organs of the phantom is less than 2 minutes. Adjusting all of the anthropometric parameters takes in average about 20 minutes depending on the amount of fat mass requiring amendment.

Monte Carlo-based dosimetry calculations

1) Internal dose from ^{18}F-FDG

Absorbed dose to organs was calculated for three anthropomorphically different phantoms in both genders including female phantoms at age group 30-40 years representing Female$_{30y\text{-}10h\text{-}10w}$ (10th height and weight percentiles) and Female$_{30y\text{-}90h\text{-}90w}$ (90th height and weight percentiles) as well as ICRP female reference phantom

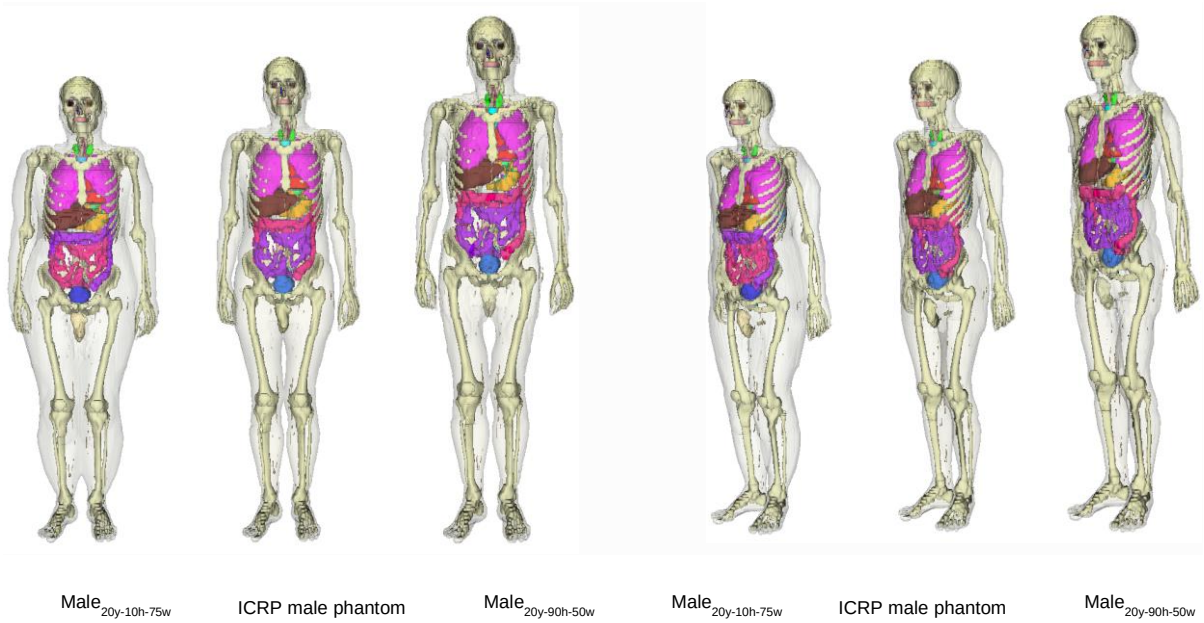


Figure 6. Frontal and rotated views of adult male phantoms at 20-30 year age groups, at 10th percentile standing height and 75th percentile of weight (Male$_{20y\text{-}10h\text{-}75w}$), ICRP voxel male phantom, 90th percentile of height and 50th percentile of weight (Male$_{20y\text{-}90h\text{-}50w}$).

They were also calculated for male phantoms in age group 30-40 years Male$_{30y\text{-}10h\text{-}15w}$ (10th height and 15th weight percentiles) at age group 50-60 years Male$_{50y\text{-}75h\text{-}50w}$ (75th height and 50th weight percentiles) in addition to ICRP male reference phantom. The simulation was designed for a whole body ^{18}F-FDG PET/CT protocol. Absorbed doses to six phantoms with different sex, age and anthropometric parameters are reported in Table 5. The absorbed dose per unit administered activity for the ICRP female reference phantom is about 35% higher than that of the obese female phantom (Female$_{30y\text{-}90h\text{-}90w}$). It is about 8.34% less in comparison with the thin phantom (Female$_{30y\text{-}10h\text{-}10w}$). The absorbed dose per unit administered activity for the ICRP male reference phantom is about 14.5% higher than the male obese phantom (Male$_{50y\text{-}75h\text{-}50w}$). It is about 1.5% less than that of the thin phantom (Male$_{30y\text{-}10h\text{-}15w}$). The total absorbed dose increases by increasing the body weights since the injected activity is proportional to patient weights (Table 5).

Table 5. Summary of ^{18}F-FDG absorbed doses for female phantom at second age group and 10th-10th height and weight percentiles, respectively (Female$_{30y\text{-}10h\text{-}10w}$), ICRP reference female phantom ICRPfemale, and female phantom at second age group and 90th-90th height and weight percentiles, respectively (Female$_{30y\text{-}90h\text{-}90w}$). Same as above for the male phantom at second age group and 10th-15th height and weight percentiles, respectively (Male$_{30y\text{-}10h\text{-}15w}$), ICRP reference male phantom ICRPmale, and male phantom at fourth age group and 75th-50th height and weight percentiles, respectively (Male$_{50y\text{-}75h\text{-}50w}$). The differences between the ICRP reference phantom and other phantoms are also shown.

Phantom ID	Weight (kg)	Height (cm)	BMI (kg/m^2)	Absorbed dose per unit administered activity		Absorbed dose for our protocol (mGy)
				(mGy/MBq)	Difference (%)	
Female$_{30y\text{-}10h\text{-}10w}$	54.75	153.8	23.145	1.51E-02	+8.34	2.89
ICRP$_{female}$	60	163.0	22.58	1.39E-02	-	2.92
Female$_{30y\text{-}90h\text{-}90w}$	106.3	172.5	35.93	9.04E-03	-35.0	3.16
Male$_{30y\text{-}10h\text{-}15w}$	72.2	167.1	25.88	1.20E-02	+1.5	3.04
ICRP$_{male}$	73	176.0	23.63	1.18E-02	-	3.02
Male$_{50y\text{-}75h\text{-}50w}$	89.2	181.3	27.23	1.01E-02	-14.5	3.16

The effective doses and absorbed doses per unit administered activity in target organs for female and male phantoms are illustrated in Figure 8. The five organs receiving the highest absorbed doses in the simulated phantoms are the heart, bladder, brain, liver and prostate for the male (uterus for female) phantom. The differences of absorbed doses in target organs between the ICRP female reference phantom and the habitus-dependent phantoms vary from -39% to 55.8%. They range between -37% and 28.7% for the male ICRP reference phantom. The differences of organ absorbed dose among phantoms depends on differences in body structures and organ masses due to the different sex, age, and anthropomorphic characteristics. From the radiation protection standpoint, the effective dose as a single metric provides a practical information to compare different radiation exposure scenarios. The effective dose differences (in mSv/MBq) between ICRP female reference model and $Female_{30y\text{-}90h\text{-}90w}$ is about -15.6% and about 4.8% with $Female_{30y\text{-}10h\text{-}10w}$. For the ICRP male reference phantom, this difference is about -3.2% for $Male_{50y\text{-}75h\text{-}50w}$ and 7.6% with $Male_{30y\text{-}10h\text{-}15w}$.

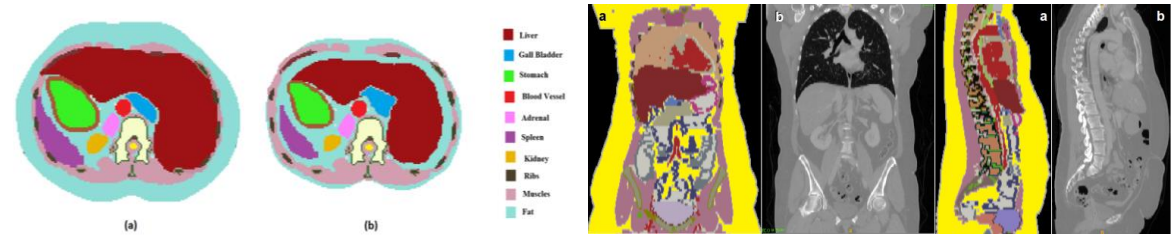


Figure 7. (right)Transaxial views of female phantoms displaying liver volume differences showing (a) 85th percentile of height and weight at age group 20-30 years and (b) 85th percentile of height and 25th percentile of weight at age group 60-70 years. (left) Coronal and sagittal views of (a) Female40y-25h-75w phantom and (b) patient CT images.

2) External dose from CT examination

The absorbed dose in target organs of the patient extracted from Radimetrics package were compared with the results of morphometrically corresponding phantom $Female_{40y\text{-}25h\text{-}75w}$ (female phantom in age group 40-50 years, 25th percentile of height and 75th percentile of weight) under the same CT scanning conditions. Coronal and sagittal slices of patient CT images and $Female_{40y\text{-}25h\text{-}75w}$ phantom are shown in Figure 7 (right).

The comparison of absorbed doses for 8 important organs and effective doses estimated by Radimetrics and our Monte Carlo calculations is illustrated in Figure 9. In most organs, the absorbed doses estimated by Radimetrics agree well with simulated results of Female40y-25h-75w with an average bias of about 16.6% (range 1.35% - 43%). The lungs present the largest deviation (about 43%). Radimetrics reported an effective dose of 17.51 mSv whereas our simulations using $Female_{40y\text{-}25h\text{-}75w}$ resulted in an effective dose of 14.35 mSv.

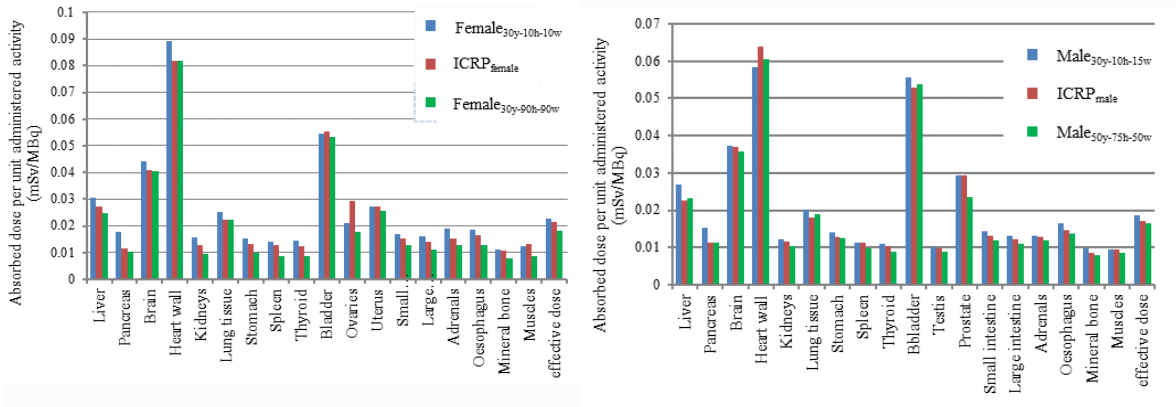


Figure 8. Effective doses and absorbed doses per unit administered activity (mSv/MBq) in target organs for female phantoms (left) at second age group and 10th-10th height and weight percentiles ($Female_{30y\text{-}10h\text{-}10w}$), 90th-90th height and weight percentiles, respectively ($Female_{30y\text{-}90h\text{-}90w}$) and the ICRP reference female phantom (ICRPfemale) and male phantoms (left) at second age group and 10th–15th height and weight percentiles, ($Male_{30y\text{-}10h\text{-}15w}$), forth age group and 75th–50th height and weight percentiles, respectively ($Male_{50y\text{-}75h\text{-}50w}$) and the ICRP reference male phantom (ICRPmale).

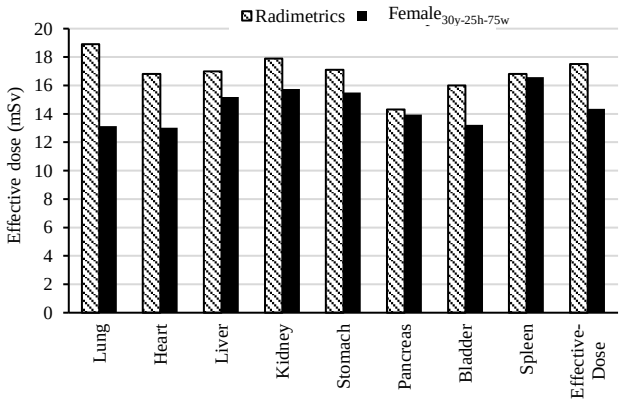


Figure 9. Comparison of organ absorbed doses and effective doses between Radimetrics and Monte Carlo simulations using $Female_{40y\text{-}25h\text{-}75w}$ phantom.

IV. Discussion

The construction of more realistic models representative of the adult population is highly desired for Monte Carlo-based simulation studies targeting a number of research applications in diagnostic and therapeutic radiology. A new series of anthropomorphic phantoms extended from the ICRP reference phantoms covering not only the diversity of anthropometric characteristics but also anatomical diversities are developed in this work. Firstly, standing height and total body mass were adjusted as the primary parameters of this library. Secondly, FFM percent was tuned by 2D scaling to make a more realistic proportion with anthropometric parameters between unscaled organs, such as muscles and bones. Although different body shapes are not defined in this work, tuning the FFM percent helps to appropriately change the body style of phantoms. In a few cases, such as 10th percentile of weight in the first age group of male phantoms, FFM percent was not applied on the reference phantom to reach the target waist circumference. By sampling from derived multiple correlations between organ masses and external anthropometric parameters (age, standing height, body weight and BMI), organ mass data does not follow a normal distribution. According to the organ mass values in the library, the brain mass increases with increasing height and decreases with increasing age. For the lung, kidney, spleen, and pancreas, it can be concluded that their masses increase by increasing the body weight and BMI without following a meaningful correlation with height and age variations. The liver and heart masses are strongly proportional to the total body mass and consequently to BMI. The behavior of the thyroid is a little different since it firstly increases and then decreases with increasing standing height. Moreover, it has a smooth increase with age and weight. As mentioned earlier, genital organs strongly depend on the age, except testis mass which increases with weight and BMI. The mass of the ovaries decreases sharply after 35 years whereas the uterus mass reaches a peak at about 45 years and decreases at higher ages. The prostate mass increases strongly with age. The masses of remaining organs, such as the muscles, bones, blood vessels, …etc. change after 3D scaling as reported by Johnson et al. [9]. The masses of most organs deviate from the mean value by about 3%-21% while it exceeds 50% for genital organs.

The internal radiation dose was estimated in the context of a whole-body ^{18}F-FDG PET protocol for morphometrically different phantoms for both genders. The calculated absorbed doses per administered activity in target organs are in agreement with those reported by the ICRP 106 [45]. The effective dose for habitus-dependent phantoms varies between 1.82 and 2.26 (mSv/100MBq) for females, and between 1.67 and 1.85 (mSv/100MBq) for males, while the ICRP 106 reports an effective dose of 2.31 (mSv/100MBq). As expected, a significant dose from ^{18}F-FDG is delivered to the heart, bladder, brain, liver, prostate for male and uterus for female. The heart, brain and liver receive a considerable dose because of their high metabolic rate and hence rapid blood supply. The accumulation of radioactive urine in the bladder not only causes a significant self-absorbed dose but also leads to a high cross-organ dose to the uterus and prostate. Overall, thin patients receive a higher internal radiation dose because the cross-irradiation between internal organs is stronger than other patients. This can be justified by the lack of subcutaneous and visceral fat which directly influences cross-organ doses.

The organ-level doses and effective dose for the CT examination calculated for the anthropomorphic fitted phantom from the developed series show acceptable dose estimation accuracy for radiation protection purposes. Since each subject has exclusive anatomical characteristics, using an independent phantom library instead of a patient-specific model may cause a deviation from the actual absorbed dose. However, it definitely provides a more accurate estimate compared to the calculations using stylized phantoms, such as the Cristy & Eckerman stylized computational phantoms [50] implemented within Radimetrics software. During the course of this study, we created a patient-specific model by segmenting a clinical CT study and compared the dose delivered to the patient-specific model, considered as reference, to results generated using the corresponding anthropomorphic phantom from our library and Radimetrics. The effective dose for the patient-specific model was estimated to be 11.72 mSv. The discrepancy between the effective dose calculated using the patient-specific model and the best fitting model from our library was 22.2%, whereas it was up to 49.2% when calculated using Radimetrics. As reported by Xie et al. [52], in addition to the anthropometric parameters, the organ-surface distance (average distance from the skin to organs) or body size shows a significant correlation with organ absorbed dose. It appears that Radimetrics overestimates the absorbed dose in target organs, particularly for the lungs.

In this work, the diversity of organ masses is considered. However, for reliable modeling, a broad database is required to derive organ masses correlated with anthropomorphic variables. In addition, defining different somatotypes for the phantoms library makes it more comprehensive. Sheldon et al. introduced three types of main somatotypes: mesomorphs who are athletically built with a low percent of body fat, ectomorphs who are underweight with a narrow skeleton frame, and endomorphs who are overweight with a pear-shaped body style [53]. In the present work, we set the FFM percent for different BMIs to consider body style in addition to the height and weight. By considering different styles in the future studies, the distribution of fat percent through the body of phantoms would be more realistic. Although the volume of intra-abdominal adipose tissue encompassed the organs is an important factor in the calculation of cross-organ dose, adjustment of this type of fat percent called visceral fat was ignored due to the lack of information, and total body mass was set by adjusting the subcutaneous fat mass.

Recent advances in deep learning are promoting a number of applications in computer vision and medical image analysis research that could be useful for constructing patient-specific models through automatic segmentation of body contours and internal organs. Considering the scarcity of large clinical databases and time-consuming classification techniques required for organs labeling, generative networks can be used for developing data-hungry deep learning algorithms. Novel unsupervised models, such as variational auto-encoders [54] or generative adversarial networks [55] have shown potential in medical image analysis and look promising for applications involving the generation of synthesized medical images to fulfill the requirement of large training datasets, e.g. auto-segmentation [56].

V. Conclusion

An algorithm was developed to consider the diversity of organ masses along with the morphometric parameters to construct a library by automatic remodeling the voxel-based ICRP adult reference phantoms. Data on 13 organ masses is culled based on information from autopsies and diagnostic examinations. Using the specific anthropometric data of each individual, it is possible to derive organ masses data and automatically construct habitus-specific phantoms according to the specific input parameters.

By using habitus-dependent anthropomorphic libraries, the calculation of absorbed doses for individuals exposed to external or internal radiation is likely to be more accurate by considering the anthropomorphic and anatomical diversity among the population. The move towards subject-specific phantoms is a major improvement taking advantage of the availability of habitus-dependent phantoms associated with morphometric parameters and classified in different somatotypes.

Acknowledgements

This work was supported by the Swiss National Science Foundation under grant SNSF 320030_176052, the Swiss Cancer Research Foundation under Grant KFS-3855-02-2016 and Iran's Ministry of Science and Technology.

References

[1] H. Zaidi, "Relevance of Accurate Monte Carlo Modeling in Nuclear Medical Imaging.,"*Med Phys,* vol. 26, pp. 574-608, 1999.

[2] H. Zaidi and X. G. Xu, "Computational Anthropomorphic Models of the Human Anatomy: The Path to Realistic Monte Carlo Modeling in Medical Imaging.,"*Annu Rev Biomed Eng,* vol. 9, no. 1, pp. 471-500, 2007.

[3] X. G. Xu, "An Exponential Growth of Computational Phantom Research in Radiation Protection, Imaging, and Radiotherapy: A Review of the Fifty-Year History.,"*Phys Med Biol,* vol. 59, no. 18, pp. R233-R302, 2014.

[4] J. Zhang, Y. H. Na, P. F. Caracappa, and X. G. Xu, "Rpi-Am and Rpi-Af, a Pair of Mesh-Based, Size-Adjustable Adult Male and Female Computational Phantoms Using Icrp-89 Parameters and Their Calculations for Organ Doses from Monoenergetic Photon Beams.,"*Phys Med Biol,* vol. 54, no. 19, pp. 5885-5908, 2009.

[5] Y. H. Na, B. Zhang, J. Zhang, P. F. Caracappa, and X. G. Xu, "Deformable Adult Human Phantoms for Radiation Protection Dosimetry: Anthropometric Data Representing Size Distributions of Adult Worker Populations and Software Algorithms.,"*Phys Med Biol,* vol. 55, no. 13, pp. 3789–3811, 2010.

[6] ICRP, *Report of the Task Group on Reference Man. Icrp Publication 23.* New York: Pergamon Press, 1975.

[7] M. Cristy, "Mathematical Phantoms Representing Children of Various Ages for Use in Estimates of Internal Dose.," Oak Ridge National Laboratory, Oak Ridge, TN, ORNL/NUREG/TM-367, 1980.

[8] C. Lee, D. Lodwick, J. L. Williams, and W. E. Bolch, "Hybrid Computational Phantoms of the 15-Year Male and Female Adolescent: Applications to Ct Organ Dosimetry for Patients of Variable Morphometry.,"*Med Phys,* vol. 35, no. 6, pp. 2366-2382, 2008.

[9] P. B. Johnson, S. R. Whalen, M. Wayson, B. Juneja, C. Lee, and W. E. Bolch, "Hybrid Patient-Dependent Phantoms Covering Statistical Distributions of Body Morphometry in the U.S. Adult and Pediatric Population.,"*Proceedings of the IEEE,* vol. 97, no. 12, pp. 2060-2075, 2009.

[10] C. Lee, D. Lodwick, J. L. Williams, and W. E. Bolch, "Hybrid Computational Phantoms of the 15-Year Male and Female Adolescent: Applications to Ct Organ Dosimetry for Patients of Variable Morphometry,"*Medical Physics,* vol. 35, no. 6, pp. 2366-2382, 2008.

[11] Y. Na, J. Zhang, X. Xu, B. Han, and P. Caracappa, "We-E-Brd-08: Next-Generation Deformable Patient Modeling for Monte Carlo Assessment of Organ Doses,"*Medical Physics,* vol. 36, no. 6, pp. 2783-2783, 2009.

[12] D. Broggio *et al.,* "Construction of an Extended Library of Adult Male 3d Models: Rationale and Results.,"*Phys Med Biol,* vol. 56, no. 23, pp. 7659-7692, 2011.

[13] G. L. de la Grandmaison, I. Clairand, and M. Durigon, "Organ Weight in 684 Adult Autopsies: New Tables for a Caucasoid Population.,"*Forensic science international,* vol. 119, no. 2, pp. 149-154, 2001.

[14] I. Clairand, L. G. Bouchet, M. Ricard, M. Durigon, M. Di-Paola, and B. Aubert, "Improvement of Internal Dose Calculations Using Mathematical Models of Different Adult Heights.,"*Physics in medicine and biology,* vol. 45, no. 10, pp. 2771-2785, 2000.

[15] V. F. Cassola, V. J. d. M. Lima, R. Kramer, and H. J. Khoury, "Fash and Mash: Female and Male Adult Human Phantoms Based on Polygon Mesh Surfaces: I. Development of the Anatomy.,"*Phys Med Biol,* vol. 55, no. 1, pp. 133-162, 2010.

[16] V. F. Cassola, F. M. Milian, R. Kramer, C. A. de Oliveira Lira, and H. J. Khoury, "Standing Adult Human Phantoms Based on 10th, 50th and 90th Mass and Height Percentiles of Male and Female Caucasian Populations.,"*Phys Med Biol,* vol. 56, no. 13, pp. 3749-3772, 2011.

[17] W. P. Segars *et al.,* "Population of Anatomically Variable 4d Xcat Adult Phantoms for Imaging Research and Optimization.,"*Med Phys,* vol. 40, no. 4, pp. 043701-11, 2013.

[18] W. P. Segars, G. Sturgeon, S. Mendonca, J. Grimes, and B. M. Tsui, "4d Xcat Phantom for Multimodality Imaging Research.,"*Med Phys,* vol. 37, no. 9, pp. 4902-4915, 2010.

[19] Y. Chen, R. Qiu, C. Li, Z. Wu, and J. Li, "Construction of Chinese Adult Male Phantom Library and Its Application in the Virtual Calibration of in Vivo Measurement,"*Physics in medicine and biology,* vol. 61, no. 5, pp. 2124-44, 2016.

[20] A. M. Geyer, S. O'Reilly, C. Lee, D. J. Long, and W. E. Bolch, "The Uf/Nci Family of Hybrid Computational Phantoms Representing the Current Us Population of Male and Female Children, Adolescents, and Adults-Application to Ct Dosimetry.,"*Phys Med Biol,* vol. 59, no. 18, pp. 5225-5242, 2014.

[21] Y. S. Yeom, J. H. Jeong, M. C. Han, and C. H. Kim, "Tetrahedral-Mesh-Based Computational Human Phantom for Fast Monte Carlo Dose Calculations.,"*Phys Med Biol,* vol. 59, no. 12, pp. 3173-3185, 2014.

[22] H. G. Menzel, C. Clement, and P. DeLuca, "Icrp Publication 110. Realistic Reference Phantoms: An Icrp/Icru Joint Effort. A Report of Adult Reference Computational Phantoms,"*Ann ICRP,* vol. 39, no. 2, pp. 1-164, 2009.

[23] C. Fryar, Q. Gu, C. Ogden, and K. Flegal, "Anthropometric Reference Data for Children and Adults: United States, 2011–2014. National Center for Health Statistics. ,"*Vital Health Stat,* vol. 3, no. 39, 2016.
[24] M. Zankl *et al.*, "Computational Phantoms of the Icrp Reference Male and Reference Female," 2008: these Proceeding.
[25] C. L. Ogden, M. D. Carroll, C. D. Fryar, and K. M. Flegal, "Prevalence of Obesity among Adults and Youth: United States, 2011–2014,"*NCHS data brief,* vol. 219, no. 219, pp. 1-8, 2015.
[26] S. Meeuwsen, G. W. Horgan, and M. Elia, "The Relationship between Bmi and Percent Body Fat, Measured by Bioelectrical Impedance, in a Large Adult Sample Is Curvilinear and Influenced by Age and Sex,"*Clinical nutrition,* vol. 29, no. 5, pp. 560-6, 2010.
[27] Y. Schutz, U. U. Kyle, and C. Pichard, "Fat-Free Mass Index and Fat Mass Index Percentiles in Caucasians Aged 18-98 Y,"*International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity,* vol. 26, no. 7, pp. 953-60, 2002.
[28] S. B. Heymsfield, M. Heo, D. Thomas, and A. Pietrobelli, "Scaling of Body Composition to Height: Relevance to Height-Normalized Indexes,"*The American journal of clinical nutrition,* vol. 93, no. 4, pp. 736-40, 2011.
[29] M. U. S. D. o. H. Centers for Disease Control and Prevention (CDC) and National Center for Health Statistics (NCHS):National Health and Nutrition Examination Survey Data. Hyattsville, 2003.
[30] S. R. Bozeman, D. C. Hoaglin, T. M. Burton, C. L. Pashos, R. H. Ben-Joseph, and C. S. Hollenbeak, "Predicting Waist Circumference from Body Mass Index,"*BMC medical research methodology,* vol. 12, p. 115, 2012.
[31] R. Mandal, A. G. Loeffler, S. Salamat, and M. K. Fritsch, "Organ Weight Changes Associated with Body Mass Index Determined from a Medical Autopsy Population,"*The American journal of forensic medicine and pathology,* vol. 33, no. 4, pp. 382-389, 2012.
[32] A. Sheikhazadi *et al.*, "Study of the Normal Internal Organ Weights in Tehran's Population,"*Journal of forensic and legal medicine,* vol. 17, no. 2, pp. 78-83, 2010.
[33] D. K. Molina and V. J. DiMaio, "Normal Organ Weights in Women: Part Ii-the Brain, Lungs, Liver, Spleen, and Kidneys,"*The American journal of forensic medicine and pathology,* vol. 36, no. 3, pp. 182-7, 2015.
[34] D. K. Molina and V. J. DiMaio, "Normal Organ Weights in Men: Part Ii-the Brain, Lungs, Liver, Spleen, and Kidneys,"*The American journal of forensic medicine and pathology,* vol. 33, no. 4, pp. 368-72, 2012.
[35] D. K. Molina and V. J. DiMaio, "Normal Organ Weights in Women: Part I-the Heart,"*The American journal of forensic medicine and pathology,* vol. 36, no. 3, pp. 176-81, 2015.
[36] D. K. Molina and V. J. DiMaio, "Normal Organ Weights in Men: Part I-the Heart,"*The American journal of forensic medicine and pathology,* vol. 33, no. 4, pp. 362-7, 2012.
[37] Q. He *et al.*, "Smaller Organ Mass with Greater Age, except for Heart,"*Journal of applied physiology,* vol. 106, no. 6, pp. 1780-4, 2009.
[38] S. B. Heymsfield, D. Gallagher, L. Mayer, J. Beetsch, and A. Pietrobelli, "Scaling of Human Body Composition to Stature: New Insights into Body Mass Index,"*The American journal of clinical nutrition,* vol. 86, no. 1, pp. 82-91, 2007.
[39] H. A. Perven, A. S. Nurunnabi, S. Ara, and M. U. Jahan, "Cadaver Study of the Volume of the Ovary in Bangladeshi Women,"*Bangladesh Medical Research Council bulletin,* vol. 40, no. 1, pp. 15-7, 2014.
[40] T. W. Kelsey, E. Ginbey, M. M. Chowdhury, L. E. Bath, R. A. Anderson, and W. H. Wallace, "A Validated Normative Model for Human Uterine Volume from Birth to Age 40 Years,"*PloS one,* vol. 11, no. 6, p. e0157375, 2016.
[41] S. J. Zhang *et al.*, "Relationship between Age and Prostate Size,"*Asian journal of andrology,* vol. 15, no. 1, pp. 116-20, 2013.
[42] S. J. Xia, X. X. Xu, J. B. Teng, C. X. Xu, and X. D. Tang, "Characteristic Pattern of Human Prostatic Growth with Age,"*Asian journal of andrology,* vol. 4, no. 4, pp. 269-71, 2002.
[43] A. Chouker *et al.*, "Estimation of Liver Size for Liver Transplantation: The Impact of Age and Gender,"*Liver transplantation : official publication of the American Association for the Study of Liver Diseases and the International Liver Transplantation Society,* vol. 10, no. 5, pp. 678-85, 2004.
[44] T. W. Kelsey *et al.*, "Ovarian Volume Throughout Life: A Validated Normative Model,"*PloS one,* vol. 8, no. 9, p. e71465, 2013.
[45] ICRP, "Radiation Dose to Patients from Radiopharmaceuticals. Addendum 3 to Icrp Publication 53. Icrp Publication 106.,"*Annals of the ICRP,* vol. 38, no. 1-2, pp. 1-197, 2008.
[46] T. Xie, W. E. Bolch, C. Lee, and H. Zaidi, "Pediatric Radiation Dosimetry for Positron-Emitting Radionuclides Using Anthropomorphic Phantoms,"*Med Phys,* vol. 40, no. 10, p. 102502, 2013.
[47] T. Xie, N. Kuster, and H. Zaidi, "Effects of Body Habitus on Internal Radiation Dose Calculations Using the 5-Year-Old Anthropomorphic Male Models,"*Phys Med Biol,* vol. 62, no. 15, pp. 6185-6206, 2017.

[48] W. E. Bolch, K. F. Eckerman, G. Sgouros, and S. R. Thomas, "Mird Pamphlet No. 21: A Generalized Schema for Radiopharmaceutical Dosimetry--Standardization of Nomenclature,"*Journal of nuclear medicine : official publication, Society of Nuclear Medicine,* vol. 50, no. 3, pp. 477-84, 2009.

[49] Bayer HealthCare. "Radimetrics Enterprise Platform: Dose Management Solution." Bayer HealthCare. http://www.radiologysolutions.bayer.com/products/ct/dosemanagement/rep/. (accessed 2016).

[50] M. Cristy and K. Eckerman, "Specific absorbed fractions of energy at various ages from internal photon sources. I methods, II one year old, III five year old, IV ten year old, V fifteen year old male and adult female, VI new-born and VII adult male," Oak Ridge Nat. Lab., Oak Ridge, TN, USA, Rep. ORNL/TM 8381/V1-V7, 1987.

[51] A. Akbarzadeh, M. R. Ay, H. Ghadiri, S. Sarkar, and H. Zaidi, "Measurement of Scattered Radiation in a Volumetric 64-Slice Ct Scanner Using Three Experimental Techniques,"*Physics in medicine and biology,* vol. 55, no. 8, pp. 2269-80, 2010.

[52] T. Xie, P.-A. Poletti, A. platon, C. Becker, and H. Zaidi, "Assessment of Ct Dose to the Fetus and Pregnant Female Patient Using Patient-Specific Computational Models,"*European Radiology*, 2017.

[53] W. H. Sheldon, S. S. Stevens, and W. B. Tucker, *The Varieties of Human Physique: An Introduction to Constitutional Psychology*. Harper, 1940.

[54] D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes,"*ARXIV,* p. eprint arXiv:1312.6114, 2014.

[55] I. Goodfellow *et al.*, "Generative Adversarial Networks.,"*ARXIV,* p. eprint arXiv:1406.2661, 2014.

[56] T. Neff, C. Payer, D. Stern, and M. Urschler, "Generative Adversarial Network Based Synthesis for Supervised Medical Image Segmentation," in *OAGM & ARW Joint Workshop*, Vienna, Austria, 2017, pp. 140-145.

Chapter 4

Construction of patient-specific computational models for organ dose estimation in radiological imaging

Tianwu Xie*, Azadeh Akhavanallaf*, and Habib Zaidi

Contribution: Study conception and design, methodological development, computer programming, analysis and interpretation of results, manuscript preparation.

Med Phys, Vol. 46, No.4, pp 1519-1522 (2019)

* Authors contributed equally

Abstract

Purpose: Diagnostic imaging procedures require optimization depending on the medical task at hand, the apparatus being used and patient physical and anatomical characteristics. The assessment of the radiation dose and associated risks plays a key role in safety and quality management for radiation protection purposes. In this work, we aim at developing a methodology for personalized organ-level dose assessment in x-ray CT imaging.

Methods: Regional voxel models representing reference patient-specific computational phantoms were generated through image segmentation of CT images for four patients. The best fitting anthropomorphic phantoms were selected from a previously developed comprehensive phantom library according to patient's anthropometric parameters, then registered to the anatomical masks (skeleton, lung and body contour) of patients to produce a patient-specific whole-body phantom. Well established image registration metrics including Jaccard's coefficients for each organ, organ mass, body perimeter, organ-surface distance and effective diameter are compared between the reference patient model, registered model and anchor phantoms. A previously-validated Monte Carlo code is utilized to calculate the absorbed dose in target organs along with the effective dose delivered to patients. The calculated absorbed doses from the reference patient models are then compared with the produced personalized model, anchor phantom and those reported by commercial dose monitoring systems.

Results: The evaluated organ-surface distance and body effective diameter metrics show a mean absolute difference between patient regional voxel models, serving as reference, and patient-specific models around 4.4% and 4.5%, respectively. Organ-level radiation doses of patient-specific models are in good agreement with those of the corresponding patient regional voxel models with a mean absolute difference of 9.1%. The mean absolute difference of organ doses for the best fitting model extracted from the phantom library and Radimetrics™ commercial dose tracking software are 15.5% and 41.1%, respectively.

Conclusion: The results suggest that the proposed methodology improves the accuracy of organ-level dose estimation in CT, especially for extreme cases (high BMI and large skeleton). Patient-specific radiation dose calculation and risk assessment can be performed using the proposed methodology for both monitoring of cumulative radiation exposure of patients and epidemiological studies. Further validation using a larger database is warranted.

I. Introduction

The use of x–ray computed tomography (CT) in the clinic has skyrocketed in the last decade and has been accounted as a major contributor to total radiation exposure of the population in developed countries [1, 2]. As the workhorse of radiology and medical imaging, CT scans can provide fast and accurate diagnostic medical images, guide surgical interventions and help planning therapeutic procedures [3]. Despite the overwhelming medical benefits of CT, there are concerns about potential cancer risks owing to the utilization of ionizing radiation. Overall, the radiation risks from CT imaging to patients are small as the absorbed doses commonly range from 1 mSv to 30 mSv depending on the CT study type [4]. However, owing to the large number of CT examinations performed annually (more than 100 million worldwide) [5], even small risks may translate into a large-scaled number of future cancers. In light of these risks, accurate estimation of the absorbed dose profile and associated risk factors for the exposed patients in CT examinations is necessary [6, 7]. Different approaches have been adopted to estimate the absorbed dose to patients from CT scans, including experimental measurements using dosimeters embedded witin physical anthropomorphic phantoms and Monte Carlo calculations using computational models. However, these approaches inherently bear a number of limitations including the difficulty of matching physical phantoms to the location of internal organs within the patient's body, the heavy workload involved for constructing patient-specific computational models and the inherent assumptions in measurements and simulation setups, which might contribute significant uncertainties to the estimated absorbed dose. A patient-specific phantom represents a realistic model enabling accurate estimation of organ-level dose; however, the segmentation of patient images is time consuming and not feasible for clinical routine applications. A potential alternative for person-specific organ dose estimation is to use a library of computational models where habitus-specific phantoms could serve as alternative models covering various anthropometric and anatomical characteristics of patients [8]. Several habitus-dependent phantom series have been developed to perform patient-specific dose estimation by matching anthropometric characteristics of patients, such as gender, age, height and body weight [9-11]. Stepusin et al. [12] suggested to match patient's data to a computational phantom from a predefined library using height-and-weight matching for patient-specific CT dosimetry. The construction of patient-specific models from regional CT images is another alternative for patient-specific dosimetry, which was adopted in a number of studies by mapping the segmented model of patient CT images to a template anatomy through a deformable registration process [13, 14]. However, this is not practical for routine clinical usage owing to the labor-intensive manual segmentation process. Kalender et al. [15] proposed to construct patient-specific whole body models from regional CT images using a simple protocol-based appending of the scan range to a reference phantom to take out-of direct field-of-view scattered radiation and over-scanning effects into account. Similar to the above mentioned method, this technique also suffers from time-consuming manual organ segmentation required for dosimetry calculations. Gao et al. [16] estimated organ doses for a large number of pediatric patients using patient-specific information implemented into VirtualDose™ CT dose calculation software. The Radimetrics™ commercial dose tracking software (Bayer HealthCare) [17] calculates patient-specific absorbed doses by matching CT images of each patient with Cristy & Eckerman stylized computational phantoms [18], taking into account the physical and anatomical characteristics of the patient. The ImPACT CT patient dosimetry calculator estimates organ-level absorbed doses and effective dose based on spreadsheet tools and adult stylized phantoms [19]. Although this approach is practical, it is impaired by the dosimetric uncertainties resulting from the large differences between the anatomy of patients and simulated computational phantoms (categorized by gender and age) in addition to the inherent uncertainties associated with protocol-based mapping of the scan location on the computational models [20]. Since Monte Carlo calculations using patient-specific models are commonly considered as reference for organ dose estimation from diagnostic imaging procedures [21], the implementation of an easy to use and reliable framework enabling to estimate patient-specific organ dose for individual patients in clinical setting is highly desirable.

In this work, we propose a methodology for constructing patient-specific computational models based on deformable registration of patient CT regional images on a habitus-dependent anchor phantom. Unlike previous works that simply append the scan range to an existing reference anthropomorphic phantom, which requires labor-intensive and time-consuming manual segmentation or definition of patient's organs inside the appended region, we employ the structural deformation of the best-fitting phantom from an existing large library of computational models through automated registration to estimate the patient-specific model. We adopted three computational

models for patient-specific CT dose estimation: the regional patient model developed by segmenting a series of CT images, the best fitting/matching anchor model selected from a phantom library based on sex, age, height-and-weight and patient-specific computational model constructed through image registration. Patient-specific radiation doses are calculated using the three computational models used as input to the N-Particle eXtended (MCNPX) Monte Carlo code. The results obtained using the anchor phantom and registered model are compared to the regional patient model serving as reference to investigate the influence of the computational model on the accuracy of patient-specific radiation dose estimation.

II. Materials and methods

Patient's regional voxel model based on CT image segmentation

The institutional ethics committee approved this retrospective study. Written informed consent was waived. CT images of four patients with different gender, age and physical characteristics referred to Geneva University Hospital for diagnostic CT examinations were included in this work. Semi-automated image segmentation was performed using the 3DSlicer software [22] enabling to identify 10 organs/tissues, including the lungs, heart wall, liver, kidneys, spleen, stomach wall, pancreas, gall bladder, urinary bladder and skeleton. The segmented organs were validated by an experienced radiologist and integrated in a voxel matrix to produce a patient regional computational model. The chemical compositions and material density for each organ were assigned according to the ICRP report 89 [23].

Patient-specific computational phantom

In previous work, we reported on the construction of a phantom library [24] extended from ICRP reference models covering different body morphometries, consisting of about 230 male and 249 female voxel adult phantoms scaled to specific age, height and weight grids based on the NHANES (2011-2014) database [25]. In these series, anatomical diversity, specifically organ masses, is implemented using a multi-correlation model to estimate organ masses based on gender, age, height, weight and BMI.Therfore, the best-fitting adult phantom is selected from the extended library to match the patient's anthropometric and anatomical characterisitics. For the pediatric models, the anchor phantoms were generated by scaling the male and female ICRP adult reference phantoms to match the height and weight of actual patients. Thereafter, anatomical masks for the skeleton, lung and body contour were generated from patients' regional CT images using auto-segmentation algorithms. Subsequently, the whole body anchor phantom is deformably registered to the patient anatomical mask model using automatic affine registration to produce a new whole body personalized computational phantom with well-defined anatomical structures, matching patient images obtained from CT examinations. Image registration was performed using the Insight Toolkit (ITK) [26]. The registration was performed in two steps: In the first step, the voxelized patient regional model was registered to the anchor phantom through automatic affine registration and the affine matrix warping the patient's regional model to the anchor phantom calculated. In the second step, the inverse affine matrix is applied to the anchor phantom to produce a new personalized computational model. The registration algorithm uses the regular step gradient descent implemented within the ITK software package and the mean squares metric between two images. The resulting model after registration is referred to as "patient-specific model" and includes 140 identified organs. Figure 2 shows representative patient-specific computational phantoms along with the regional model of corresponding patients as well as the selected best-fitting models.

Dosimetry calculations

Computational models are commonly coupled with Monte Carlo techniques for dosimetry calculations through full simulation of the CT scanner and parameters used by the scanning protocol under which the patients were examined. The studies were acquired on the GE 750HD CT scanner (GE Healthcare, Waukesha, WI). The geometry of the system was accurately modelled and validated against experimental measurements as described in previous work [27]. The patient-specifc acquisition parameters, including the table speed, revolution time, pitch factor, total collimation width, tube voltage and modulated tube current , extracted from the DICOM headers, were modeled in this simulation setup. The examined body part was defined automatically through mapping the skeletal mask obtained from patient CT images to the whole body computational phantoms. Obtaining the complete tube

current modulation profile was not possible owing to the lack of sufficient information (proprietary data undisclosed by the manufacturer). We therefore extracted the longitudinal tube current modulation from the DICOM header file of patient's CT images and didn't consider angular tube current modulation [16]. The absorbed doses to predefined target organs were calculated using a previously validated C++ computer code [27, 28]. The simulations of each helical CT scan were repeated six times with x-ray tube starting at angles differing by 60 degrees since the actual tube starting angles in the actual examinations was unknown. The effective dose was estimated according to ICRP report 103 [29]:

$$ED = \sum_T \omega_T \sum_R \omega_R D_{R,T}, \qquad (1)$$

where E is the effective dose, ω_R is the radiation weighting factor for radiation type R, $D_{R,T}$ is the contribution of radiation type R to the absorbed dose, and ω_T is the tissue weighting factor for organ or tissue T reflecting its relative radiation sensitivity.

Subsequently, clinical CT images of the patients were imported into Radimetrics™ dose tracking software [17] for calculation of organ absorbed doses and effective dose. Radimetrics™ collects CT scans directly from the hospital's picture archiving and communication system and matches patient images with Cristy & Eckerman stylized computational phantoms [18] according to age, gender and body size. The software extracts scanning parameters (tube voltage, mAs, scan range, etc.) from the DICOM files' header information and calculates overnight patient-specific absorbed dose at the organ level through Monte Carlo simulations.

In this work, the results obtained from Monte Carlo simulations of the patient's regional model were considered as reference to which dose profiles calculated using computational models and dose monitoring software were compared.

Quantitative analysis

To investigate the impact of anatomical metrics of computational models on radiation dose estimation, habitus-dependent parameters of the computational models were compared using established metrics including the Jaccard's coefficient for each organ, organ mass, mean body perimeter, organ-surface distance and mean body effective diameter. The similarity between the patient regional model and patient-specific model and the selected best-fitting model was evaluated through the Jaccard's coefficient:

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \qquad (2)$$

where ***A*** refers to the organ volume of the patient's regional model whereas B refers to the volume of the same in the patient-specific model or the selected best-fitting model. This metric enables the assessment of organs overlap between the two investigated models. The mean body perimeter was determined by the average outer perimeter of the patient in the scan range. Surface-organ distance is defined as the average distance from the skin to the organ in all slices. The body effective diameter is defined based on lateral (LAT) and anterior-posterior (AP) dimentions:

$$effective\ diameter = \sqrt{AP \times LAT} \qquad (3)$$

III. Results

Computational models

Figure 1 shows the segmented regional model of the considered patients along with the original CT images. The best-fitting model for adult patients were selected from the extended phantom library [24] based on age, gender and height-weight matching, while for the pediatric patients, the best-fitting models were constructed by deforming the ICRP reference model to reach the height-weight target values. The anthropomorphic characteristics of the 4 patients included in this study are summarized in Table 1. Subsequently, the selected best-fitting model was registered to the patient's regional model to generate a patient-specific model. Figure 2 shows representative patient-specific computational models together with regional computational models of the corresponding patients and best-fitting phantoms.

Table 1. Anthropomorphic characteristics of the patients and the best-fitting models (matched by gender, age, height and weight).

Physical parameters	Patients				Best-fitting model			
	#1	#2	#3	#4	#1	#2	#3	#4
Sex	AF	AM	PM	PM	AF	AM	PM	PM
Age (y)	50	25	7	3	40-50	20-30	-	-
Weight (kg)	88	104	26	13	89.3	105.7	26	13
Height (cm)	160	185	131	93	158.1	184.3	131	93
BMI (kg/cm2)	34.37	30.38	15.1	15	35.77	31.22	15.1	15
Body region	Th-Ab	Th-Ab	Th-Ab	Th-Ab	WB	WB	WB	WB

AM: Adult Male, AF: Adult Female, PM: Pediatric Male, Th-Ab: Thorax-Abdomen, WB: Whole body

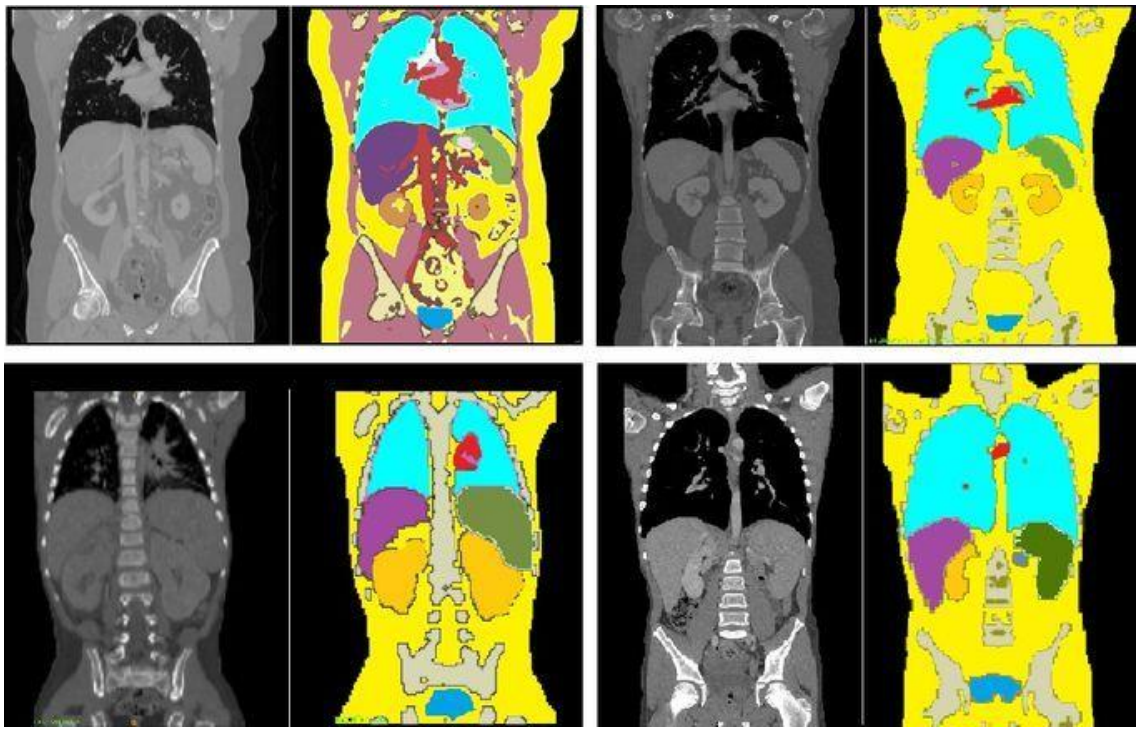


Figure 1. Representative slices showing coronal views of segmented regional patient models for patient case #1 (top left), case #2 (top right), case #3 (bottom left), and case #4 (bottom right).

The differences in terms of anthropometric metrics between the different computational models are summarized in Tables 2-6. When the results obtained using the patient's regional model serve as reference, the magnitude of the relative difference is reported based on the subtraction of the target metric from the reference value. The mean Jaccard coefficient, which describes the similarity between models, for the best fitting models and the patient-specific models are 0.2 ± 0.17 and 0.19 ± 0.18, respectively (Table 2). The absolute difference of organ mass between the actual patient and the best-fitting model ranges from -241.4% to 83.6% while the difference between the patient and the patient-specific model ranges from -251.8% to 82.2%. The mean body perimeter in the range of each organ of the patients are in the range 81.6 ± 23 cm, while the mean absolute difference for the best-fitting model and patient-specific model are 10% and 9.8%, respectively. The organ-surface distances of the patient are in the range 13.5 ± 4.1 cm for different organs while the mean absolute difference between the patient and the best-fitting model and patient-specific model are 7.1% and 4.4%, respectively. The body effective diameters for each organ of the patient are in the range 24.6 ± 7.9 cm, while the mean absolute differences for the best-fitting model and patient-specific model are 6.6% and 4.5%, respectively.

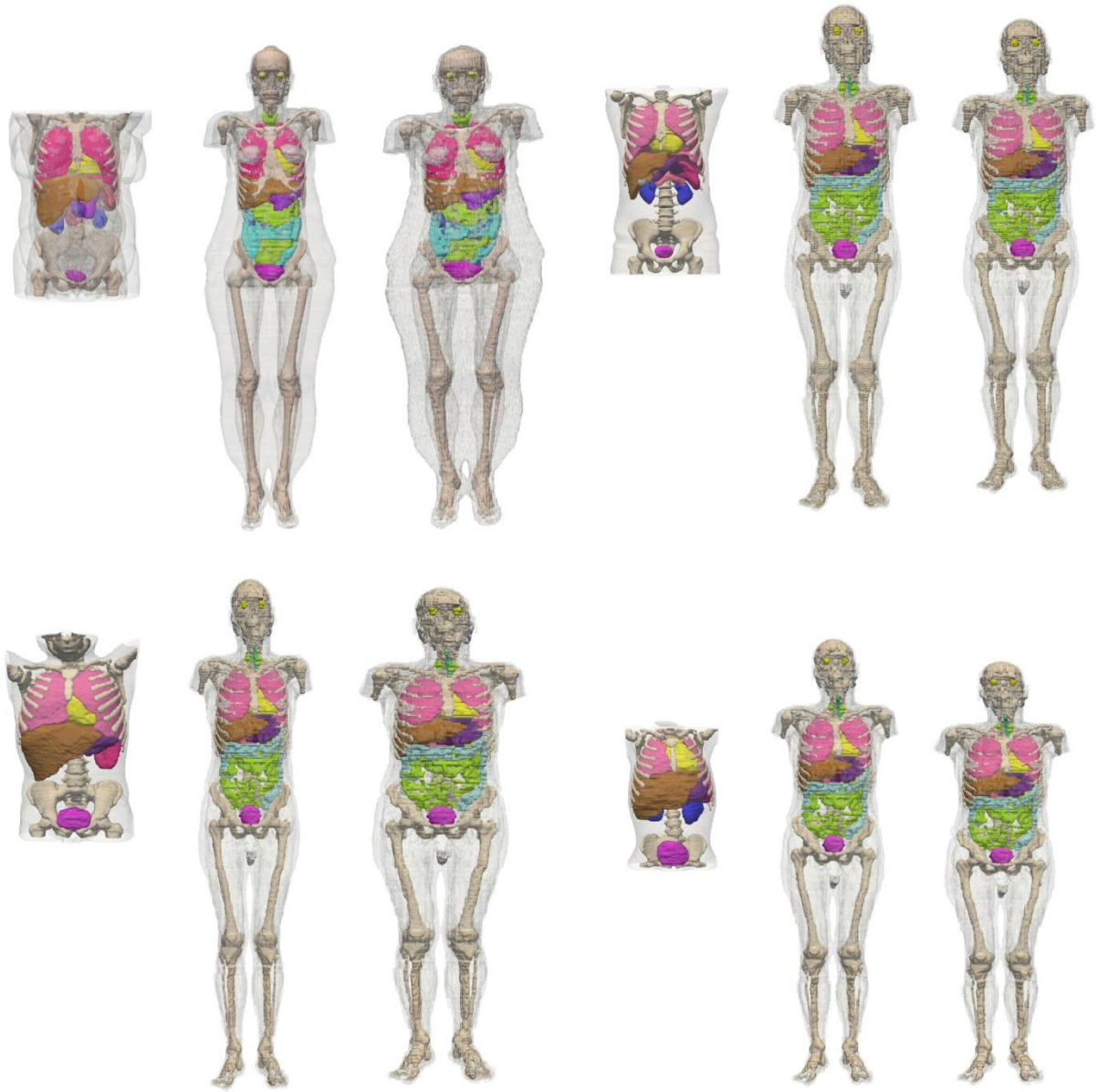


Figure 2. Three-dimensional views of, from left to right, patients' regional models, the corresponding anchor phantoms, and patient-specific phantoms. Patient case#1 (top left), case#2 (top right), case#3 (bottom left), case#4 (bottom right).

Table 2. Jaccard's coefficients for organs between the patient regional model and other computational models.

	JACCARD COEFFICIENTS							
	Best-fitting phantom				**Patient-specific phantom**			
Organs	#1	#2	#3	#4	#1	#2	#3	#4
Lung	0.345	0.48	0.44	0.54	0.56	0.48	0.47	0.50
Heart	0.054	0.19	0.13	0.17	0.12	0.19	0.16	0.17
Liver	0.346	0.37	0.49	0.52	0.52	0.38	0.35	0.53
Kidney	0.287	0.15	0.15	0.22	0.21	0.16	0.14	0.19
Stomach	0.069	0.04	0.07	0.09	0.05	0.03	0.07	0.07
Pancreas	0.122	0.08	0.07	0.04	0.16	0.07	0.16	0.02
Bladder	0.016	0.04	0.0	0.04	0.01	0.04	0.09	0.05
Spleen	0.306	0.01	0.24	0.16	0.08	0.02	0.02	0.18

Table 3. Comparison of organ masses between the regional patient models, best-fitting and registered models.

	ORGAN MASS DIFFERENCE (%)							
	Patient vs. best-fitting model				Patient vs. patient-specific model			
Organs	#1	#2	#3	#4	#1	#2	#3	#4
Lungs	29.9	39.7	46.5	-44.7	-54.0	37.9	46.6	-59.5
Heart	-24.5	-22.5	0.9	31.1	-1.9	-26.2	2.1	22.5
Liver	33.1	-51.7	17.3	46.8	0.0	-55.6	17.2	41.3
Kidneys	31	-10.8	-15.3	68.6	-1.0	-14.2	-13.7	65.5
Stomach	-20.2	-188.0	-121.8	39.7	0.0	-192.2	-122.0	36.3
Pancreas	15.7	-241.4	-17.5	-225.6	2.8	-251.8	-14.3	-241.1
Bladder	-30.8	-233.3	21.5	51.1	-68.9	-242.2	20.5	44.6
Spleen	45	27.6	63.3	83.6	1.9	26.4	63.3	82.2

Table 4. Comparison of mean body perimeter at organ longitudinal scan range between the regional patient model and best-fitting phantom and registered models.

Organs	**MEAN BODY PERIMETER DIFFERENCE** (%)							
	Patient vs. best-fitting model				Patient vs. patient-specific model			
	#1	#2	#3	#4	#1	#2	#3	#4
Lungs	14.4	-15.5	-12.2	-22.1	-28.8	-19.1	-26.1	-31.1
Heart	20.3	-17.3	-10.7	-20.6	-21.2	-21.2	-24.4	-29.5
Liver	22.1	-5.5	-5.4	-3.5	-17.1	-9.8	-18.7	-11.0
Kidneys	16.4	-0.9	-1.2	-1.9	-25.0	-5.2	-13.9	-9.3
Stomach	18.3	-5.5	-0.6	-1.8	-22.5	-9.9	-13.3	-9.1
Pancreas	16.3	-3.2	-0.1	2.2	-24.8	-7.8	-12.8	-5.1
Bladder	-4.3	-7.1	-9.3	-22.9	-52.7	-10.1	-23.0	-31.9
Spleen	21.4	-7.6	-5.7	-3.8	-18.0	-11.9	-19.0	-11.1

Table 5. Comparison of organ-surface distances between the regional patient model and best-fitting phantom and registered models.

	ORGAN-SURFACE DISTANCE DIFFERENCE (%)							
	Patient vs. best-fitting model				Patient vs. patient-specific model			
Organs	#1	#2	#3	#4	#1	#2	#3	#4
Lungs	19.2	-6.2	-0.5	-14.1	-5.2	-9.9	-12.8	-22.7
Heart	22.5	-7.6	-0.7	-11.8	-3.2	-11.8	-13.0	-20.2
Liver	19.9	-1.1	-1.0	0.7	-4.9	-5.0	-13.4	-6.9
Kidneys	17.4	-1.4	0.7	0.5	-8.7	-5.6	-11.3	-7.2
Stomach	16.1	-5.7	4.7	1.5	-9.9	-10.3	-7.0	-6.2
Pancreas	16.6	1.3	2.7	5.8	-9.8	-3.1	-9.2	-1.5
Bladder	-2.5	-2.4	-5.5	-13.6	-35.3	-6.6	-18.5	-22.4
Spleen	18	-3.9	-0.5	1.4	-6.3	-7.8	-12.8	-6.1

Table 6. Comparison of mean body effective diameters at organ longitudinal scan range between the different computational models.

	Mean Body Effective Diameter Difference (%)							
	Patient vs. best-fitting model				Patient vs. patient-specific model			
Organs	#1	#2	#3	#4	#1	#2	#3	#4
Lungs	20.2	-3.1	2.9	-9.7	-7.0	-7.9	-8.9	-18.1
Heart	24.7	-4.6	3.3	-10.4	-2.1	-9.4	-8.4	-18.7
Liver	18.4	1.2	0.6	0.0	-9.5	-3.4	-11.6	-7.7
Kidneys	13.7	4.5	3.0	0.7	-14.7	0.0	-8.6	-7.1
Stomach	16.5	0.8	4.2	1.7	-11.4	-3.8	-7.6	-5.8
Pancreas	12.3	2.9	3.5	3.8	-14.9	-1.6	-8.1	-264.4
Bladder	3.1	2.9	-4.6	-14.5	-29.2	-1.7	-17.3	-23.3
Spleen	19.3	0.1	0.6	0.4	-7.5	-4.5	-11.8	-7.1

Organ absorbed dose and effective dose

Organ radiation doses to the studied patients were calculated using the MCNPX Monte Carlo code using three computational models as input. The results were compared with the values estimated by Radimetrics™ dose monitoring software. Figure 3 shows the organ absorbed doses from CT examinations for regional computational models of patients included in this study. The mean absorbed doses to the patients are 9.2 mGy for the lung, 8.78 mGy for the heart, 10.14 mGy for the liver, 10.2 mGy for the kidney, 9.48 mGy for the stomach, 8.7 mGy for the pancreas, 9 mGy for the bladder and 10.63 mGy for the spleen.

As shown in Figure 4, when the results calculated using the patient's regional model serve as reference, the mean absolute discordance of organ doses of the best-fitting model, the patient-specific model and Radimetrics™ are 15.5%, 9.1% and 41.1%, respectively. The effective dose to the patient models are 11.74 mSv, 5.96 mSv, 10.52 mSv and 9.5 mSv for case#1, case#2, case#3 and case#4, respectively. Conversely, the values calculated using the best-fitting model, patient-specific model and Radimetrics™ are within the range of 9.5 ± 3.3 mSv, 9.2 ± 2.14 mSv and 10.5 ± 5.5 mSv, respectively. The absolute difference between the effective dose reported for the patient regional model and the best-fitting model, the patient-specific model, and Radimetrics™ are 15%, 5.7%, and 58.6%, respectively.

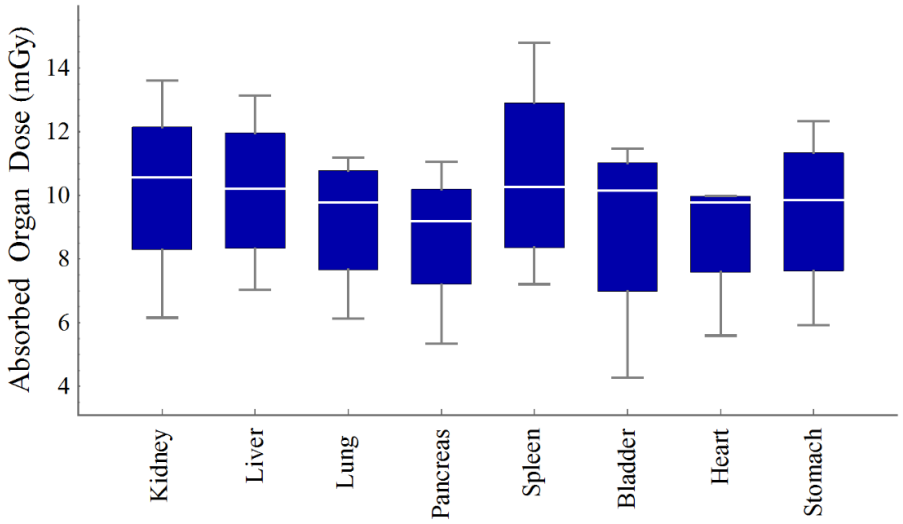


Figure 3. Absorbed doses for segmented target organs.

IV. Discussion

The patients' dose profiles from CT scans were calculated using different computational phantoms: the regional patient model, the best-fitting model selected using height-and-weight matching from previously developed phantom library and the registered patient-specific model. When the regional patient model is used as reference, the registered patient-specific model is superior to the selected best-fitting model in terms of error in estimated organ dose. The results extracted from Radimetrics™ dose tracking software showed a considerable deviation from the reference even though the morphometric characterisitics of the patient have been taken into account. Through image registration, the proposed approach allows matching the patient's data to a whole-body phantom presenting with similar location and anatomical morphometry. The evaluated metrics, including organ mass, body perimeter, organ-surface distance and effective diameter increased in patient-specific models owing to the nonlinear deformation of the best-fitting phantom during the registration process to match the obtained anatomical masks (skeleton, lung and body contour) of patients to the corresponding anchor phantom. The Jaccard coefficients do not show a good similarity in some cases owing to the dependency of this index on organs' volume and the spatial location of organs for the different computational models (Table 2).

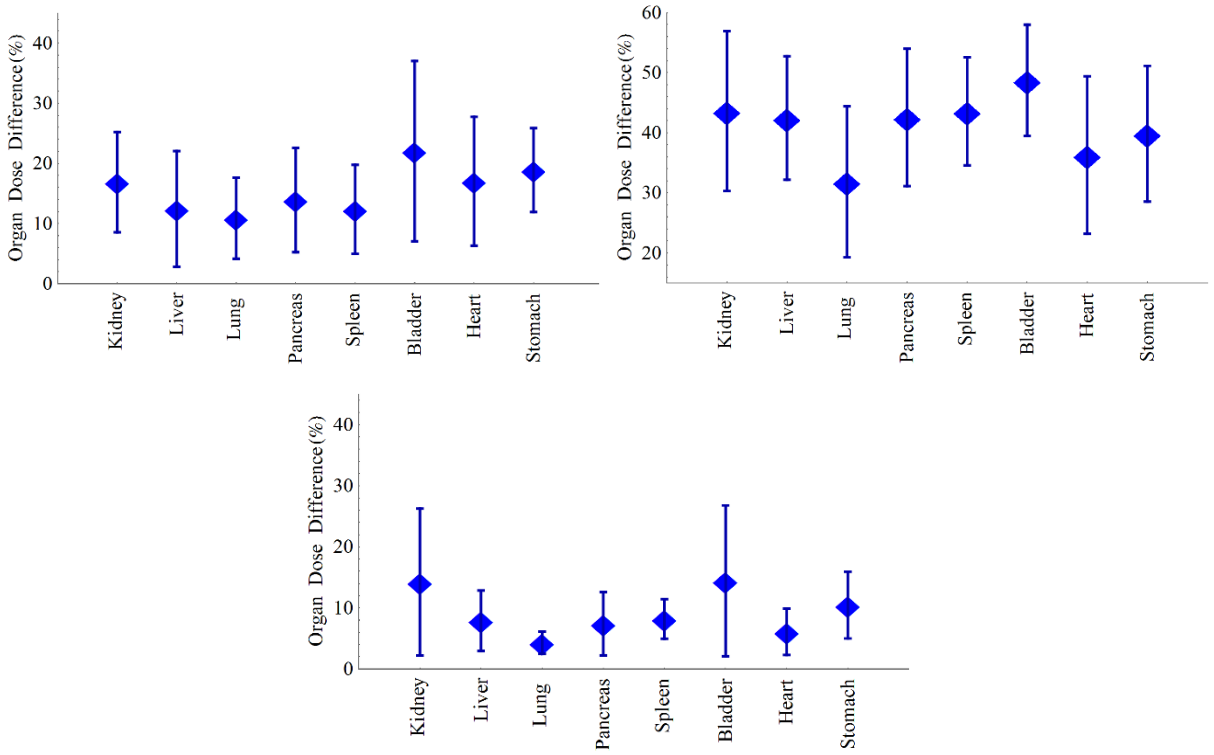


Figure 4. Relative differences of absorbed doses in segmented target organs for (top-left) best-fitting models, (top-right) Radimetrics software and (bottom) patient-specifc model with respect to patient regional models.

In this work, we selected a group of patients presenting diverse anatomical characteristics to evaluate our methodology on a representative patient population: Case #1, a patient with high BMI and large skeleton size, selected to evaluate our methodology on extreme cases and case #3, a pediatric patient presenting with a large skeleton structure. In these two cases, there is a noticeable difference in dose estimates between the reference values and the best fitting model selected from the phantom library, even though matching patient age, height and weight were carried out. Patient-specific models for these cases improve the dosimetric results by about 15%. Case #2, is an athletic male with small-sized internal organs whereas case #4 is a morphometrically normal pediatric patient where the best-fitting model provides organ dose estimates deviating by ~10% from the reference values without any remarkable improvement in organ dose calculation using patient-specific models.

The habitus-dependent phantom library approach suffers from the limited number of anatomies and morphometries that cause an uncertainity in dose estimation [30]. This approach provides acceptable organ doses for anthromorphically and anatomically normal patients. The current methodology building patient-specific computational models from patients' CT images demonstrated noticeable improvement in the accuracy of organ dose calculation for extreme cases. The calculated absorbed dose in the lungs is significantly improved owing to the excellent matching between the lungs mask obtained from patient CT images and the best-fitting phantom. Although Radimetrics™ dose tracking software estimates patient size from the scout scan and the examined anatomical region is determined using image registration methods, the deviation from the reference doses can be caused by the oversimplified anatomy and body contour of stylized models and the limited number of phantoms available that do not cover the diversity of different anatomies and morphometries. In addition, implementing the scan range into the simulation based on predefined CT acquisition protocols (anatomical landmarks) is another source of errors in Radimetics™.

This study bears a number of limitations, including the construction of the regional patient models for evaluation of the domestric results and the development of the phantom library. Regarding the regional patient model, the segmentation of internal organs was performed manually, where prior anatomical knowledge guides identification of organs and delimitation of their boundaries. This approach is not feasible for routine clinical application. Fortunately, the body countour, skeleton and lungs can be automatically segmented from CT images

and used during model registration to produce a patient-specific model. Likewise, assumptions regarding chemical compositions and density for organs may also introduce errors in dose estimations. Another limitation related to the selection of the best-fitting model is that the developed library contains a limited number of phantoms and cannot cover all patient morphometries. This limitation may potentially be addressed by expanding the library in future work. The limited number of patients included in this study is among the limitations of this work. In addition, the posture differences of the anchor phantom and patient studies (arms up vs. arms down) may introduce a displacement of internal organs. The registration of the anchor phantom to regional patient images can be further optimized to improve the matching of the patient-specific model and patient regional model. The TCM model adopted in the simulation process didn't consider the angular modulation owing to the lack of sufficient information (proprietary raw data format undisclosed by the manufacturer). Other limitations of this work include the few organs considered and the use of only one CT scanner model. The construction of patient-specific models for accurate dosimetry calculations remains a challenging issue requiring further research and development efforts [31]. Deep learning approaches have brought revolutionary advances in the field of medical image analysis that could be useful for constructing patient-specific models through automatic segmentation of medical images (body contours and internal organs).

V. Conclusion

The aim of this study is to quantify the dosimetric characteristics of patient-specific computational models in CT dose estimation. Although using height-weight matching to select the best-fitting model from a comprehensive phantom library is feasible in clinical setting, the estimated organ dose may differ from the reference by up to 36% as demonstrated in this study. If, however, patient CT images are available, a reference computational model can be matched to the patient data to produce a patient-specific computational model for radiation dosimetry calculations, thus improving the accuracy of organ dose estimation.

Acknowledgements

This work was supported by the Swiss National Science Foundation under grant SNSF 320030_176052 and Iran's Ministry of Science and Technology.

References

[1] W. A. Kalender, "Dose in x-ray computed tomography,"*Phys Med Biol,* vol. 59, no. 3, pp. R129-50, 2014.

[2] Y. Ohno, H. Koyama, S. Seki, Y. Kishida, and T. Yoshikawa, "Radiation dose reduction techniques for chest CT: Principles and clinical results,"*European journal of radiology*, 2018.

[3] C. H. McCollough, A. N. Primak, N. Braun, J. Kofler, L. Yu, and J. Christner, "Strategies for reducing radiation dose in CT.,"*Radiologic clinics of North America,* vol. 47, no. 1, pp. 27-40, 2009.

[4] R. Smith-Bindman *et al.*, "Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer.,"*Arch Intern Med,* vol. 169, no. 22, pp. 2078-2086, 2009.

[5] World Health Organization, *Communicating radiation risks in paediatric imaging: Information to support healthcare discussions about benefit and risk.* Geneva, Switzerland: World Health Organization, 2016.

[6] G. Wang, C. R. Crawford, and W. A. Kalender, "Multirow detector and cone-beam spiral/helical CT.,"*IEEE Trans Med Imaging,* vol. 19, no. 9, pp. 817-821, 2000.

[7] M. Cros, R. M. S. Joemai, J. Geleijns, D. Molina, and M. Salvado, "SimDoseCT: dose reporting software based on Monte Carlo simulation for a 320 detector-row cone-beam CT scanner and ICRP computational adult phantoms,"*Phys Med Biol,* vol. 62, no. 15, pp. 6304-6321, 2017.

[8] B. M. Moore, S. L. Brady, A. E. Mirro, and R. A. Kaufman, "Size-specific dose estimate (SSDE) provides a simple method to calculate organ dose for pediatric CT examinations,"*Med Phys,* vol. 41, no. 7, p. 071917, 2014.

[9] M. Bostani *et al.*, "Erratum: "Estimating organ doses from tube current modulated CT examinations using a generalized linear model" [Med. Phys. Vol 44 (4), 1500-1513 (2017)].",*Med Phys,* vol. 44, no. 7, p. 3883, 2017.

[10] P. B. Johnson, A. Geyer, D. Borrego, K. Ficarrotta, K. Johnson, and W. E. Bolch, "The impact of anthropometric patient-phantom matching on organ dose: a hybrid phantom study for fluoroscopy guided interventions.,"*Med Phys,* vol. 38, no. 2, pp. 1008-17, 2011.

[11] Y. H. Na, B. Zhang, J. Zhang, P. F. Caracappa, and X. G. Xu, "Deformable adult human phantoms for radiation protection dosimetry: anthropometric data representing size distributions of adult worker populations and software algorithms.,"*Phys Med Biol,* vol. 55, no. 13, pp. 3789–3811, 2010.

[12] K. McMillan *et al.*, "Estimating patient dose from CT exams that use automatic exposure control: Development and validation of methods to accurately estimate tube current values.,"*Med Phys,* vol. 44, no. 8, pp. 4262-4275, 2017.

[13] W. P. Segars *et al.*, "Patient specific computerized phantoms to estimate dose in pediatric CT,"*SPIE Medical Imaging,* vol. 7258, p. 9, 2009.

[14] P. Sahbaee, W. P. Segars, and E. Samei, "Patient-based estimation of organ dose for a population of 58 adult patients across 13 protocol categories,"*Med Phys,* vol. 41, no. 7, p. 072104, 2014.

[15] W. A. Kalender, N. Saltybaeva, D. Kolditz, M. Hupfer, M. Beister, and B. Schmidt, "Generating and using patient-specific whole-body models for organ dose estimates in CT with increased accuracy: Feasibility and validation,"*Physica Medica: European Journal of Medical Physics,* vol. 30, no. 8, pp. 925-933, 2014.

[16] Y. Gao *et al.*, "Patient-specific organ and effective dose estimates in pediatric oncology computed tomography,"*Physica medica : PM : an international journal devoted to the applications of physics to medicine and biology : official journal of the Italian Association of Biomedical Physics,* vol. 45, pp. 146-155, 2018.

[17] "Radimetrics Enterprise Platform: Dose Management Solution. Bayer HealthCare. http://www.radiologysolutions.bayer.com/products/ct/dosemanagement/rep/." (accessed 2019).

[18] M. Bostani *et al.*, "Estimating organ doses from tube current modulated CT examinations using a generalized linear model.,"*Med Phys,* vol. 44, no. 4, pp. 1500-1513, 2017.

[19] G. Wang and M. W. Vannier, "Helical CT image noise-analytical results.,"*Med Phys,* vol. 20, no. 6, pp. 1635-1640, 1993.

[20] B. Quinn, Z. Dauer, N. Pandit-Taskar, H. Schoder, and L. T. Dauer, "Radiation dosimetry of 18F-FDG PET/CT: incorporating exam-specific parameters in dose .estimates,"*BMC Med Imaging,* vol. 16, no. 1, p. 41, 2016.

[21] H. Zaidi and M. Ay, "Current status and new horizons in Monte Carlo simulation of X-ray CT scanners,"*Med Biol Eng Comput,* 10.1007/s11517-007-0207-9 vol. 45, no. 9, pp. 809-817, 2007.

[22] R. Kikinis and S. Pieper, "3D Slicer as a tool for interactive brain tumor segmentation,"*Conf Proc IEEE Eng Med Biol Soc,* vol. 2011, pp. 6982-6984, 2011.

[23] I. J. Feain, L. Court, J. R. Palta, S. Beddar, and P. Keall, "Innovations in Radiotherapy Technology,"*Clin Oncol (R Coll Radiol),* vol. 29, no. 2, pp. 120-128, 2017.

[24] H. J. Aerts *et al.*, "Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach.,"*Nature communications*, vol. 5, p. 4006, 2014.
[25] Vital and Health Statistics, "Anthropometric reference data for children and adults: United States, 2011–2014.," in "National Health Statistics Reports," Centers for Disease Control, Hyattsville, Maryland, DHHS Publication No. 2016–1604, 2016. [Online]. Available: http://www.cdc.gov/nchs/data/nhsr
[26] L. Ibanez, W. Schroeder, L. Ng, and J. Cates, "The ITK software guide," 2005.
[27] T. Xie, P. A. Poletti, A. Platon, C. D. Becker, and H. Zaidi, "Assessment of CT dose to the fetus and pregnant female patient using patient-specific computational models.,"*Eur Radiol*, vol. 28, no. 3, pp. 1054–1065, 2018.
[28] H. J. Johnson, M. M. McCormick, L. Ibanez, and and the Insight Software Consortium, "The ITK software guide.,"
[29] X. Zhang *et al.*, "Feasibility study of micro-dose total-body dynamic PET imaging using the EXPLORER scanner,"*Journal of Nuclear Medicine*, vol. 55, no. supplement 1, p. 269, 2014.
[30] H. Zaidi and X. G. Xu, "Computational anthropomorphic models of the human anatomy: The path to realistic Monte Carlo modeling in medical imaging.,"*Annu Rev Biomed Eng*, vol. 9, no. 1, pp. 471-500, 2007.
[31] T. Xie, P. Zanotti-Fregonara, A. Edet-Sanson, and H. Zaidi, "Patient-specific computational model and dosimetry calculations for a patient pregnant with twins undergoing a PET/CT examination.,"*J Nucl Med*, vol. 59, no. 9, pp. 1451-1458, 2018.

Chapter 5

Assessment of uncertainties associated with Monte Carlo-based personalized dosimetry in clinical CT examinations

Azadeh Akhavanallaf, Tianwu Xie, and Habib Zaidi

Contribution: Study conception and design, methodological development, computer programming, analysis and interpretation of results, manuscript preparation.

Phys Med Biol, Vol. 65, No. 4, pp 045008 (2020)

Abstract

The clinical value of x-ray computed tomography (CT) has skyrocketed in the last decade while at the same time being the main source of medical exposure to the population. Concerns regarding the potential health hazards associated with the use of ionizing radiation were raised and an appropriate estimation of absorbed dose to patients is highly desired. In this work, we aim to validate our developed Monte Carlo CT simulator using in-phantom dose measurements and further assess the impact of personalized scan-related parameters on dosimetric calculations. We developed a Monte Carlo-based CT simulator for personalized organ level dose calculations, in which the CT source model, patient-specific computational model and personalized scanning protocol were integrated. The CT simulator was benchmarked using an ionization chamber and standard CT Dose Index phantom while the dosimetry methodology was validated through experimental measurements using thermoluminescent dosimeters (TLDs) embedded within an anthropomorphic phantom. Patient-specific scan protocols extracted from CT raw data and DICOM image metadata, respectively, were fed as input into the CT simulator to calculate individualized dose profiles. Thereby, the dosimetric uncertainties associated with using different protocol-related parameters were investigated. The absolute absorbed dose difference between measurements and simulations using the ionization chamber was less than 3%. In the case of the anthropomorphic phantom, the absolute absorbed dose difference between simulations and TLD measurements ranged from −8.3% to 22%, with a mean absolute difference of 14% while the uncertainties of protocol-related input parameters introduced an extra absolute error of 15% to the simulated results compared with TLD measurements. The developed methodology can be employed for accurate estimation of organ level dose from clinical CT examinations. The validated methodology can be further developed to produce an accurate MC simulation model with a reduced computational burden.

I. Introduction

Computed Tomography (CT) has become a key diagnostic imaging modality in clinical diagnosis of a wide range of diseases. The number of CT scans in the US had an average annual increase of 10% between 1995 and 2015 [1, 2]. The sharp utilization trend of CT imaging in clinical setting has raised health concerns regarding potential risks of ionizing radiation on patients undergoing CT examinations. Although medical exposure brings individual benefits to the patient, it should still follow the principles of justification and optimization [3, 4]. In this context, each medical radiological procedure has to be optimized for the specific task and individual patient. Efforts devoted to designing patient-specific CT scanning protocols depending on the target task, scanner model and patient anatomy, can greatly benefit from a unified methodology for radiation dose estimation [5, 6].

Direct measurement of energy deposition in the different tissues/organs within the patient's body is not conceivable in clinical setting. Therefore, experimental measurements using dosimeters embedded within physical phantoms and Monte Carlo simulations using realistic anthropomorphic computational phantoms served as substitutes. Monte Carlo simulation, deemed to be the gold standard technique for dosimetry calculations, should be carefully validated against experimental measurements because the input parameters related to the imaging system, patient's anatomical model and the scanning protocol dictate the accuracy of the obtained results. A number of studies reported on the use of Monte Carlo programs benchmarked using standard CT Dose Index (CTDI) and other anthropomorphic physical phantoms [7-10]. The paradigm shift introduced by advances in personalized medicine and precision medicine stimulated the development of strategies for patient-specific dosimetry and protocol optimization during the last few years. In this context, Segars *et al.* constructed personalized computational models by mapping the segmented model of patient CT images to a template anatomical model using a deformable registration algorithm [11]. Li *et al.* developed a Monte Carlo code for patient-specific dose calculation [12] by constructing patient-specific computational models through manual segmentation that are fed as input to the Monte Carlo program. Kalender *et al.* constructed personalized whole-body phantoms from regional CT images through appending the scan range to an anchor phantom to assess the effect of scatter and overscanning on organ doses [13]. Xie *et al.* proposed a methodology for constructing patient-specific computational models based on deformable registration of patient CT regional images on a habitus-dependent anchor phantom [14, 15]. They employed the structural deformation of the best-fitting phantom from a previously developed library of computational models through automated non-rigid registration to estimate the patient-specific model [16]. They calculated patient's organ-level radiation dose using the obtained personalized computational phantom and imaging protocol implemented into the MC simulation. The Radimetrics™ commercial dose tracking software (Bayer HealthCare, Berlin, Germany) provides Monte Carlo-based organ-level dose profile from patient CT images where the scan parameters are extracted from CT image DICOM header information [17]. Radimetrics™ simply maps the patient's regional CT images on Cristy & Eckerman mathematical computational phantoms categorized by age and gender [18]. It further adjusts the phantom's diameter according to the effective diameter of the patient obtained from CT images.

The dosimetric impact of scan parameters (e.g. x-ray energy spectrum, beam filtration, tube current modulation, tube start angle, over-ranging, etc) fed into Monte Carlo simulations has been investigated in previous studies [19-22] and more comprehensively in the recent AAPM report No. 246 [23]. The input data are commonly provided by CT scanner manufacturers, experimental measurements, or extracted directly from scanner control console or the generated radiation dose structured report after the examination. However, there are uncertainties associated with all sources of input parameters. Muryn *et al.* examined the impact of deviations related to the input parameters on the simulated dose profiles [24]. They studied the parameters linked to the CT scanner (e.g. x-ray spectrum, beam filtration, beam width) and the scan setup (e.g. tube start angle, scan length, isocenter position) to address the uncertainties introduced on the simulated dose when the input parameters deviate from the actual values. However, in this work, the tube current modulation and the impact of patient's anatomy were not taken into account. Lee *et al.* performed organ-level dose estimation for a large cohort of CT examinations to investigate the dosimetric impact of uncertainties on patient-related and empirical scan-related parameters [25].

In this work, we aimed to provide a computationally-efficient framework for accurate patient-specific dose estimation. To this end, we developed a unified methodology for patient-specific dosimetry from CT examinations. Unlike previous works requiring manual segmentation of CT images to construct patient-specific computational model, we adopted methodology that automatically builds patient-specific computational models from CT images.

Subsequently, the validated CT source model, patient-specific computational phantom and scan parameters were integrated in the Monte Carlo code to calculate organ-level absorbed dose. Therefore, the dosimetry results were benchmarked against experimental measurements. To further assess the impact of uncertainties associated with simulation input parameters on the organ level personalized dosimetry, we compared different scenarios where the patient-specific scan-related parameters were extracted from the CT image DICOM header file and more detailed CT raw data, respectively.

II. Materials and methods

Monte Carlo simulations

In Monte Carlo-based CT dosimetry studies, three essential components are incorporated into the simulations. This includes the CT source model, computational phantom and protocol-related parameters. CT data were acquired on the Somatom Definition Edge scanner (Siemens Healthcare, Erlangen, Germany). The geometry of the CT system was realistically modelled by using an x-ray energy spectrum generated using SpekCalc [26] and tuning the inherent filtration of the scanner's x-ray tube and half value layer (HVL) associated with the x-ray energy. The applied inherent filtration and HVL were extracted from system owner manual [27] and were matched with the results reported by Yang *et al.* [28]. The Bowtie or shaped filter and beam collimators, including dynamic collimators as well as fixed collimation components were elaborately modelled based on the information provided by the manufacturer. The geometry of the gantry (e.g. focal spot size, the distance between focal spot and isocenter), fan angle, etc. were taken into account in the simulation as shown in Figure 1. This CT source model was created within the MCNPX general purpose Monte Carlo radiation transport code (version 2.6) [29]. The computational phantom with its detailed anatomy was constructed using a previously developed methology based on automated registration of the patient's CT images to an anchor phantom [5]. The protocol-related parameters, including scan mode, tube potential, total beam collimation, revolution time, exposure time, table speed, pitch factor and tube current modulation were extracted from DICOM header information using a MATLAB (The MathWorks Inc., Natick, MA, USA) function and implemented in the simulation code. The tube current modulation (longitudinal and angular), tube start angle and over-ranging information were extracted from the DICOM header information of both CT images and CT raw data and used in the simulation setup.

Validation of the CT scanner model

The conversion of the relative MCNPX dose tallies to absolute dose value was performed by conducting free in air measurement using a 10 cm RaySafe™ Solo pencil ionization chamber (Unfors RaySafe GmbH, Germany). The ionization chamber was placed at the isocenter of the CT scanner with its active volume aligned with the axis of gantry rotation. Free in air measurements were performed in single axial scans where the absorbed dose derived from simulations was calculated as:

$$D_{estimated} = D_{simulated} \times N \times \Omega \times mAs \times CF \qquad (1)$$

where $D_{simulated}$ is the simulated absorbed dose per photon emitted from the source (F6 tally in unit of MeV/g); N is the number of photons emitted from the source per solid angle per mAs; Ω is the solid angle of the fan-beam; mAs is the effective tube current-time product value; and CF is a calibration factor to correct for uncertainties introduced in Monte Carlo simulations to calculate the absorbed dose values from simulations in absolute units of mGy. To validate the developed CT source model, we benchmarked our simulation results against standard CTDI phantom measurement for both head (16 cm diameter) and body (32 cm diameter) cylindrical phantoms. To this end, firstly, we performed free in air measurements to estimate the CF as the ratio of the absolute measured dose to the simulated dose. In Eq. (1), there is a linear relationship between the measured dose and mAs, which is regularly checked during routine CT scanner quality control procedures. While the factors N and Ω depend on tube potential and total collimation, respectively, the relationship between these factors and the absorbed dose is not ideally linear [27]. For this reason, we reduced Eq. (1) to $D_{estimated} = D_{simulated} \times mAs \times CF$ by providing a unique calibration factor associated with the specific acquisition parameters (e.g. tube potential and total beam collimation) during CT examinations. Consequently, we measured the absolute dose in the CTDI phantom in helical mode with the same acquisition parameters used for free in air experiment to compare the simulation results

against the measurements using the estimated calibration factor. Three million photons were used in this simulation to achieve a statistical error less than 2%.

Anthropomorphic phantom dose measurements

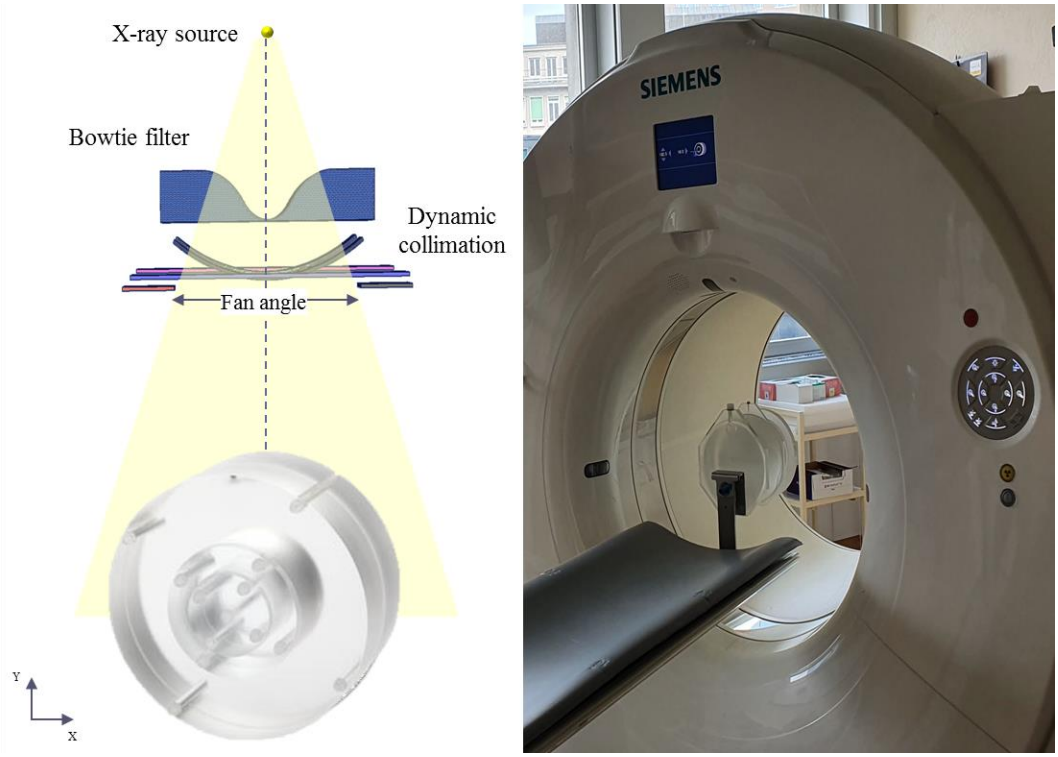


Figure 1. CT source model simulated in the MCNPX Monte Carlo code (left) and experimental set-up used to measure dosimetric metrics in the CTDI phantom (right).

To benchmark the accuracy of the developed Monte Carlo code for patient-specific dosimetry, absorbed doses were measured in the CIRS ATOM® dosimetry verification phantom (CIRS, Inc., Norfolk, VA, USA). This anthropomorphic physical phantom, axially sliced in 25 mm thick, consists of three types of materials, including skeleton, lung and soft tissue. The physical phantom was matched to its corresponding computational twin where the boundaries of all internal organs were mapped on the physical slice consistent with their delineation in the corresponding computational model (Figure 2). We used the previously developed program to construct the patient-specific phantom from CT images through automated deformable registration [5]. CT images of the ATOM phantom were employed to provide the anatomical masks of the skeleton, the lungs and body contour which constitute the basis of information used by the registration algorithm. Hence, the computational model of the phantom with detailed anatomy was constructed using the automated deformable registration algorithm.

TLDs (LiF, Harshaw TLD-100) in the form of 3.2×3.2×0.9 mm^3 chips were inserted within the tissue equivalent dosimeter holders embedded within the phantom. TLDs were individually calibrated in terms of absorbed dose in water for Co-60 radiation. A correction factor was multiplied by the TLD results to account for the TLDs response for the specific beam quality, i.e. x-ray energy and half value layer of the beam. A total number of 103 TLDs were distributed in the phantom. Depending on the size of organs, from two to several measurement points were used enabling accurate volume-averaged organ dose calculation. The background noise was determined using separate dosimeters that were not exposed. The measured quantity from reading the TLDs is the absorbed dose in water. Calculation of the absorbed dose in other tissues involved application of a correction factor, calculated as the ratio of the average mass energy absorption coefficients in the tissue in question per mass energy absorption in water. The average of the TLD readings for each organ was used as the measured organ absorbed dose.

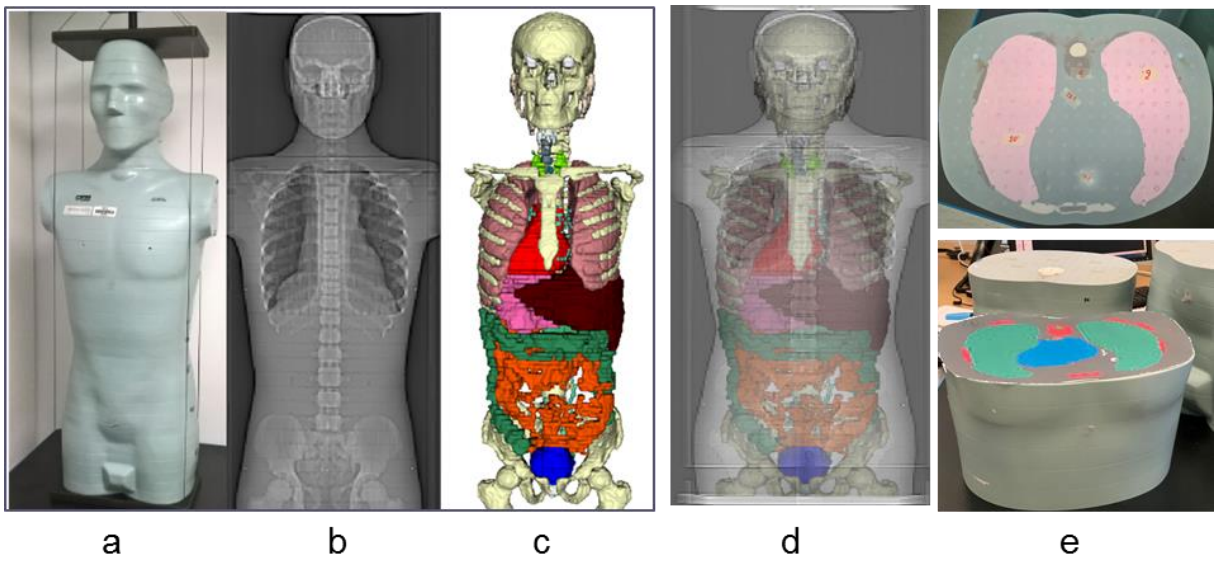


Figure 2. (a) Frontal view of the ATOM physical phantom, (b) phantom CT images, (c) computational model, (d) registered anatomy on phantom CT images; (e) axial view of the physical phantom (top) and the printed computational model (bottom).

Patient-specific organ-level dose simulation

Validation of the developed methodology in estimating patient-specific organ-level doses from CT examinations involved the comparison of simulation results with experimental measurements using the anthropomorphic physical phantom. The CT source model, computational model and scan parameters extracted from DICOM headers of both CT images and CT raw data were integrated in the MCNPX code. The acquisition parameters, including tube voltage, total collimation width, table speed, revolution time, pitch factor, and modulated tube current, start angle and over-ranging length were modelled in this simulation setup. To simulate helical whole body scanning (90 cm length), Monte Carlo simulations were run for 16'854 discrete source positions (576 source positions per rotation) taking into account the complete tube current modulation (longitudinal and angular modulation). Considering the linearity of the radiation output or measured dose with the mAs, the obtained energy deposition tally (Gy/particle) was multiplied by the mAs in each simulation point. To calculate organ doses in absolute units of mGy, the unique calibration factor that depends on the beam energy spectrum, filtration, and beam collimation was used. To calculate the specific calibration factor associated with the conducted CT examination, the value representing scan-specific radiation output ($CTDI_{vol}$) was simulated according to the 21 CFR 1020.33 guidelines [30]. Thereby, the scan-specific calibration factor was defined as the ratio of the simulated $CTDI_{vol}$ to the $CTDI_{vol}$ appearing in the dose report of anthropomorphic phantom CT examination. To evaluate the effect of dynamic collimation on dosimetric results, a correction factor defined as the time-weighted average of collimator during the scan, divided by the nominal collimator width for the scan was applied to the simulation results. The absorbed radiation dose associated with the topogram scan was also added to the simulation results.

To evaluate the accuracy of the developed simulation framework, organ-level dose profiles obtained from experimental TLD measurements were compared with simulation results as well as the doses reported by Radimetrics^{TM} commercial dose tracking software.

Uncertainties associated with simulation input parameters

The accuracy of results obtained from Monte Carlo radiation transport simulations depends directly on the input parameters fed into the simulator. In this section, we estimated organ-level doses using the information extracted from DICOM header information of CT images, referred to as image-based simulation, which basically contains longitudinal tube current modulation and lacks information about tube start angle and over-ranging length. In the exact simulation where the detailed input parameters were obtained from CT raw data, the angular component has been also taken into account in addition to the longitudinal tube current modulation. The complete tube current modulation was obtained from CT raw data based on CAREDose4D module in the Siemens CT scanner, which

employs patient size information from the CT localizer scout scan to predict the longitudinal and angular modulation functions. Overall, the tube current value reported in the DICOM header at each table position is calculated from the moving average of complete tube current over one rotation while the high frequency components (i.e. angular modulation) are smoothed (Figure 3). To simulate a helical scan in the image-based simulation, the source positions were modelled based on fixed intervals in the Z direction, while the angular position was determined according to the gantry revolution time and table speed obtained from the DICOM header information. However, it has been reported that over-ranging length is around half of the collimation width. This parameter is proprietary information for the different manufacturers [31]. Therefore, we extracted over-ranging length from CT raw data to accurately simulate this feature. In the image-based simulation, the random start angle was modelled and the over-ranging length was ignored. Here, the dosimetric impact of the tube start angle and number of simulation points across the entire scan was investigated and the results compared with those obtained from exact simulations.

Statistical analysis

The comparison between the results obtained from exact simulations and experimental measurements using TLDs, serving as reference, underwent statistical analysis. Furthermore, the intraclass correlation coefficient (ICC), as a measure of the reliability of organ dose calculation methods, was considered.

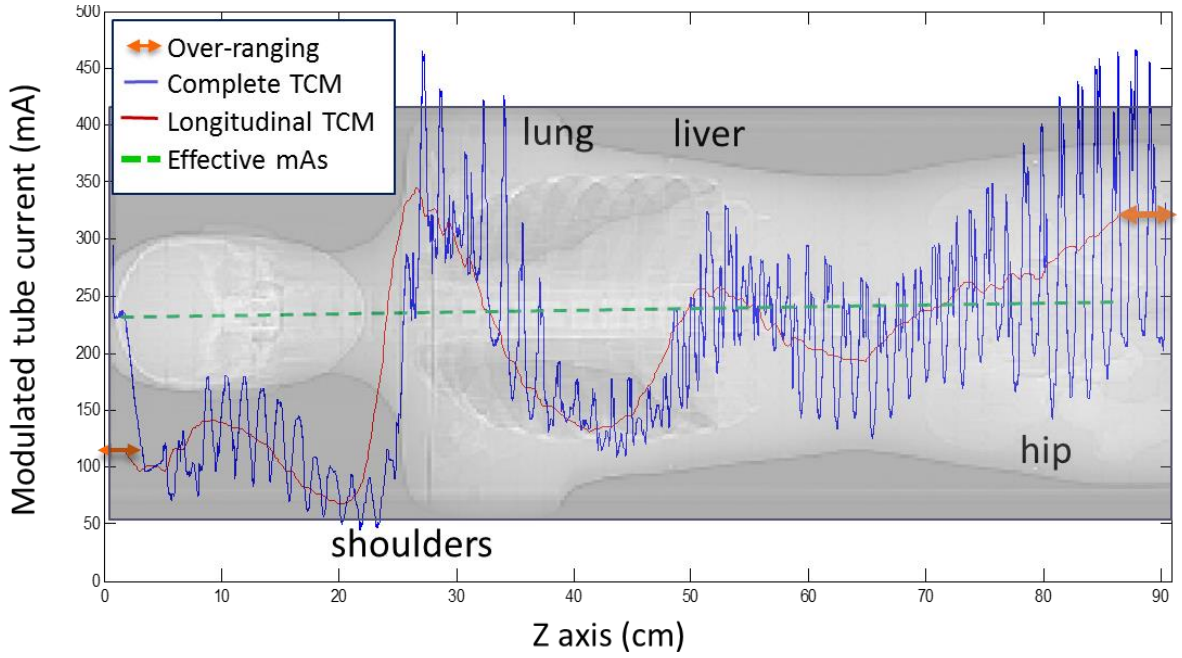


Figure 3. Extracted tube current modulation profile overlaid on phantom topogram. The solid blue line represents the complete modulated tube current schemes extracted from the raw projection data. The solid red line represents the longitudinal modulated tube current (moving average over one rotation). The dashed green line represents the effective time-product tube current (mAs) across the entire scan (mean mAs/pitch). The orange arrow represents pre- and post-spiral over-ranging.

III. Results

Validation of Monte Carlo simulation model

The CT source model was defined in the MCNPX code based on information provided by manufacturer by combining simple geometries. The model was validated through comparison with experimental measurements using the standard CTDI head and body phantoms. The acquisition parameters associated with the examination and the corresponding calibration factors obtained from free in air measurements are illustrated in Table 1. Table 2 summarizes the central and peripheral $CTDI_{100}$ and CTDIvol for the protocols specified in Table 1 demonstrating absolute mean differences between simulations and measurements around 6.4%. The computational model of the ATOM physical phantom derived from CT images is shown in Figure 2. To validate the developed methodology

for patient-specific organ-level dosimetry, the experimental CT acquisition illustrated in Table 1 using the physical phantom was performed (CTDIvol = 4.97 mGy). The results from the in-phantom measurements were considered as reference and compared with simulation results as well as organ dose profiles reported by Radimetrics™ software. The mean absolute difference between TLD measurements and MC simulations is about 14% (range [-8.3 - 22%]), while it exceeds 33% for Radimetrics™ (Figure 4). The effective dose obtained from TLD measurements was about 11.44 mSv while the effective dose calculated from simulation and Radimetrics™ was about 12.44 mSv and 7.35 mSv, respectively.

Table 1. CT acquisition parameters for experimental measurements performed in the CTDI phantoms and the anthropomorphic physical phantom.

Acquisition parameters	Head standard (perfusion)	Head standard	Body standard	Physical phantom (whole-body)
Tube voltage (KeV)	80	120	120	100 KeV
Total collimation (mm)	64*0.6	12*1.2	64*0.6	64*0.6
Tube current (mA)	200	350	252 (ref)/210 (eff)	268 (ref)/126 (eff)
Rotation time (sec)	1	1	0.5	0.5 sec
Pitch factor	-	-	-	0.8
Mode	axial	axial	axial	helical
Calibration factor	25.3	67.7	48.7	77.6

Table 2. Comparison between measured and Monte Carlo-based calculations of the $CTDI_{100}^{C}$ (central),$CTDI_{100}^{P}$ (peripheral) and $CTDI_{vol}$ in body and head CT dosimetry phantoms.

Acquisition type	Voltage (KeV)	Measurements			Simulations			$CTDI_{vol}$ Difference (%)
		$CTDI_{100}^{C}$ (mGy)	$CTDI_{100}^{P}$ (mGy)	CTDIvol (mGy)	$CTDI_{100}^{C}$ (mGy)	$CTDI_{100}^{P}$ (mGy)	$CTDI_{vol}$ (mGy)	
Head standard (perfusion)	80	8.39	8.94	8.76	7.68	8.47	8.21	6.3
	100	16.95	17.61	17.39	14.95	17.57	16.70	4.0
Body standard	120	7.88	14.71	12.43	6.91	17.53	13.99	12.5
Head standard	120	7.14	7.34	7.28	7.25	7.62	7.5	3.0

Uncertainties associated with input parameters

To assess the impact of input parameters on the simulation results, we obtained the input data from two sources: CT raw data and DICOM header. The complete tube current modulation obtained from CT raw data was compared with longitudinal tube current modulation obtained from DICOM header (Figure 3). The organ dose resulted from image-based simulation was compared with reference values and results from exact simulation. The mean absolute difference between the absorbed dose from the image-based simulation compared with reference organ dose values is about 29%. The effective dose from the image-based simulation was calculated about 11.68 mSv.

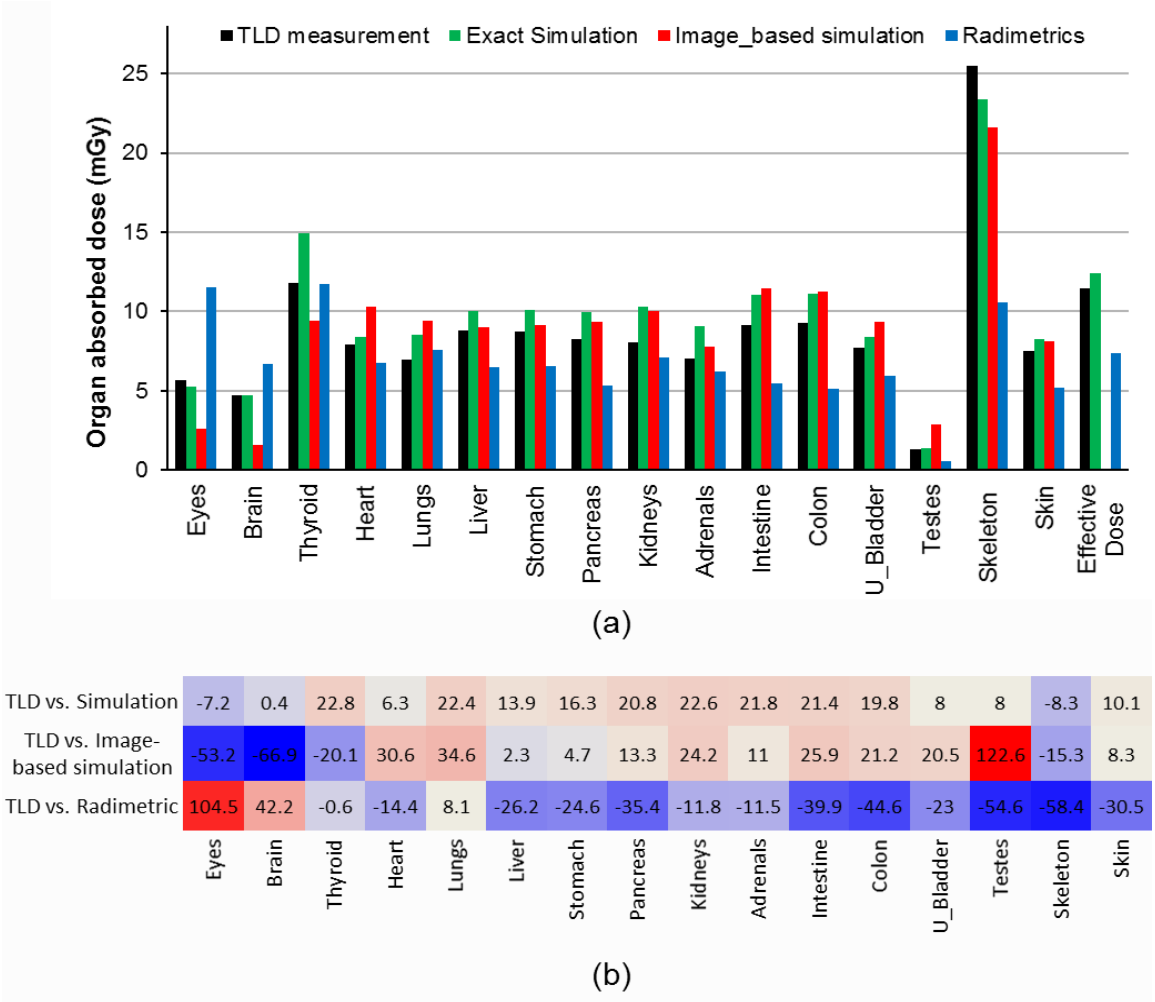


Figure 4. Comparison between (a) absolute organ absorbed doses and (b) relative differences of organ absorbed doses measured using TLDs against estimated using Monte Carlo simulations based on raw projection data information (exact simulation) and DICOM header information (image-based simulation) and Radimetrics™ dose-tracking software.

To illustrate the dosimetric impact of tube start angle parameter on the simulation results, Figure 5 shows the results from energy deposition tally over one complete rotation around the concerned organs. The absolute difference of the total deposited energy for the thyroid between simulation results based on exact tube start angle compared with those produced using 90° and 180° deviated start angle is about 137% and 27.5%, respectively. For the liver, this difference reduces to 26% and 23%, respectively.

The sources of uncertainties in organ dose simulation studied in this work originated from tube current modulation, tube start angle and the number of simulation points. To investigate the dosimetric impact of these factors, organ doses obtained from different simulation scenarios were compared against exact simulations in terms of percent difference, illustrated in Figure 6. By implementing only the longitudinal modulation scheme to the simulated dose values the absolute mean difference compared to the organ doses resulted from exact simulation was about 3.9% while the absolute difference for small peripheral organs like thyroid and testis exceeds 10% and 13 %, respectively. By implementing the reported effective time product tube current as a fixed tube current into the simulation, the absolute mean difference of estimated organ doses against exact simulation was calculated about 15.15% while it exceeds 50% for brain and eye lenses.

The dosimetric impact of the number of simulation points on the organ doses was investigated by modelling the different simulation intervals in the z-direction where the projection angles were conserved according to the exact tube start angle and fixed table speed. The absolute mean difference between the results from exact simulation compared to 2 mm interval, 6 mm interval and 10 mm interval is about 11.5, 11.45 and 13%. The mean

absolute difference between exact simulation against approximate ones starting at 50°, 90° and 180° with 2 mm interval of simulation points in z direction is about 12%, 13.8% and 14%, respectively.

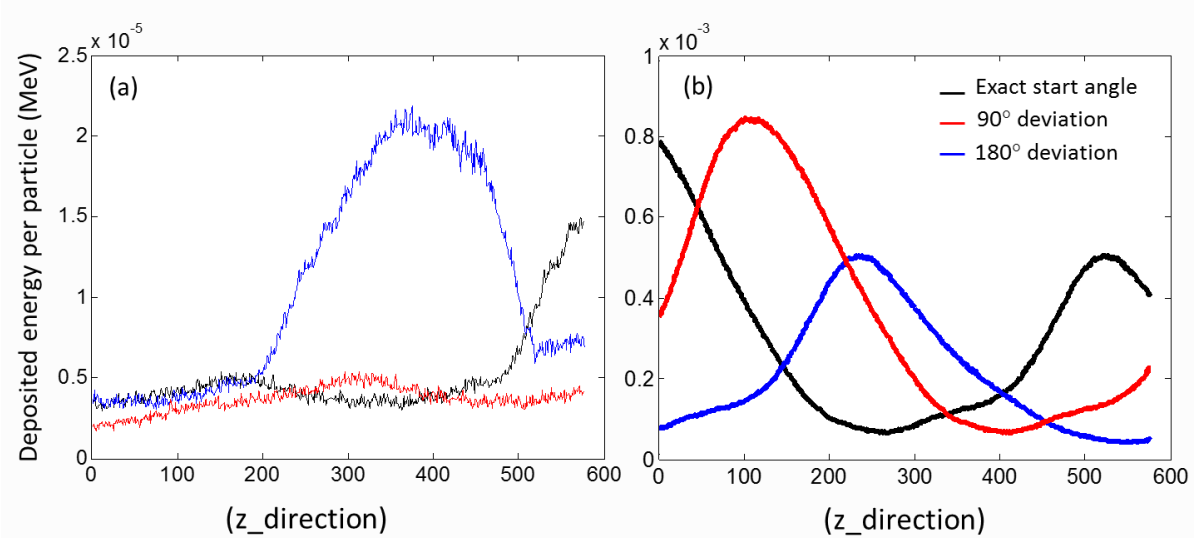


Figure 5. Deposited energy per particle over one rotation for different tube start angles in (a) the thyroid and (b) the liver.

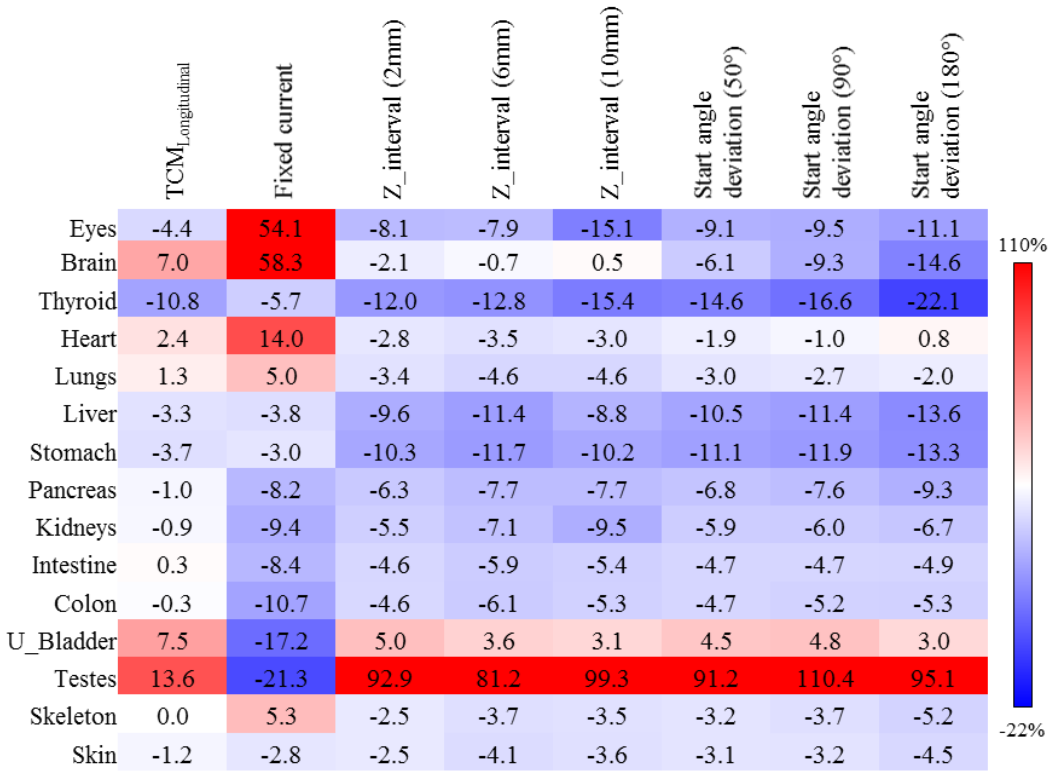


Figure 6. Heat map displaying the relative differences between organ absorbed doses obtained from exact Monte Carlo simulations (information extracted from raw projection data) and image-based simulations (information extracted from CT image DICOM header information) considering longitudinal tube current modulation, fixed tube current (effective mAs), simulation point intervals in Z direction (2 mm, 6 mm and 10 mm) and start angle deviation from the actual angle (50°, 90° and 180°).

Statistical analysis

The differences between experimental measurements using TLDs and exact simulations were not statistically significant (P-value=0.53). The ICC showed that the results obtained from exact simulations are in excellent agreement with experimental measurements with a consistency of 0.98 and absolute agreement of 0.97.

IV. Discussion

We developed a Monte Carlo framework for radiation dose assessment from CT examinations toward patient-specific organ-level dose monitoring. To this end, we integrated a CT source model, computational anthropomorphic phantom and protocol parameters into a dedicated Monte Carlo program. Experimental measurements using an ionization chamber and TLD chips were performed to validate the developed methodology. Considering the dependency of the calibration factor on CT acquisition parameters, we calculated, for each scenario, a unique calibration factor associated with scan-specific parameters. Since the $CTDI_{vol}$ index represents the radiation output of a specific CT scan, we used this value to benchmark the developed CT source model. The calibration factor was determined as the ratio of the reported $CTDI_{vol}$ on the scanner control console and $CTDI_{vol}$ obtained from simulation. Furthermore, it has been previously reported that the variation of organ doses obtained from scanner-specific simulations across different scanner manufacturers are close to the variation in scanner-specific $CTDI_{vol}$ [32]. In this context, calculation of a unique calibration factor associated with a specific CT examination (scanner-specific and scan-specific) enables to utilize our Monte Carlo program for a variety of protocols and scanner models.

A reasonable agreement was observed between experimental in-phantom measurements and Monte Carlo simulations in organ-level dosimetry. The TLD results had an uncertainty of 7% where standard dosimeter calibration, background noise subtraction, and response correction for specific energy range of x-ray beams has been taken into account. Likewise, there is a statistical uncertainty associated with Monte Carlo simulation results, estimated to be less than 2% for all simulations carried out in this work. The simulated organ doses overestimated the measured ones in most of the cases. These deviations are mainly caused by the differences between the constructed computational model and the actual physical phantom employed for measurements. These errors are partly caused by the registration process between the computational model and the physical phantom, partly related to differences in terms of material composition, since the physical phantom is made of three different tissue equivalent materials (bone, lung, and soft tissue), while the elemental compositions of the different organs have been implemented into the simulations according to the ICRP report 89 [33]. For skeleton dose measurements, most of the TLDs were inserted in the spinal cord while in simulations all bones contributed to skeleton dose, resulting in 8.3% underestimation. Radimetrics™ software underestimated organ doses compared to TLD measurements by a mean absolute difference of 26% for most organs, except the brain and eyes. A slight misalignment of CT images of the physical phantom with the stylized phantom used by Radimetrics™ was observed in the head region, which explains the overestimation of the absorbed dose to these two organs by about 73%. However, the physical phantom utilized in this work is anthropomorphically similar to the reference man computational phantom, whereas the stylized computational phantom used by Radimetrics™ doesn't reflect the anatomical features of this model. Radimetrics™ provides a simple protocol-based registration of CT scan localizer or topogram to predefined anatomical landmarks in the stylized phantom without resorting to any form of deformable registration. The acquisition parameters that Radimetrics™ used in this simulation were extracted from DICOM header information and may introduce extra errors to the results. Therefore, the results presented in Figure 4 confirm the good agreement between exact simulations and experimental measurements (within the range [-8.3% - 22%]). The differences between estimated organ doses obtained from TLD measurements and exact simulations were not statistically significant (P-value=0.53).

The accuracy of organ dose estimation is directly dependent on the accuracy of the constructed computational model representing patient's anatomy and the modelling of exposure conditions [23]. The uncertainty associated with the construction of patient-specific computational phantoms was investigated in our previous study, where the mean absolute differences between organ doses estimated from a reference model (manual segmentation) and those estimated from the constructed patient-specific model were within the range [0.5% - 29%] with a mean value of 9.1% [5]. In this work, we further analysed the organ dose uncertainties associated with irradiation conditions. In this context, the patient's dose profile calculated from exact simulations (input parameters obtained from raw CT projection data) against image-based simulations (input parameters derived from DICOM header of CT images) was investigated. The extra errors introduced to the simulation results caused by the smoothed tube current modulation, lack of knowledge about tube start angle and ignoring the overranging distance were considered. In the exact simulations, we simulated 16'854 projection points where the mAs values were known for each point. For the image-based simulation, the number of simulation points was determined based on the intervals in the Z direction (scan length / Z-interval) and the mAs values were reported in each axial slice of CT images. According

to Figure 6, using only longitudinal current modulation introduces an extra error within the range [-10% - 13%] (mean=3.85%) to the simulation results. The value of total tube current time product of the examination obtained from CT raw data is 6.6% higher than that calculated from CT DICOM images owing to the smoothing of the angular current modulation. As seen in Figure 6, there are substantial differences between organ doses produced by exact simulations and the results based on the deviated tube starting angles within the range [-22% - 110%]. The impact of tube start angle on the calculated doses for small and superficial organs (e.g. thyroid and testes) is significant. The overranging length and dynamic collimation was modelled in this simulation. The dose efficiency of dynamic collimation has been reported to be in the range of 90% in case of full beam collimation on Siemens CT scanners, which significantly reduces the dosimetric impact of overranging [27]. The impact of the number of simulation points on organ doses was investigated in the condition where the start angle was exactly modelled according to the information extracted from CT raw data and accordingly the angular positions of the simulated source points were exactly matched to the exact simulation scenario. By increasing the simulation intervals in the z-direction, the uncertainty for estimation of the organ dose slightly increases. However, the difference between 2 mm interval and 6 mm interval does not show any significant impact on the dosimetric results. For large organs, the impact of simulation parameters on organ absorbed doses is less than that for small organs like the testis and thyroid. Since Monte Carlo simulations provide the mean deposited energy per particle for a specific source position, organ absorbed doses are calculated based on the summation of the deposited energy multiplied by the tube current time product for all simulated source positions. Therefore, it is expected that for large organs, the simulation parameters are compensated during this summation while for small organs or partially irradiated organs in the border of CT examination, the simulation parameters play important roles in organ level dosimetry. We also anticipate improved modelling accuracy using the new version of MCNP code (version 6.2).

This study bears a number of inherent limitations. First, the experimental measurements using the anthropomorphic phantom were performed only once using a limited number of TLDs, which might introduce some statistical uncertainties. Second, this study is limited to a single CT scanner and a single set of acquisition parameters. Third, a single physical phantom (adult male) was studied while it can be extended to other categories, e.g. female and paediatrics. Lastly, the personalized computational model was constructed using deformable registration where the uncertainties associated with the registration algorithm introduced some extra errors to the simulation results. Thanks to advances in deep learning algorithms, patient-specific dosimetry is becoming feasible in the clinic. Using deep neural network algorithms, patient-specific computational models can be constructed from CT images through automated segmentation [15, 34]. Furthermore, the dose map of an individual patient commonly obtained from computationally expensive Monte Carlo simulations can be directly generated through deep learning approaches [35].

V. Conclusion

An experimental setup was performed in this work to evaluate the accuracy of Monte Carlo-based personalized organ-level dosimetry from CT examinations. Individual patient dose profiles can be accurately estimated using the developed simulation framework. Investigations considering different CT scanners and scanning protocols can be conducted to optimize CT technologies and scanning protocols. The validated CT scanner model could be employed in personalized CT dosimetry where the patient-specific computational model is constructed using different approaches. We also assessed the dosimetric impact of input parameters in organ-level dose simulation. It can be concluded that, when the information from the CT raw projection data is not available, the simulation results could be acceptable if the input parameters obtained from CT image DICOM header are correctly employed in the simulation setup. In this context, the longitudinal tube current modulation should be implemented at least by averaging simulations with three random tube start angles. The number of simulation points should be defined appropriately in the Z direction and in case of dynamic collimation, over-ranging length can be ignored. Hence, the methodology can be further expanded to produce an accurate MC simulation toolkit with a reduced computational burden.

Acknowledgments

This work was supported by the Swiss National Science Foundation under grant SNRF 320030_176052, Qatar National Research Fund under grant NPRP10-0126-170263 and Iran's Ministry of Science.

References

[1] A. Ferrero, N. Takahashi, T. J. Vrtiska, A. E. Krambeck, J. C. Lieske, and C. H. McCollough, "Understanding, justifying, and optimizing radiation exposure for CT imaging in nephrourology.,"*Nature reviews. Urology*, vol. 16, no. 4, pp. 231-244, 2019.

[2] IMV Publising, "CT Market Outlook Report, IMV Medical Information Division," Arlington, VA, 2018.

[3] A. B. De González *et al.*, "Projected cancer risks from computed tomographic scans performed in the United States in 2007,"*Arch Intern Med*, vol. 169, no. 22, pp. 2071-2077, 2009.

[4] A. Sodickson *et al.*, "Recurrent CT, cumulative radiation exposure, and associated radiation-induced cancer risks from CT of adults,"*Radiology*, vol. 251, no. 1, pp. 175-184, 2009.

[5] T. Xie, A. Akhavanallaf, and H. Zaidi, "Construction of patient-specific computational models for organ dose estimation in radiological imaging.,"*Med Phys*, vol. 46, no. 5, pp. 2403-2411, 2019.

[6] A. Lahham and H. ALMasri, "estimation Of Radiation Doses From Abdominal Computed Tomography Scans,"*Radiat Prot Dosim*, vol. 182, no. 2, pp. 235-240, 2018.

[7] J. DeMarco *et al.*, "A Monte Carlo based method to estimate radiation dose from multidetector CT (MDCT): cylindrical and anthropomorphic phantoms,"*Phys Med Biol*, vol. 50, no. 17, p. 3989, 2005.

[8] P. Deak, M. Van Straten, P. C. Shrimpton, M. Zankl, and W. A. Kalender, "Validation of a Monte Carlo tool for patient-specific dose simulations in multi-slice computed tomography," *Eur Radiol* vol. 18, no. 4, pp. 759-772, 2008.

[9] X. Li *et al.*, "Patient-specific radiation dose and cancer risk estimation in CT: Part I. Development and validation of a Monte Carlo program,"*Med phys*, vol. 38, no. 1, pp. 397-407, 2011.

[10] D. J. Long *et al.*, "Monte Carlo simulations of adult and pediatric computed tomography exams: validation studies of organ doses with physical phantoms,"*Med phys*, vol. 40, no. 1, p. 013901, 2013.

[11] W. Segars *et al.*, "Patient specific computerized phantoms to estimate dose in pediatric CT," presented at the Physics of Medical Imaging, 2009.

[12] X. Li *et al.*, "Patient-specific radiation dose and cancer risk estimation in CT: Part II. Application to patients,"*Med phys*, vol. 38, no. 1, pp. 408-419, 2011.

[13] W. A. Kalender, N. Saltybaeva, D. Kolditz, M. Hupfer, M. Beister, and B. Schmidt, "Generating and using patient-specific whole-body models for organ dose estimates in CT with increased accuracy: Feasibility and validation,"*Phys Medica*, vol. 30, no. 8, pp. 925-933, 2014.

[14] T. Xie, P. A. Poletti, A. Platon, C. D. Becker, and H. Zaidi, "Assessment of CT dose to the fetus and pregnant female patient using patient-specific computational models.,"*Eur Radiol*, vol. 28, no. 3, pp. 1054–1065, 2018.

[15] T. Xie and H. Zaidi, "Estimation of the radiation dose in pregnancy: an automated patient-specific model using convolutional neural networks.,"*Eur Radiol*, vol. 29, no. 12, pp. 6805–6815, 2019.

[16] A. Akhavanallaf, T. Xie, and H. Zaidi, "Development of a library of adult computational phantoms based on anthropometric indexes,"*IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 3, no. 1, pp. 65-75, 2019.

[17] Bayer HealthCare. "Radimetrics Enterprise Platform: Dose Management Solution." Bayer HealthCare. http://www.radiologysolutions.bayer.com/products/ct/dosemanagement/rep/. (accessed 2016).

[18] M. Cristy and K. Eckerman, "Specific absorbed fractions of energy at various ages from internal photon sources. I Methods, II one year old, III five year old, IV ten year old, V fifteen year old male and adult female, VI new-born and VII adult male.,"*Oak Ridge National Laboratory, Oak Ridge, TN*, 1987.

[19] A. J. Hardy *et al.*, "Estimating lung, breast, and effective dose from low-dose lung cancer screening CT exams with tube current modulation across a range of patient sizes,"*Med Phys*, vol. 45, no. 10, pp. 4667-4682, 2018.

[20] K. Fujii, K. Nomura, Y. Muramatsu, S. Obara, K. Akahane, and M. Kusumoto, "Organ Dose Evaluations Based on Monte Carlo Simulation for CT Examinations Using Tube Current Modulation.,"*Radiat Prot Dosimetry*, vol. 174, no. 3, pp. 387-394, 2017.

[21] W. Chen, D. Kolditz, M. Beister, R. Bohle, and W. A. Kalender, "Fast on-site Monte Carlo tool for dose calculations in CT applications.,"*Med Phys*, vol. 39, no. 6, pp. 2985-2996, 2012.

[22] M. Bostani, K. McMillan, J. J. DeMarco, C. H. Cagnon, and M. F. McNitt-Gray, "Validation of a Monte Carlo model used for simulating tube current modulation in computed tomography over a wide range of phantom conditions/challenges.,"*Med Phys*, vol. 41, no. 11, p. 112101, 2014.

[23] AAPM, "Report No. 246 - Estimating Patient Organ Dose with Computed Tomography: A Review of Present Methodology and Required DICOM Information," AAPM, AAPM Report No. 246, 2019. [Online]. Available: https://www.aapm.org/pubs/reports/detail.asp?docid=190

[24] J. S. Muryn *et al.*, "Analysis of uncertainties in Monte Carlo simulated organ and effective dose in chest CT: scanner-and scan-related factors,"*Phys Med Biol*, vol. 62, no. 8, p. 3175, 2017.

[25] C. Lee *et al.*, "Organ dose estimation accounting for uncertainty for pediatric and young adult ct scans in the united kingdom,"*Radiat Prot Dosim*, vol. 184, no. 1, pp. 44-53, 2018.

[26] G. Poludniowski, G. Landry, F. DeBlois, P. Evans, and F. Verhaegen, "SpekCalc: a program to calculate photon spectra from tungsten anode x-ray tubes,"*Phys Med Biol*, vol. 54, no. 19, p. N433, 2009.
[27] Siemens AG, "Computed Tomography, SOMATOM Definition Edge, System Owner Manual," Forchheim/Germany, 2012.
[28] K. Yang, X. Li, X. George Xu, and B. Liu, "Direct and fast measurement of CT beam filter profiles with simultaneous geometrical calibration,"*Med phys*, vol. 44, no. 1, pp. 57-70, 2017.
[29] L. S. Waters, "MCNPX user's manual,"*Los Alamos National Laboratory*, 2002.
[30] J. A. Bauhs, T. J. Vrieze, A. N. Primak, M. R. Bruesewitz, and C. H. McCollough, "CT dosimetry: comparison of measurement techniques and devices,"*Radiographics*, vol. 28, no. 1, pp. 245-253, 2008.
[31] A. Schilham, A. J. van der Molen, M. Prokop, and H. W. de Jong, "Overranging at multisection CT: an underestimated source of excess radiation exposure,"*Radiographics*, vol. 30, no. 4, pp. 1057-1067, 2010.
[32] A. C. Turner *et al.*, "The feasibility of a scanner-independent technique to estimate organ dose from MDCT scans: Using to account for differences between scanners,"*Med phys*, vol. 37, no. 4, pp. 1816-1825, 2010.
[33] ICRP, "Basic anatomical and physiological data for use in radiological protection: reference values. ICRP Publication 89", *Annals of the ICRP*, vol. 32, no. 3-4, pp. 1-277, 2002.
[34] G. Zhang *et al.*, "Preliminary study on automatic organ segmentation using machine learning and application in personalized dosimetry for nuclear medicine," in *7th International Workshop on Computational Human Phantoms*, Neuherberg, Germany, , 22-24 July, 2019.
[35] M. S. Lee, D. Hwang, J. H. Kim, and J. S. Lee, "Deep-dose: a voxel dose estimation method using deep convolutional neural network for personalized internal dosimetry,"*Sci Rep*, vol. 9, no. 1, p. 10308, 2019.

Chapter 6

The promise of artificial intelligence and deep learning in PET and SPECT imaging

Hossein Arabi, Azadeh Akhavanallaf, Isaac Shiri, Amirhossein Sanaat, and Habib Zaidi

Contribution: Study conception and design, literature review, analysis and interpretation of results, manuscript preparation.

Phys Med, Vol. 83, pp 122-137 (2021)

Abstract

This review sets out to discuss the foremost applications of artificial intelligence (AI), particularly deep learning (DL) algorithms, in single-photon emission computed tomography (SPECT) positron emission tomography (PET) imaging. To this end, the underlying limitations/challenges of these imaging modalities are briefly discussed followed by a description of AI-based solutions proposed to address these challenges. This review will focus on mainstream generic fields, including instrumentation, image acquisition/formation, image reconstruction and low-dose/fast scanning, quantitative imaging, image interpretation (computer-aided detection/diagnosis/prognosis), as well as internal radiation dosimetry. A brief description of deep learning algorithms and the fundamental architectures used for these applications is also provided. Finally, the challenges, opportunities, and barriers to full-scale validation and adoption of AI-based solutions for improvement of image quality and quantitative accuracy of PET and SPECT images in the clinic are discussed.

I. Introduction

Artificial intelligence (AI) approaches, particularly deep learning (DL) techniques, have received tremendous attention during the last decade owing to their remarkable success in offering novel solutions to solve complex problems. Novel AI/DL-based solutions have created opportunities in clinical and research settings to automate a number of tasks deemed to depend on human cognition, and hence require his intervention to facilitate the decision-making process [1]. State-of-the-art AI/DL algorithms have exhibited exceptional learning capability from high dimensional and/or highly complex data, accomplishing daunting challenging tasks in image and data analysis/processing in general and multimodality medical imaging in particular.

In the context of medical imaging, challenging tasks, such as image segmentation/classification, data correction (such as noise or artifact reduction), image interpretation (prognosis, diagnosis, and monitoring of response to treatment), cross-modality image translation or synthesis, and replacing computationally demanding algorithms (such as Monte Carlo calculations) have been broadly revisited and evolved ever since the adoption of deep learning approaches [2, 3]. AI-based solutions have been proposed to address the fundamental limitations/challenges faced by image acquisition and analysis procedures on modern molecular imaging technologies. Considering the superior performance of deep learning approaches compared to conventional techniques, a paradigm shift is expected to occur provided that task-specific pragmatic developments of these algorithms continue to evolve in the right direction.

Single-photon emission computed tomography (SPECT) and Positron emission tomography (PET) imaging provide the *in vivo* radiotracer activity distribution maps, representative of biochemical processes in humans and animal species. The introduction of hybrid imaging combining functional and anatomical imaging modalities in the form of combined PET/CT and PET/MRI systems has remarkably thrived the widespread adoption and proliferation of these modalities in clinical practice. In this light, AI-based algorithms/solutions are developed to overcome the major shortcomings or to enhance the current functionality of these modalities.

The applications of AI-based algorithms in PET and SPECT imaging ranges from low-level electronic signal formation/processing to high-level internal dosimetry and diagnostic/prognostic modeling. For developments in instrumentation, deep learning approaches have been mostly employed to improve the timing resolution and localization accuracy of the incident photons aiming at enhancing the overall spatial and time-of-flight (TOF) resolutions in PET. Image reconstruction algorithms are being revisited through the introduction of deep learning algorithms wherein the whole image reconstruction process or certain critical components (analytical models) are being replaced by machine learning models. A large body of literature is dedicated to quantitative SPECT and PET imaging aiming at reducing the impact of noise, artifact, and motion, or to correct for physical degrading factors, including attenuation, Compton scattering, and partial volume effects. The lack of straightforward techniques for generation of the attenuation map on organ-specific standalone PET scanners or hybrid PET/MRI systems inspired active scientists in the field to devise suitable strategies to enhance the quantitative potential of molecular imaging. High-level image processing tasks, such as segmentation, data interpretation, image-based diagnostic and prognostic models as well as internal dosimetry based on SPECT or PET imaging have substantially evolved owing to the formidable power and versatility of deep learning algorithms.

AI/DL-based solutions have been proposed to undertake certain tasks belonging to the long chain of processes involved in image formation, analysis, and extraction of quantitative features for the development of disease-specific diagnosis/prognosis models from SPECT and PET imaging. In this review, the applications of AI/DL in these imaging modalities are summarized in six key sections focusing on the major challenges/opportunities and seminal contributions in the field. A concise overview of machine learning methods, in particular deep learning approaches, is presented in section 2. The following section describes AI-based techniques employed in PET instrumentation, image acquisition and formation, image reconstruction and low-dose scanning, quantitative imaging (attenuation and scatter corrections), image analysis and computer-aided detection/diagnosis/prognosis, as well as internal radiation dosimetry. The last section provides in perspective the major challenges and opportunities for AI/DL-based solutions in PET and SPECT imaging.

II. Principles of machine learning and deep learning

Machine learning algorithms are considered as a subset of non-symbolic artificial intelligence, which tends to automatically recognize a pattern and create/extract a desirable representation from raw data [4]. In machine learning algorithms, the system attempts to learn certain patterns from the extracted features. Likewise, in deep learning algorithms, a subtype of machine learning techniques, feature extraction, feature selection, and ultimate tasks of classification or regression are carried out automatically in one step [5]. Different deep learning algorithms have been proposed and applied in nuclear medicine [2, 6], including convolutional neural networks (CNNs) [7], convolutional encoders-decoders (CEDs) [8], and generative adversarial networks (GANs) [5]. Some applications of machine learning algorithms, such as classification, segmentation, and image-to-image translation, have attracted more attention [9].

A number of deep learning architectures became popular in the field of medical image analysis, including CED networks consisting of encoder and decoder parts designed to convert input images to feature vectors and feature vectors to target images, respectively [8]. In addition, GANs consist of two major components: a generator, mostly a CED network, and a discriminator, a classifier to differentiate the ground truth from the synthetic images/data [8]. Different architectures based on these models were developed and applied on medical images for different tasks, including image segmentation and image to image translation [10]. U-Net [11] is one the most popular architectures built upon the CED structure via adding some skip connections for context capturing and for creating a symmetric expanding path, which enables more efficient feature selection. Upgrading networks with different modules, such as attention blocks/components [12] for highlighting salient features in the input data, and residual connections [13] to prevent gradient vanishing, are intended to improve the overall performance of the networks. Conventional GAN architectures have been upgraded in different ways, leading to conditional GAN (cGAN) [14] and cycle consistency GANs (Cycle-GAN) [15] models, which consist of a CED in the generator and discriminator components and task-specific loss functions. Cycle-GAN [15] is an unsupervised model for image-to-image transformation, which does not require paired (labeled) datasets. In the Cycle-GAN model, two generator and discriminator components are jointly involved in the training process, wherein images from two different domains are used as input and output within a cycle consistency scheme. In the cycle consistency scheme, the output of the generator component is used as input and vise versa with the calculated loss between the input and output acting as regularization of the generator model [15].

Overall, deep learning-based algorithms outperformed conventional approaches in various applications [5]. AI-based approaches, especially deep learning algorithms, do not require handcraft features extraction, specific data preprocessing, or user intervention within the learning and inferring processes [5]. The major applications of deep learning approaches in SPECT and PET imaging are summarized in figure 1. Deep learning methods face many challenges, including the fact that they are data hungry, require high computation burden for the training process, and their black box nature (which hampers systematic analysis of their operation/performance) [7]. To reach peak performance, these algorithms require a large number of clean and cured datasets for the training process. However, data collection remains the main challenge owing to patients' privacy and complexity of ethical issues. Moreover, task-specific deep learning algorithms (i.e. for a particular organ/body region or radiotracer) are able to exhibit superior performance compared to more general models which are commonly more sensitive variability in image acquisition and reconstruction protocols scanner model, etc. Another challenge faced by the application of deep learning algorithms in medical imaging is the high computational burden owing to the large size of clinical data in terms of number of subjects and individual images (large 3-dimensional images or sinograms) which might cause memory or data management issues.

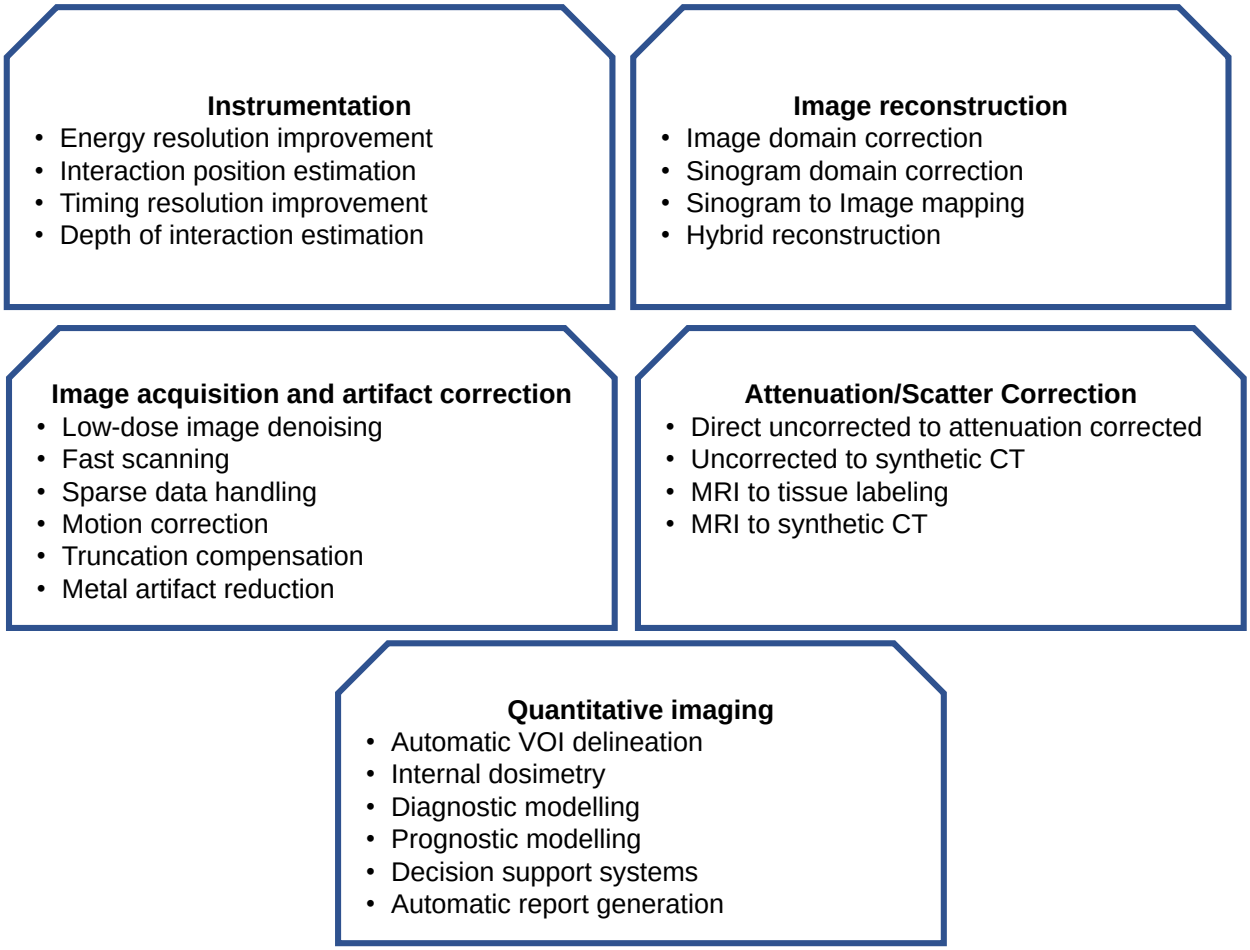


Figure 1. Main applications of deep learning-based algorithms in PET and SPECT imaging.

III. Applications of deep learning in SPECT and PET imaging

Instrumentation and image acquisition/formation

Detector modules play a key role in the overall performance achieved by PET scanners. An ideal PET detector should have a good energy and timing resolution and capable of accurate event positioning. Energy resolution is a metric that determines how accurately a detector can identify the energy of incoming photons and as a result, distinguish scatter and random photons from true coincidences. These parameters affect the scanner's sensitivity, spatial resolution, and signal-to-noise ratio (true coincidence versus scatters or randoms). Despite significant progress in PET instrumentation, there are a number of challenges that still need to be addressed and where machine learning approaches can offer alternative solutions to complex and multi-parametric problems.

Accurate localization of the interaction position inside the crystals improves the overall spatial resolution of PET scanners. Since optical photons distribution is stochastic, particularly near the edges of the crystal, and owing to multiple Compton scattering and reflection, accurate positioning of the interaction within the crystal is challenging. In comparison with other positioning algorithms, such as Anger logic and correlated signal enhancement, which rely on determination of the centre of gravity, machine learning algorithms led to a better position estimation particularly at the crystal edges [16]. In this regard, Peng *et al.* trained a CNN classifier that was fed with signals from each Silicon photomultiplier's channel to the coordinates of the scintillation point for a quasi-monolithic crystal [17]. Another study applied a multi-layer perceptron to predict the 3D coordinates of the interaction position inside a monolithic crystal and compared the performance of this positioning algorithm with anger logic for a preclinical PET scanner based on NEMA NU4 2008 standards [18]. Figure 2 depicts the adopted

deep learning-based event-positioning scheme in monolithic detectors. To address the challenge of determining the depth of interaction, a gradient tree boosting supervised machine learning algorithm was used to extract the scintillation position, resulting in a spatial resolution of 1.4 mm full width half maximum (FWHM) for a 12 mm thick monolithic block [19]. Recently, Cherenkov-based detectors attracted much attention owing to their superb performance in terms of time and spatial resolution. Hashimoto *et al.* studied the performance of a deep learning model for 3D positioning in this type of detectors through a Monte Carlo simulation study [20]. They demonstrated that in comparison with conventional positioning methods, such as the centre of gravity determination and principal component analysis, the deep learning model led to significantly improved spatial resolution.

Time resolution is another crucial factor in PET instrumentation which determines the achievable performance using TOF imaging as well as the efficiency of randoms and scatter rejection. This factor depends on the physical characteristics of the scintillator, photodetector quantum efficiency, and electronic circuits that convert the scintillation light to electrical signals. Considering the physics of photon interactions within a crystal, only a portion of produced scintillation photons reach the photodetector and contribute to positioning and timing. The consequence of this is noticeable statistical uncertainty and noise-induced bias. Straightforward approaches, such as feeding a CNN model with detector signals to estimate TOF information produced promising results. In a recent study, a training dataset (reference) obtained by scanning a ^{68}Ga point source shifted repeatedly with steps of 5 mm across the field-of-view of the PET scanner was used to train a deep learning algorithm [21]. The authors reported a TOF resolution of about 185 ps, exhibiting significant improvement with respect to conventional methods with a resolution of 210 to 527 ps. Gladen *et al.* developed a machine learning method, referred to as self-organized map (SOM) algorithm, for estimating the arrival time of annihilation photons in a high purity germanium detector (HPGe). SOM was able to cluster the TOF bins based on the signal shape and its raising edge [22].

Recent studies substantiated the applicability of deep learning techniques to reliably estimate the interaction position, energy, and arrival time of incident photons within the crystal with improved accuracy and robustness to noise. One of the major difficulties in developing such models is the creation of labelled data (used as reference), which require extensive experimental measurements. For example, preparing a training dataset for position estimation requires a precise and reproducible setup of a single pencil beam and several measurements at any possible spot within the field-of-view. A number of recent studies came up with novel ideas to perform these tasks for monolithic crystal through using uniform or fan-beam sources or applying clustering to the training dataset [19]. Likewise, for TOF training dataset, hundreds of point source positionings and data acquisitions are required to create a realistic range of TOF variations. In this regard, artificial ground-truth data creation was proposed through switching the PET detector waveforms forward and backward in the time domain [21].

Sophisticated machine learning-based algorithms for event positioning, timing, and/or calibration are envisioned on next generation SPECT and PET systems on the front-end electronics using dedicated application-specific integrated circuits (ASICs) and field-programmable gate arrays (FPGAs) [23]. Furthermore, developing a single model for extracting time, position, and energy simultaneously from photodetector outputs would be an interesting approach that can potentially improve the overall performance of the nuclear imaging systems.

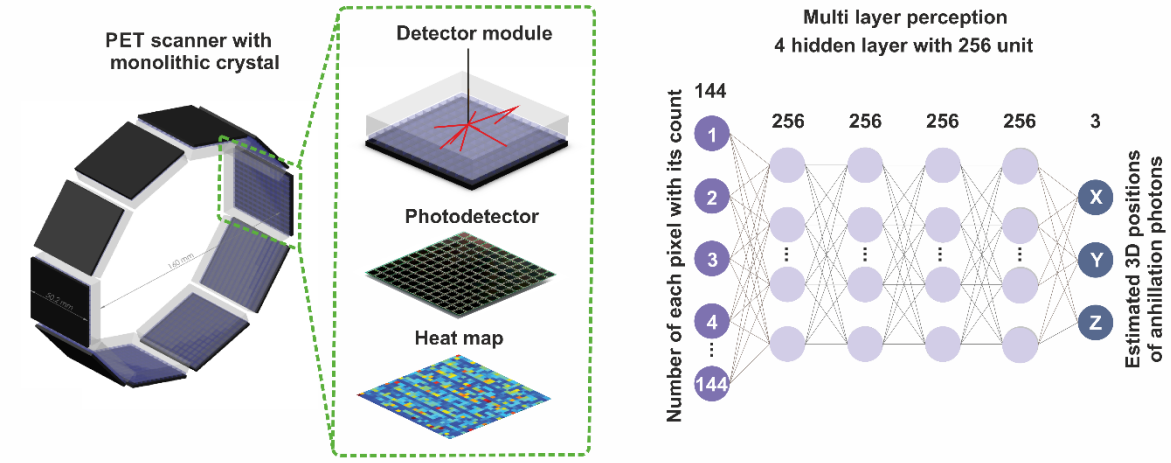


Figure 2. Deep learning-based event positioning in monolithic detectors.

Image reconstruction and low-dose/fast image acquisition

Deep learning algorithms have recognized capabilities in solving complex inverse problems, such as image reconstruction from projections. The process of image reconstruction for CT, PET, and SPECT using deep learning techniques entails roughly the same procedure. Overall, four strategies were adopted for image reconstruction using deep learning algorithms. The first approach consists of image-to-image translation in the image space, wherein a model is trained to convert reconstructed images to another representation to improve image quality through, for instance, noise removal, supper resolution modelling, motion correction, etc. [24]. The second approach implements the training of the deep learning model in the projection space prior to image reconstruction to avoid the sensitivity and dependence on reconstruction algorithms. In the third approach, a model learns to develop non-linear direct mapping between information in the sinogram and image domains [25, 26]. The fourth approach, referred to as hybrid domain learning, relies simultaneously on analytical reconstruction and machine learning approaches to reach an optimal solution for the image reconstruction problem [27, 28].

Two companies released AI-based solutions for image reconstruction in CT that were approved by the FDA [29, 30]. DeepPET is one of the earliest works suggesting direct reconstruction from sinograms to images through a deep learning approach [25]. Likewise, FastPET, is a machine learning-based approach for direct PET image reconstruction using a simple memory-efficient architecture implemented to operate for any tracer and level of injected activity [31].

Decreasing the injected activities is often desired owing to potential hazards of ionizing radiation for pediatric patients or subjects undergoing multiple serial PET or SPECT scans over time for monitoring of disease progression or in longitudinal studies. Moreover, decreasing the acquisition/scanning time increases scanners throughput and enhances patients' comfort, particularly elderly patients and those suffering from neurodegenerative diseases where the risk of involuntary motion during scanning is more common.

Reducing the injected activity amplifies Poisson noise, thus impacting image quality, lesion detectability, and quantitative accuracy of PET images. Devising optimized low-dose scanning protocols that preserve the critical information in the images is desirable. Although there is a fundamental difference between fast and low-dose scanning, both approaches have been interchangeably used in the literature. While both strategies produce noisy images, the content and information collected by these scanning modes are completely different. In a fast scan, the acquired data reflect the radiotracer kinetics in a short time course. For instance, if the scan starts right after injection, much information would be missing owing to insufficient and/or slow uptake in some organs. Fast acquisition protocols are also less sensitive to motion artifacts, though the patient's effective dose is similar to standard protocols. Conversely, low-dose scanning is performed with standard acquisition time, with a much lower injected activity, which obviously decreases the effective dose.

There might be a need to redesign/optimize reconstruction algorithms for low-dose scanning to reach an optimal trade-off between noise level and signal convergence. In low-dose/fast imaging, much critical information

would be buried under the increased noise level wherein an efficient denoising algorithm would be able to recover genuine signals [32].

To address the above-mentioned challenges, a number of denoising techniques to generate full-dose PET images from corresponding noisy/low-dose counterparts have been proposed. Conventional techniques include post-reconstruction processing/filtering algorithms [33, 34], anatomically-guided algorithms [35], statistical modelling during iterative reconstruction [36], and MRI-guided joint noise removal and partial volume correction [37]. Although these approaches attempted to minimize noise and quantitative bias, they still suffer from loss of spatial resolution and over-smoothing. By introducing image super-resolution techniques, such as sparse representation [38], canonical correlation analysis [39], and dictionary learning [40], effective noise reduction and signal recovery in low-dose images is expected with minimum artifacts or information loss. The widespread availability of hybrid imaging enabled to incorporate anatomical information in the reconstruction of low-dose PET images [41].

In the last few years, AI algorithms have been widely used in the field of image reconstruction and enhancement of image quality [42]. In most previous works, low-dose images were considered as the model's input whereas full-dose images were considered as the target to perform an end-to-end mapping between low-dose and full-dose images [43-46]. Such models with a single input channel (only low-dose images) suffer from the lack of sufficient information (for instance anatomical structures) to distinguish noise from genuine biological signals. Therefore, adding anatomical priors into the training procedure would make the model more accurate and robust. For resolution recovery, high-resolution anatomical information obtained from MR imaging was employed along with spatially-variant bluring kernels to avoid information loss during image reconstruction [47]. Some groups devised strategies for deep learning-guided denoising models for synthesizing full-dose sinograms from their corresponding low-dose sinograms [48].

An elegant study by Xu *et al.* proposed a U-Net model with concatenation connection and residual learning for full-dose reconstruction from a single 200^{th} low-dose image [49]. Xiang *et al.* presented a novel deep auto-context CNN model for synthesizing full-dose images from low-dose images complementing T1-weighted MR images. In comparison with state-of-the-art methods, their proposed model was able to generate comparable image quality while being 500 faster [44]. Another study employed a multi-input U-Net to predict 2D transaxial slices of ^{18}F-Florbetaben full-dose PET images from corresponding low-dose images, taking advantage of available T1, T2, and Diffusion-weighted MR sequences [43]. Liu *et al.* employed three modified U-Net architectures to enhance the noise characteristics of PET images through concurrent MR images without the need for full-dose PET images with a higher signal-to-noise ratio [50]. In addition, Cui *et al.* [51] proposed a 3D U-Net model for denoising of PET images acquired with two different radiotracers (^{68}Ga-PRGD2 and ^{18}F-FDG) where the model was trained with MR/CT images and prior high-quality images as input and noisy images as training labels. Using original noisy images instead of high-quality full-dose images makes the training of the model more convenient. Unsupervised networks are always desirable in medical image analysis due to the fact that data collection with accurate labels is challenging and/or time-cosuming. A foremost drawback of the above-mentioned models is that model training was performed in 2D rather than 2.5D or 3D.

The 3D U-Net architecture was able to reduce the noise and PET quantification bias while enhancing image quality of brain and chest ^{18}F-FDG PET images [52]. To compensate for the limited training dataset, they pre-trained the model using simulation studies in the first stage and then fine-tuned the last layers of the network with realistic data. Kaplan *et al.* [53] trained a residual CNN separately for various body regions, including brain, chest, abdomen, and pelvis to generate full-dose images from $1/10^{th}$ of the standard injected tracer activity. Training and testing of the model were performed on only two separate whole-body ^{18}F-FDG PET datasets.

GAN networks are widely used for image-to-image transformation tasks, especially image denoising. Conditional GANs (cGAN) and cycle GANs (Cycle-GAN) are two well-established architectures commonly used for style and domain transformation. In cGAN, unlike regular GAN, the generator and discriminator's output is regularized by an extra-label. For instance, Wang *et al.* estimated the generator error and used it beside the discriminator loss to train the generator of a 3D cGAN more efficiently for denoising low-dose brain PET images [45].

Cycle-GAN models do not necessarily require paired images as the model can learn in an unsupervised way to map input images from source to target domains. Because of the iterative feature extraction process and the presence of the inverse path in this architecture, the underlying characteristics of input/output data can be extracted

from unrelated images to be used in the image translation process. Zhou *et al.* proposed a 2D Cycle-GAN for generating full-dose with around 120 million true coincidences (for each bed position) from a low-dose image with only one million true coincidences [54]. Lei *et al.* claimed that their Cycle-GAN model is able to predict whole-body full-dose ^{18}F-FDG PET images from 1/8th of the injected activity [55]. They used a generator with residual blocks to learn the difference between low-dose and full-dose images to effectively reduce the noise. The same group presented a similar model incorporating CT images to guide low-dose to full-dose transformation using a relatively small dataset [56]. Their results revealed that the incorporation of CT images can improve the visibility of organ boundaries and decrease bias especially in regions located near bones.

More recent studies implemented the training process using deep learning models in the projection space instead of image space, demonstrating that training a model in the sinogram space could lead to more efficient learning compared to training in the image space. Sanaat *et al.* trained a U-Net model with a dataset consisting of 120 brain ^{18}F-FDG PET full-dose studies in the sinogram space [48]. The proposed model predicted full-dose from low-dose sinograms and demonstrated the superior performance of deep learning-based denoising in the sinogram space versus denoising in the image space (Figure 3). Furthermore, another study proposed a prior knowledge-driven deep learning model for PET sinogram denoising [57]. Hong *et al.* [58] combined Monte Carlo simulations and deep learning algorithms to predict high-quality sinograms from low-quality sinograms produced by two PET scanners equipped with small and large crystals, respectively. In whole-body PET imaging, Sanaat *et al.* compared the performance of two state-of-the-art deep learning approaches, namely Cycle-GAN and ResNet, to estimate standard whole-body ^{18}F-FDG PET images from a fast acquisition protocol with 1/8th of the standard scan time [59]. Cycle-GAN predicted PET images exhibited superior quality in terms of SUV bias and variability as well as the lesion conspicuity.

Though most of the above-described approaches could be applied to SPECT imaging, few studies dedicatedly addressed low-dose and/or fast SPECT imaging studies. Recently, a supervised deep learning network was employed to reduce the noise in myocardial perfusion SPECT images obtained from 1/2th, 1/4th, 1/8th, and 1/16th of the standard-dose protocol across 1052 subjects [60]. Similarly, Shiri *et al.* exploited a residual neural network to predict standard SPECT myocardial perfusion images from half-time acquisitions [61]. Raymann *et al.* used a U-Net architecture and XCAT phantom simulation studies of different regions of the body to reduce noise in SPECT images [62].

Generalizability and robustness of deep learning models are two significant factors that show how much a model is trustable and the results robust and reproducible for normal/abnormal unseen datasets. These two factors are largely linked to the diversity and number of training samples. It is very common to exclude abnormal cases prior to training or evaluation of a model to create a homogeneous training/test sample. Although this results in better results, it will reduce robustness to a realistic dataset with a broad range of abnormalities. It is strongly recommended to use both healthy/normal and unhealthy/abnormal subjects with a realistic distribution of the samples. Moreover, to avoid overfitting and guarantee effective training of the model, application of relevant data augmentation techniques is also recommended.

Using recurrent neural networks to decrease the scanning time and/or injected activity, especially in low-count dynamic PET imaging studies would be an interesting field of research. In addition, applying self-attention concepts to deep learning models would effectively enhance the performance of these models through indirect down-weighting/elimination of irrelevant regions and information in low-dose images while emphasizing the prominent/meaningful properties/information during the training process. Using realistic simulations to produce gold standard data sets beside clinical images would help deep learning models to learn noise distributions from a larger representative sample.

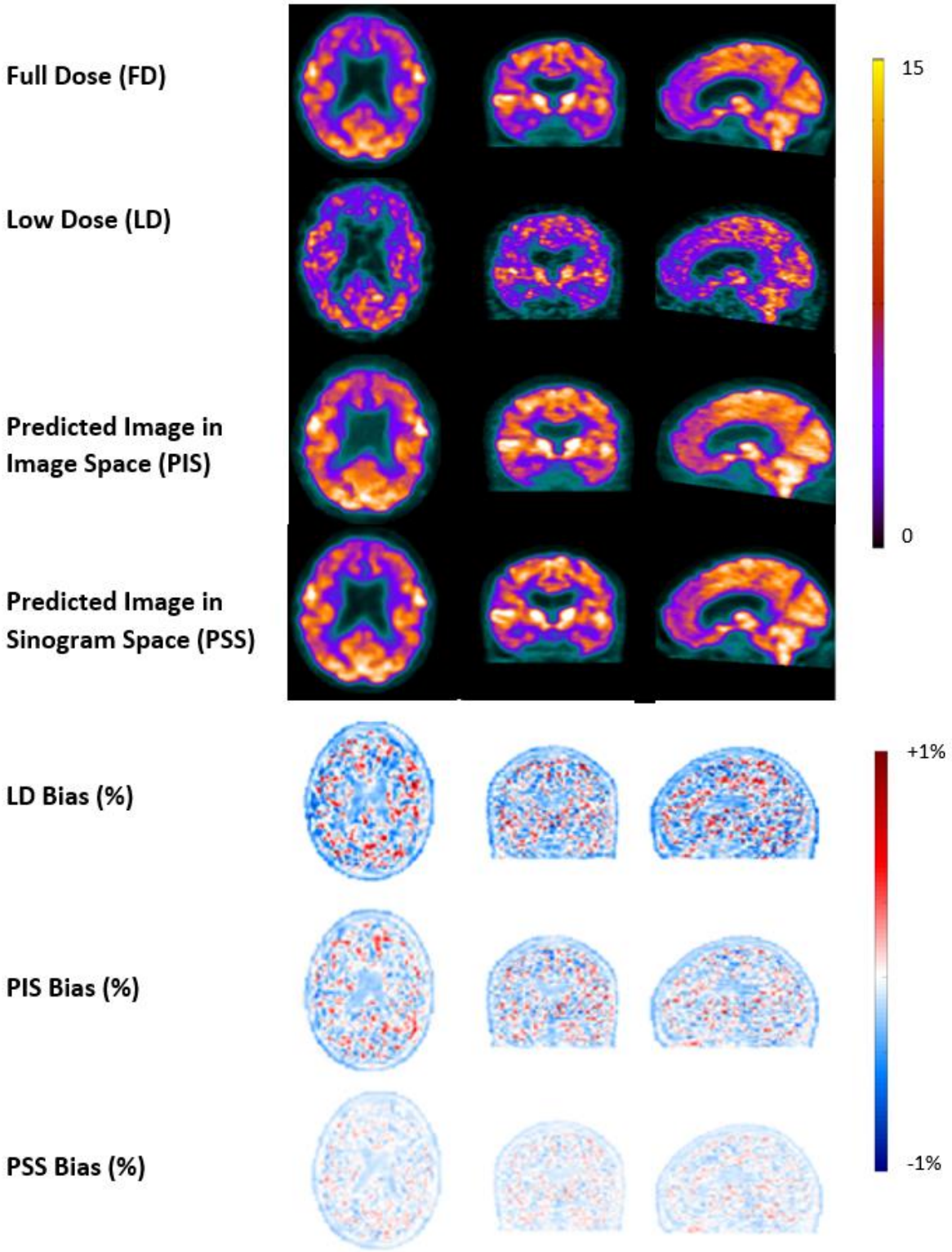


Figure 3. Comparison between full-dose and low-dose brain PET image predictions in the sinogram and image domains.

Quantitative imaging

A significant number of emitted photons undergo attenuation and Compton scatter interactions before they reach PET and SPECT detectors. Scatter and attenuation lead to over- and under-estimation of activity concentration, consequently resulting in large quantification errors [63]. To link the detected photons to the radiotracer activity concentration, attenuation and scatter correction (ASC) should be performed in SPECT and PET imaging [63, 64]. In hybrid PET/CT and SPECT/CT images, the attenuation maps reflecting the distribution of linear attenuation coefficients are readily provided by CT images.

The main challenge for ASC arises in SPECT-only, PET-only, as well as PET/MR and SPECT/MR imaging since MR images are not directly correlated to electron density, and as such, do not provide information about attenuation coefficients of biological tissues [65, 66]. For SPECT-only and PET-only systems, emission-based algorithms have been developed to address this issue [67]. The main advantage is the capability to account for metallic implants and truncation artefacts [65, 66]. Including TOF information and anatomical prior improved the quantitative accuracy of emission-based algorithms [68-70]. However, application of this methodology across different radiotracers warrants further investigation.

In addition to emission-based algorithms, MR image-based algorithms, including segmentation and atlas-based algorithms have been developed to estimate attenuation coefficients from concurrent MR images [66]. In segmentation-based algorithms, different MR sequences, including T1, T2, ultra-short echo (UTE), and zero-time echo (ZTE) have been used to delineate major tissue classes followed by assignment of pre-defined linear attenuation coefficients to each tissue class. In Atlas-based algorithms [71, 72], pairs of co-registered MR and CT images (considered as template or atlas) are aligned to the target MR image to generate a continuous attenuation map. The main disadvantage of atlas-based algorithms is the high dependence on the atlas dataset and sub-optimal performance for subjects presenting with anatomical abnormalities [73, 74].

Deep learning-based algorithms were proposed to address the challenges of conventional ASC approaches in PET and SPECT imaging [2, 6]. Liu *et al.* [75] proposed converting non-attenuation corrected (NAC) brain PET images to synthetic CT (sCT) images. A GAN model was trained using 100 patients (in 2D mode) and tested on 28 patients achieving a relative error of less than 1% within 21 brain regions. Dong *et al.* [76] applied a similar approach in whole-body PET imaging using Cycle-GAN [76] reporting a mean PET quantification bias of 0.12% ±2.98%. Shi *et al.* [77] proposed a novel approach to generate sCT images in ^{99m}Tc-tetrofosmin myocardial perfusion SPECT imaging taking advantage of two images produced using different energy windows providing different representations of scattered and primary photon distributions. A multi-channel conditional GAN model was trained using SPECT images reconstructed using different energy windows as input to predict the corresponding sCT image. This model exhibited a normalized mean absolute error (NMAE) of 0.26± 0.15%.

Hwang *et al.* [78] used emission-based generated activity distributions and μ-maps as input to generate high-quality sCT images for ^{18}F-FDG brain PET studies. They reported less than 10% errors for CT values using CED and U-Net models. The same group applied the same approach in whole-body PET imaging using U-Net, achieving a relative error of 2.22 ±1.77% across 20 subjects [79]. Arabi and Zaidi [80] proposed the estimation of attenuation correction factors from the different TOF sinogram bins using ResNet, reporting an absolute SUV bias of less than 7% in different regions of the brain.

In addition to generating sCTs using PET emission data, direct generation of attenuation and scatter corrected images from NAC images was reported. Shiri *et al.* [81] and Yang *et al.* [82] trained a 2D U-Net network using brain ^{18}F-FDG PET studies reporting PET quantification bias of less than 5% in different regions of the brain. Arabi *et al.* [83] applied this approach to different brain molecular imaging probes, including ^{18}F-FDG, ^{18}F-DOPA, ^{18}F-Flortaucipir, and ^{18}F-Flutemetamol and reported SUV bias of less than 9% in different brain regions (figure 4). Shiri *et al.* [84] trained 2D, 3D, and patch-based ResNets on 1000 whole-body ^{18}F-FDG images and tested the proposed models on unseen 150 subjects. They performed ROI-based and voxel-based assessments and reported a relative error of less than 5%. Dong *et al.* [56] trained a 3D patch-based Cycle-GAN for whole-body ^{18}F-FDG images and reported a mean relative error of less than 5% calculated on malignant lesions. Emission-based ASC approaches using deep learning algorithms are summarized in Table 1.

The generation of sCT from MR images using deep learning-based regression approaches were reported in a number of studies. Li *et al.* used a 2D CED model to generate a 3-class probability map from T1-weighted images for ^{18}F-FDG brain images and reported an average bias of less than 1% in different brain regions [85]. Arabi *et al.* reported on the development of a novel adversarial semantic structure GAN model using T1-weighted MR

images to generate synthetic CT images for brain PET studies [86]. They reported a relative error of less than 4% in 64 anatomical brain regions. Leynes *et al.* used ZTE and Dixon MR sequences in a multi-channel input framework to train a U-Net model [87]. The network was trained on 10 subjects using a patch extraction strategy and tested on 16 subjects consisting of the external validation set, reporting a quantification bias less than 5% in different ROIs defined on bones and soft tissues of ^{18}F-FDG and ^{68}Ga-PSMA-11 PET images. Ladefoged *et al.* evaluated 3D U-Net architectures with UTE MR sequence as input and reported a mean relative error of -0.1% in brain tumours [88]. The main contributions of deep learning-assisted MRI-guided attenuation and scatter correction in emission tomography are summarized in Table 2.

Most deep learning-based ASC studies focused on brain imaging, which is less challenging compared to whole-body imaging where the anatomical structures are more complex with juxtapositions of various tissues having diverse attenuation properties and irregular shapes. There is obviously a need to evaluate these algorithms in more challenging heterogeneous regions, such as the chest and abdomen [84]. Moreover, the majority of these studies were performed using only one radiotracer (mostly ^{18}F-FDG) which raises questions regarding the generalizability of the models and the need for retraining and reevaluation on other tracers [83]. The size of training and evaluation sets is another limitation of deep learning-based ASC as the performance of these algorithms depends on the training sample. To the best of our knowledge, only two studies, one focusing on brain imaging [89] and the other on whole-body imaging [84], which used a large number of training sets. Most deep learning-based ASC studies were performed in PET imaging with a limited number of works reported for SPECT imaging [77, 90].

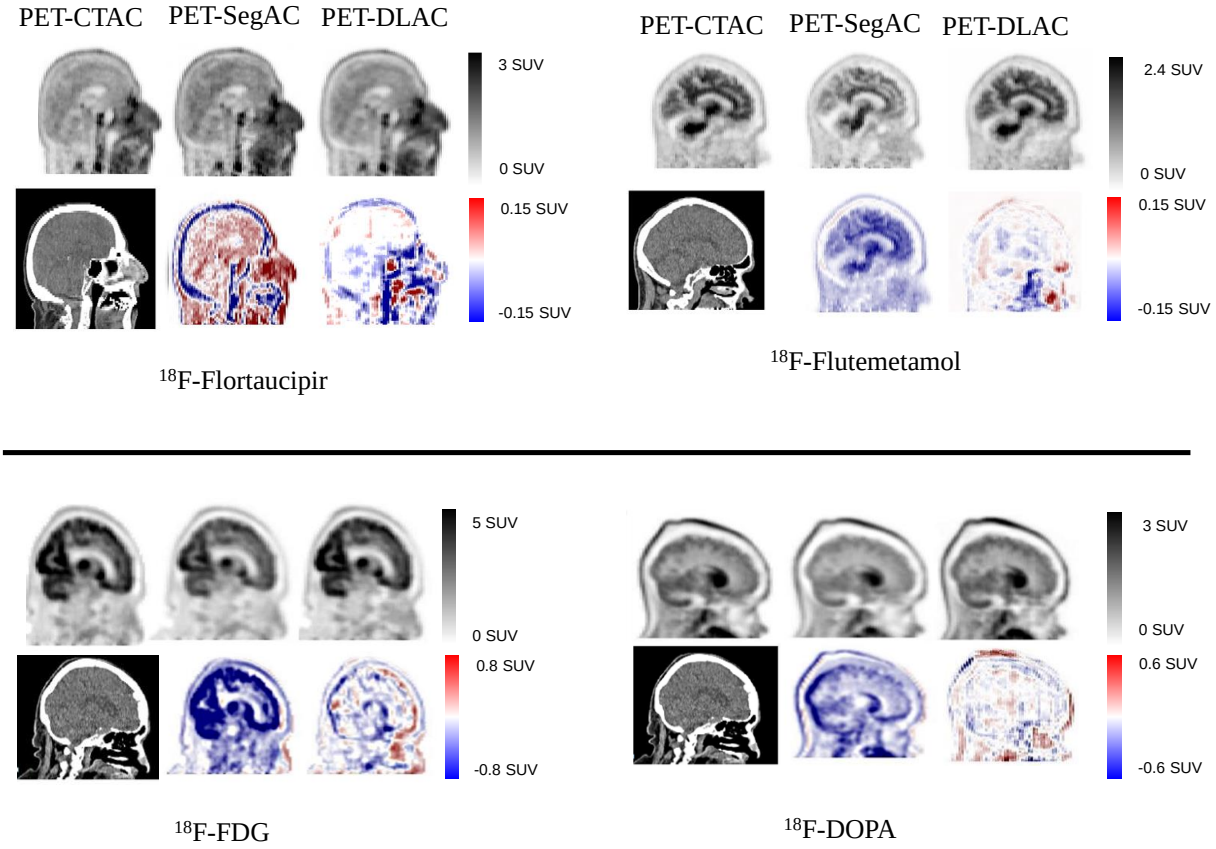


Figure 4. Comparison of PET images corrected for attenuation using CT-based, segmentation-based (containing background air and soft-tissue) (SegAC), and deep learning-guided (DLAC) approaches together with the reference CT image for ^{18}F-FDG, ^{18}F-DOPA, ^{18}F-Flortaucipir, and ^{18}F-Flutemetamol radiotracers. Difference SUV error maps are also presented for segmentation- and deep learning-based approaches.

Table 1.Summary of studies performed for emission-based ASC using deep learning algorithms.

Authors	Modality	Radiotracer	Approach	Algorithm	Body region	Training	Training/Test	Input	Output	Evaluation	Outcome	Loss Function
Liu et al. [75]	PET	18F-FDG	NAC to sCT	CED	Brain	2D (200×180)	100/28	NAC	sCT	21 VOIs + whole brain	Average PET quantification bias − 0.64 ± 1.99	L2
Armanious et al. [91]	PET	18F-FDG	NAC to sCT	GAN	Brain	2D (400×400)	50/40	NAC	sCT	7 VOIs + Whole brin	< 5% Average PET quantification bias	Perceptual
Dong et al. [76]	PET	18F-FDG	NAC to sCT	Cycle-GAN	Whole-body	Patches (64×64×16)	80/39	NAC	sCT	7 VOIs in different regions	0.12% ±2.98% Mean PET quantification bias	Adversarial loss + cycle consistency loss
Colmeiro et al. [92]	PET	18F-FDG	NAC to sCT	GAN	Whole-body	3D (128×128×32)	108/10	NAC	sCT	---	SUV not reported MAE 88.9 ± 10.5 (HU)	
Shi et al. [77]	SPECT	99mTc-tetrofosmin	NAC to sCT	GAN Conditional	Cardiac	3D (16×16 ×16)	40/25	Photo peak (126–155 keV) and (114–126 keV)	sCT	Voxelwise	NMAE 0.26%± 0.15%	L2+LGDL
Arabi et al. [80]	PET	18F-FDG	NAC to ACF	ResNet	Brain	2D (168×200) 7 input channels and 1 output channel	68/4 CV	TOF sinogram bins	attenuation correction factors (ACFs)	63 brain regions	< 7% absolute PET quantification bias	L2norm
Hwang et al. [78]	PET	18F-FP-CIT	MLAA to sCT	CAE and U-Net	Brain	2D (200×200)	40/5 CV	MLAA-generated activity distribution and μ-map	sCT	4 VOIs of brain	PET quantification bias ranging from −8% to −4%	L2-norm
Hwang et al. [79]	PET	18F-FDG	MLAA to sCT	U-Net	Whole-body	Patches (64×64×16)	80/20	MLAA-generated activity distribution and μ-map	sCT	bone lesions + soft-tissues	PET quantification bias bias% Bone lesions: 2.22 ±1.77% Soft-tissue lesions: 1.31% ± 3.35%)	L1 norm
Shi et al. [93]	PET	18F-FDG	MLAA to sCT	U-Net	Whole-body	Patches (32×32×32)	80/20	MLAA-generated activity distribution and μ-map	sCT	Region-wise	NMAE 3.6%	Line-integral projection loss

Shiri et al. [81]	PET	18F-FDG	NAC to MAC	U-Net	Brain	2D (256×256)	111/18	NAC	AC	83 VOIs	PET quantification bias − 0.10 ± 2.14	MSE
Yang et al. [82]	PET	18F-FDG	NAC to MAC	U-Net	Brain	2D (256×256)	25/10	NAC	AC	116 VOIs	PET quantification bias 4.0% ± 15.4%	Mean squared error (or L2 loss)
Arabi et al. [83]	PET	18F-FDG 18F-DOPA 18F-Flortaucipir 18F-Flutemetamol	NAC to MAC	ResNet	Brain	2D (128×128)	180	NAC	AC	7 brain regions	< 9% Absolute PET quantification bias	L2norm
Dong et al. [56]	PET	18F-FDG	NAC to MAC	Cycle-GAN	Whole-body	Patches (64×64×64)	25 leave-one-out + 10 patients × 3 sequential scan tests	NAC	AC	6 VOIs in lesions	ME 2.85 ± 5.21	Wasserstein loss
Shiri et al. [84]	PET	18F-FDG	NAC to MAC	ResNet	Whole-body	2D (154×154) Patch (64×64×64) 3D (154×154×32)	1000/150	NAC	AC	Voxelwise and region-wise	RE % < 5 %	L2norm
Xiang et al. [90]	SPECT	90Y	Input: μ-map + SPECT projections, Output: scatter projections	DCNN (VGG and ResNet)	Chest + Abdomen	2D (128×80)	Phantom + 6 patients	Projected attenuation map SPECT projection	Estimated scatter projections	Voxelwise	NRMSE 0.41	MSE

CED: Convolutional Encoder Decoder, GAN: Generative Adversarial Network, NAC: Non-Attenuation Corrected, sCT: Pseudo CT, VOI: Volume of Interest, HU: Hounsfield Unit, MAE: Mean Absolute Error, ACF: Attenuation correction factor, CV: Cross-Validation, TOF: Time of Flight, ME: Mean Error, RE: Relative Error, NRMSE: Normalized Root Mean Square Error

Table 2. Summary of studies performed on MRI-guided synthetic CT generation using deep learning approaches.

Authors	Modality	Radiotracer	Approaches	Algorithm	Organ	Training	Training/Test	Input	Output	Evaluation	Error	Loss Function
Bradshaw *et al.* [94]	PET	^{18}F-FDG	MRI to tissue labeling	DeepMedic	Pelvis	Patch (25×25×25)	12/6	T1/T2	4-class probability map	16 soft-tissue lesions	MSE 4.9%	Cross-entropy loss
Jang *et al.* [95]	PET	^{18}F-FDG	MRI to tissue labeling	SegNet	Brain	2D (340 ×340)	Pretraining: 30 MRI, Training 6 MRI Evaluation: 8 MRI	UTE	sCT	23 VOIs+ whole brain	< 1%	Multi-class soft-max classifier
Liu *et al.* [85]	PET	^{18}F-FDG	MRI to tissue labeling	CED	Brain	2D (340×340)	30/10 MRI to label 5 PET/MRI	T1-weighted	3-class probability map	23 VOIs + whole brain	Average error <1% in the whole brain	Cross-entropy
Arabi *et al.* [86]	PET	^{18}F-FDG	MRI to tissue labeling	GAN	Brain	3D (224 × 224 × 32)	40 /2 CV	T1	3-class probability map	63 brain regions	less than 4%	Cross-entropy
Mecheter *et al.* [96]	PET	^{18}F-FDG	MRI to Segment	SegNet	Brain	2D (256×256)	12/3	T1/T2	3 Tissue	–	–	Cross-entropy
Leynes *et al.* [87]	PET	^{18}F-FDG ^{68}Ga-PSMA-11	MRI to sCT	U-Net	Pelvis	Patch (32×32×16)	10/16	ZTE and Dixon (fat/water) multi-input	sCT	30 bone lesions and 60 soft-tissue lesions	RMSE 2.68% in bone and 4.07% in soft-tissues	L1-loss, gradient difference loss (GDL), and Laplacian difference loss (LDL)
Gong *et al.* [97]	PET	^{18}F-FDG	MRI to sCT	U-Net	Brain	2D (144×144)	40 /5 CV	Dixon and ZTE	sCT	8 VOIs + whole brain	MRE 3%	L1 norm
Ladefoged *et al.* [88]	PET	^{18}F-FET	MRI to sCT	U-Net	Brain	3D (192 ×192×16)	79/4 CV	UTE	sCT	36 brain tumor VOIs	Mean relative difference -0.1%	Mean squared-error
Blanc-Durand *et al.* [98]	PET	^{18}F-FDG	MRI to sCT	U-Net	Brain	Patch (64×64×16)	23/47	ZTE	sCT	70 VOIs + whole brain	Average error -0.2%	Mean squared error
Spuhler *et al.* [99]	PET	^{11}C-WAY-100635 ^{11}C-DASB	MRI to sCT	U-Net	Brain	2D (256 × 256)	56/11	T1	sCT	20 brain regions (VOIs)	PET quanitifaction error within VOIs −0.49±1.7% 11C-WAY-100635 −1.52±0.73% 11C-DASB	L1 error

Torrado-Carvajal *et al.* [100]	PET	^{18}F-FDG ^{18}F-choline	MRI to sCT	U-Net	Pelvis	2D (256 × 256)	28/4 CV	Dixon-VIBE	sCT	Regionwise and voxelwise	< 1%	Mean absolute error
Gong *et al.* [101]	PET	^{11}C-PiB ^{18}F-MK6240	MRI to sCT	U-Net	Brain	2D (160×160) Multichannel input of 5 and 35	35/5 CV	1 UTE image and 6 multi-echo Dixon with different TEs	sCT	8 VOIs	< 2%	L1-norm
Gong *et al.* [102]	PET	^{18}F-FDG	MRI to sCT	Cycle-GAN	Brain	Patch (144×144×25)	32 /4 CV	Dixon	sCT	16 VOIs	< 3%	L1-norm loss
Ladefoged *et al.* [89]	PET	^{18}F-FDG	MRI to sCT	U-Net	Brain	3D (192× 192 ×16) Multichannel	732/305	Dixon VIBE T1 UTE	sCT	16 VOIs	< 1%	Mean squared error
Leynes *et al.* [103]	PET	^{18}F-FDG ^{68}Ga-PSMA-11 ^{68}Ga-DOTATATE	MRI to sCT	Bayesian DCNN U-Net	Pelvis	Patch (64×64×32)	10/19	Dixon ZTE	sCT	ROIs on lesion	< **5**%	L1-loss+gradient difference loss (GDL)+ Laplacian difference loss
Pozaruk *et al.* [104]	PET	^{68}Ga-PSMA-11	MRI to sCT	GAN, U-Net	Pelvis	2D (192×128)	18/10	Dixon	sCT	ROIs on the prostate	< 3%	mean absolute error
Tao *et al.* [105]	PET	Not reported	MRI to sCT	Conditional GAN	Brain	2D (256×256)	9/2	ZTE	sCT	Voxel wise	<5% CT HU bias	L1 loss and GAN loss

CV: Cross-Validation, ROI: Region of Interest, VOIs: Volume of Interest, HU: Hounsfield Unit.

Image interpretation and decision support

Image segmentation, registration, and fusion

Computer-aided tools for the analysis and processing of medical images have been developed to improve the reliability and robustness of the extracted features. Advanced machine-learning techniques are being developed to learn 1) effective similarity features, 2) a common feature representation, or 3) appearance mapping, in order to provide a model that can match large appearance variations [106, 107].

Accurate organ/tumor delineation from molecular images is mainly used in the context of oncological PET imaging studies for quantitative analysis targeting various aspects, including severity scoring, radiation treatment planning, volumetric quantification, radiomic features extraction, etc. However, this is challenging owing to the poor spatial resolution and high statistical noise of molecular images. In current clinical practice, image segmentation is typically performed manually, which tends to be labor-intensive and prone to intra- and inter-observer variability. A number of recent studies explored the potential of DL-based automated tumor segmentation from PET or hybrid PET/CT examinations [108, 109]. Zhao *et al.* used a U-Net architecture for tumor delineation from ^{18}F-FDG PET/CT images within the lung and nasopharyngeal regions [110, 111]. Blanc-Durant *et al.* demonstrated the feasibility of ^{18}F-fluoro-ethyl-tyrosine (^{18}F-FET) PET lesion segmentation using a CNN model [112]. Leung *et al.* developed a modular deep-learning framework for primary lung tumor segmentation from FDG-PET images with a small-size clinical training dataset, generalized across different scanners, achieving a Dice index of 0.73. They addressed the limitations of the small size of the training dataset as well as the accuracy and variability of manual segmentations used as ground truth by using a realistic simulation dataset [113]. Wang *et al.* proposed a deep learning-assisted method for automated segmentation of the left ventricular region using gated myocardial perfusion SPECT [114].

Roccia *et al.* used a DL algorithm to predict the arterial input function for quantification of the regional cerebral metabolic rate from dynamic ^{18}F-FDG PET scans [115]. Park *et al.* developed an automated pipeline for glomerular filtration rate (GFR) quantification of ^{99m}Tc-DTPA from SPECT/CT scans using a 3D U-Net model through kidney segmentation [116].

AI-assisted diagnosis and prognosis

AI algorithms have been employed to build models exploiting the information extracted from medical images to perform a specific clinical task, e.g. object detection/classification, severity scoring, clinical outcome prediction, treatment planning, and monitoring response to therapy [117]. Numerous works reported on automated detection and classification of various pathologies (e.g. malignant vs. benign) in nuclear medicine [118]. For benign diseases, cardiovascular SPECT and brain PET imaging were the main focus of AI applications [119]. Xu *et al.* developed an automated pipeline using two cascaded V-NETs for lesion prediction and segmentation to detect multiple myeloma bone lesions from ^{68}Ga-Pentixafor PET/CT [120]. Togo *et al.* demonstrated the feasibility of cardiac sarcoidosis detection from ^{18}F-FDG PET scans using Inception-v3 network (83.9% sensitivity and 87% specificity), which outperformed conventional SUV_{max}- (46.8% sensitivity and 71.0% specificity) and coefficient of variance (CoV)-based (65.5% sensitivity and 75.0% specificity) approaches [121]. Ma *et al.* modified a DenseNet architecture for the diagnosis of thyroid disease using SPECT images into three categories: Graves' disease, Hashimoto, and subacute thyroiditis [122].

^{18}F-FDG PET is extensively used as a diagnostic tool in neurodegenerative disorders, especially Alzheimer Disease (AD) to improve diagnosis and monitor disease progression. The role of AI in AD diagnosis has been recently reviewed by Duffy *et al.* [123]. Lu *et al.* developed an AI-based framework for the early diagnosis of AD using multimodal ^{18}F-FDG PET/MR and multiscale deep neural network (82.4% accuracy and 94.23% sensitivity) [124]. Choi and Jin proposed a straightforward deep learning algorithm based on only ^{18}F-FDG PET images for early detection of AD (84.2% accuracy) that outperformed conventional feature-based quantification approaches, e.g. Support-Vector-Machine (76.0% accuracy) and VOI-based (75.4% accuracy) techniques [125]. Machine learning algorithms have shown promising results in the classification of AD using brain PET images. Liu *et al.* proposed a classification algorithm of FDG PET images composed of 2D CNNs and recurrent neural networks (RNNs) [126]. The CNN model was trained to extract the features in 2D, while the RNN extracted the features in 3D mode (95.3% accuracy for AD vs controls and 83.9% for mild impairment vs controls). In a follow-up work,

they proposed a cascaded CNN model to train the multi-level features of multimodal PET/MRI images. First, a patch-based 3D CNN was constructed, and then, a high-level 2D CNN followed by a softmax layer was trained to collect the high-level features. Finally, all features were concatenated followed by a softmax layer for AD classification [127]. The flexibility of AI algorithms enables learning the characteristics from heterogeneous data that have meaningful correlations but not obvious for the human interpreter. Zhou *et al.* developed a deep learning model for AD diagnosis using genetic input data, e.g. single nucleotide polymorphism in addition to radiological brain images that outperformed classification performance relative to other state-of-the-art methods [128].

Radiomics and precision medicine

Radiomics refers to a quantitative set of features, e.g. intensity, texture, and geometrical characteristics obtained from radiological images to discriminate quantifiable phenotypes that cannot be extracted through qualitative assessment of images. A radiomics model is commonly built through 4 steps: i) image acquisition/reconstruction; ii) VOI segmentation; iii) quantification/hand-crafted feature extraction; iv) statistical analysis [129]. While data-driven deep learning approaches are different from feature-driven approaches, deep learning has the ability to directly learn discriminative features from data in their natural raw form without the necessity to define VOIs or extract engineered features [130].

SPECT and PET images represent biological and physiopathological characteristics that can be quantitatively expressed using radiomics. Most studies focused on ^{18}F-FDG PET images for prognosis (staging) or outcome prediction using handcrafted radiomics [131-133]. Delta radiomics, as a metric for treatment outcome, has been developed based on multiple time-point images [134]. Some studies investigated the advantage of using hybrid images, e.g. PET/CT and PET/MR [135], extending the feature extraction to non-primary tumor volumes, such as bone marrow and metastatic lymph nodes [136], and deriving features from parametric PET images [137]. Application of radiomics in SPECT has also been recently investigated by Ashrafnia *et al.* for prediction of coronary artery calcification in [^{99m}Tc]-sestamibi SPECT myocardial perfusion scans [138]. Rahmim *et al.* evaluated the extraction of radiomic features from longitudinal Dopamine transporter (DAT) SPECT images for outcome prediction in Parkinson's disease [139]. DL-based radiomics was compared with feature-driven methods to highlight the advantages of CNNs compared to handcrafted radiomics for response prediction of chemotherapy in oesophageal cancer [140]. Wang *et al.* reported that CNNs did not outperform traditional radiomics in the classification of mediastinal lymph nodes of non-small lung cancer. Yet, it was preferred, since it was more user-friendly and required less data handling, and was less prone to feature selection bias [118].

Internal radiation dosimetry

AI has significantly impacted other fields of nuclear medicine through developing methods for radiation dose monitoring, dose reduction strategies, building theranostic decision trees, and dose limit compliance. In the era of precision medicine, personalized dosimetry is increasingly used in nuclear medicine. Targeted Radionuclide Therapy (TRT) has been recently merged with the concept of theranostics, a promising technique in radiation oncology. Despite the growing interest in dosimetry-guided patient-specific TRT, the one-fits-all approach is still used in routine clinical practice. In the context of individualized dose profiling, the construction of patient-specific computational models is the first step toward this goal [141]. Numerous works focused on the development of pipelines for the construction of patient-specific computational models applicable in personalized dosimetry in either therapy or diagnostic procedures [142-144]. Fu *et al.* developed a framework for automated generation of computational phantoms from CT images [145]. They used cascaded modules consisting of i) registration of patient CT images to an anchor phantom, ii) segmentation of organs using UNet structure, and iii) registration of segmented organs inside the deformed anchor phantom to generate an individualized computational model that is applicable for personalized dosimetry in both diagnostic and therapeutic procedures. Besides, the automatic segmentation of organs at risk for various application sites of TRT has been extensively studied. Jackson *et al.* developed a framework for automated monitoring of absorbed dosed in the kidneys of patients undergoing ^{177}Lu-PSMA therapy [146]. They used a 3D CNN architecture for kidney segmentation to provide organ-level dosimetry from post-treatment SPECT imaging to estimate renal radiation doses from TRT. Tang *et al.* proposed a CNN-based algorithm for liver segmentation for personalized selective internal radiation therapy [147]. Kidney segmentation has been conducted using a 3D UNet architecture on ^{177}Lu SPECT images for uptake quantification and dosimetry [148].

MC simulations using patient-specific anatomical and metabolic features constitute the current gold standard for internal dosimetry calculations. However, the approach suffers from exhaustive computational burden. Recently deep learning approaches have been employed in patient-specific dosimetry for monitoring or treatment plan optimization using molecular images (SPECT and PET). Akhavanallaf *et al.* developed an AI-based framework based on ResNet architecture for personalized dosimetry in nuclear medicine procedures [149]. They extended the key idea behind the voxel-based MIRD (Medical Internal Radiation Dose) approach through the prediction of specific S-values according to the density map derived from CT images followed by calculation of the cumulated activity map from the predicted specific kernels (Figure 5). A physics-informed deep neural network (DNN) was designed to predict the energy deposited in the volume surrounding a unit radioactive source in the center of the kernel. The input channel was fed with a density map whereas the output was MC-based deposited energy maps of the given radiotracer, referred to as specific S-value kernels. Lee *et al.* proposed a methodology employing deep learning for the direct generation of dose rate maps from ^{18}F-FDG PET/CT images [150]. Gotz *et al.* used a modified U-Net network for dose map reconstruction of patients receiving ^{177}Lu-PSMA [151]. They further extended their work for patient-specific dosimetry of ^{177}Lu compounds by predicting specific dose voxel kernels using AI algorithms [152]. Xue *et al.* developed a GAN model to predict post-therapy dosimetry for ^{177}Lu-PSMA therapy using pre-therapy ^{68}Ga-PSMA PET/CT examinations [153].

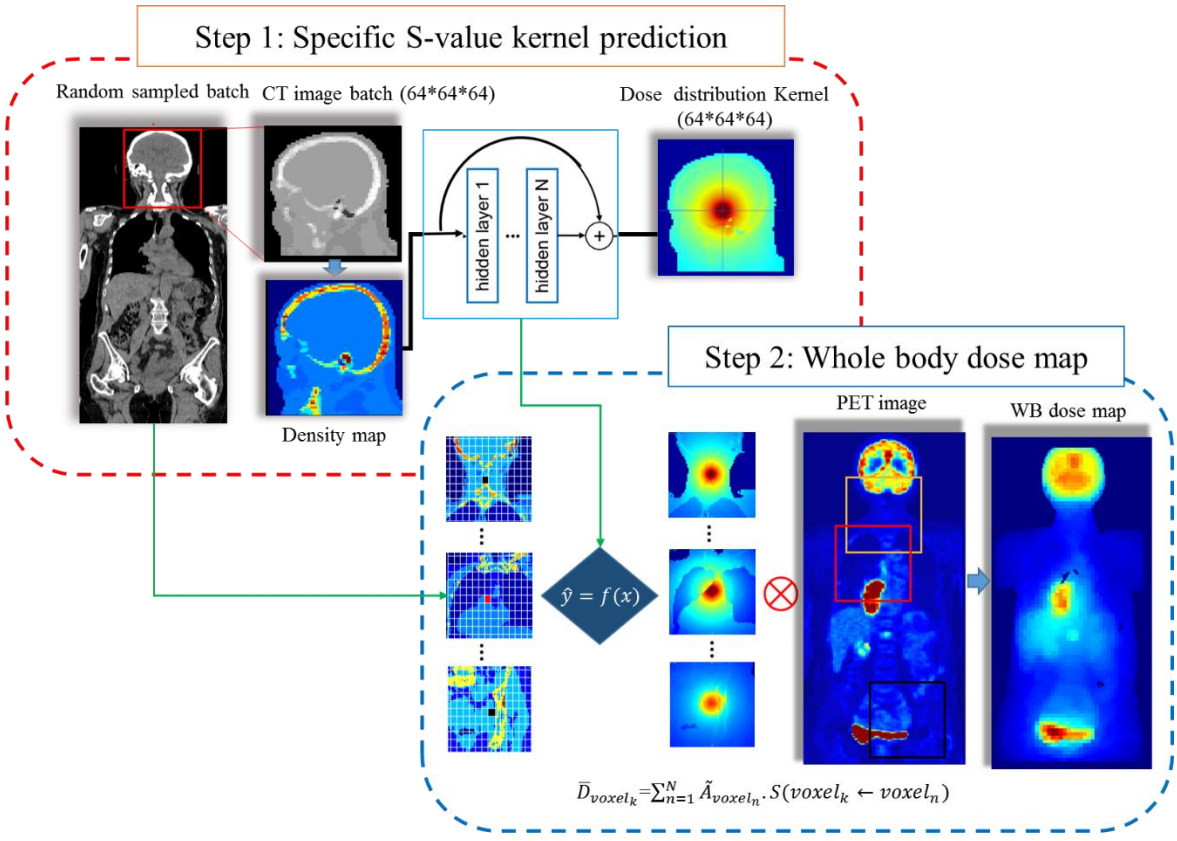


Figure 5. Schematic representation of the voxel-scale dosimetry procedure. The top and bottom panels show the deep learning-based specific S-value kernel prediction and MIRD-based voxel dosimetry formalism. Adapted from Ref. [149].

Despite the substantial growth and widespread adoption of patient-specific TRT, the "one-size-fits-all" approach is still commonly used in the clinic. Del Prete *et al.* reported that in TRT, organs at risk rarely reach the conservative threshold dose while most tumors receive submaximal doses, thus leading to undertreatment of patients [154]. Therefore, retrospective studies involving patients receiving TRT allows the evaluation of the treatment response to the one-dose-fits-all approach and would demonstrate the critical nature of the transition to adaptive dosimetry-guided treatment planning. This technique requires a tool incorporating a module for automatic segmentation of tumors/organs at risk along with a fast and accurate personalized dosimetry module.

IV. Challenges/opportunities and outlook

Over the past decade, there have been significant advances in deep learning-assisted developments which have impacted modern healthcare. The potential of AI-based solutions in various molecular imaging applications has been thoroughly explored in academic and corporate settings during the last decade. This article may, therefore, be viewed as an early album covering some of the many and varied snapshots of this rapidly growing field. At this time, these tools are still available only to experts in the field but there are many reasons to believe that it will be potentially available for routine use in the near future.

The proposed AI-based solutions in PET and SPECT imaging can be divided into two groups: (i) Techniques solely proposed to replace the current algorithms/frameworks due to their superior performance and (ii) approaches that have rendered previously impractical/unfeasible scenarios/frameworks using conventional methods feasible. In the first category, the promise of deep learning approaches consists in providing even slightly better functionality/performance compared to existing methods rather than undertaking an unprecedented functionality previously inconceivable. For example in PET instrumentation, Anger logic is used to determine the location of the interaction within the detector modules. Novel approaches based on deep learning methods tend to solely replace the Anger logic to achieve better localization and energy resolution. In this regard, novel deep learning approaches play the same role and compete with existing methods.

Likewise, in MRI-guided synthetic CT generation, deep learning approaches serve as alternative to atlas- or MRI segmentation-based techniques, whereas in the domain of noise reduction, current analytical models/algorithms are being replaced by deep learning methods. In this regard, the proposed deep learning methods would not revolutionarily alter the current frameworks or produce a paradigm shift, though they hold the promise of providing more accurate outcomes or requiring less human intervention, and easy adaptability to new input data. In this light, this category of AI-based solutions are more likely to be fully employed in clinical practice or on commercial systems since less standardization, protocols and frame redefinition, and staff retraining is required. For instance, deep learning-guided CT image reconstruction developed by GE Medical Systems obtained FDA approval [30].

Conversely, the extraordinary power of deep learning approaches has rendered many previously impractical/nonfeasible scenarios/frameworks feasible. This includes tasks, such as attenuation and scatter correction in the image domain, estimation of synthetic CT images from the non-attenuation corrected emission images, object completion of truncation date, image translation, and internal dosimetry. These processes are inherently ill-posed and in many cases, there is a lack of a mathematical framework associated with these problems. Such AI-based solutions, though offering unprecedented opportunities in PET and SPECT imaging, face thoughtful challenges with respect to their deployment in clinical practice as they require extensive validation using large clinical databases and a wide range of conditions.

Overall, a clear distinction should be made between the applications of AI-based solutions as processing or decision support tools or the replacement of experts or clinicians in clinical practice. Considering the superior performance of deep learning approaches, some algorithms are sufficiently mature and robust to be deployed in clinical practice as decision support tools. These algorithms are supposed to replace conventional methods owing to their superior performance or robustness. In this regard, any possible failure of the AI-based solution would be treated in a similar way to existing approaches. Conversely, AI-based solutions deemed to fully replace the experts are still considered as fantasy or science-fiction. Such algorithms still require additional development and remarkable evolution to be independently employed in clinical setting. Nevertheless, these algorithms could play a significant role in the short run as decision support tools to create a synergy between the capabilities of AI and human expertise.

It is gratifying to see in overview the progress that AI has made, from early developments in neural networks to complex deep learning architectures, and more recently towards continuous learning AI in radiology [47]. Challenges remain, particularly in the areas of clinical validation and liability towards wider adoption, ethical and legal aspects and a number of other issues that need to be settled [155].

Acknowledgments

This work was supported by the Swiss National Science Foundation under grant SNRF 320030_176052 and the Private Foundation of Geneva University Hospitals under grant RC-06-01.

References

[1] F. Nensa, A. Demircioglu, and C. Rischpler, "Artificial intelligence in nuclear medicine,"*Journal of Nuclear Medicine,* vol. 60, no. Supplement 2, pp. 29S-37S, 2019.

[2] H. Arabi and H. Zaidi, "Applications of artificial intelligence and deep learning in molecular imaging and radiotherapy,"*European Journal of Hybrid Imaging,* vol. 4, no. 1, pp. 1-23, 2020.

[3] K. Gong, E. Berg, S. R. Cherry, and J. Qi, "Machine learning in PET: From photon detection to quantitative image reconstruction.,"*Proceedings of the IEEE,* vol. 108, no. 1, pp. 51-68, 2020.

[4] E. Alpaydin, *Introduction to machine learning*. MIT press, 2020.

[5] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning,"*nature,* vol. 521, no. 7553, pp. 436-444, 2015.

[6] T. Wang *et al.*, "Machine learning in quantitative PET: A review of attenuation correction and low-count image reconstruction methods,"*Phys Med,* vol. 76, pp. 294-306, 2020.

[7] G. Lee and H. Fujita, *Deep learning in medical image analysis: challenges and applications*. Springer, 2020.

[8] J. Masci, U. Meier, D. Cireşan, and J. Schmidhuber, "Stacked convolutional auto-encoders for hierarchical feature extraction," in *International conference on artificial neural networks*, 2011: Springer, pp. 52-59.

[9] D. Shen, G. Wu, and H.-I. Suk, "Deep learning in medical image analysis,"*Annual review of biomedical engineering,* vol. 19, pp. 221-248, 2017.

[10] F. Altaf, S. M. Islam, N. Akhtar, and N. K. Janjua, "Going deep in medical image analysis: Concepts, methods, challenges, and future directions,"*IEEE Access,* vol. 7, pp. 99540-99572, 2019.

[11] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, 2015: Springer, pp. 234-241.

[12] O. Oktay *et al.*, "Attention u-net: Learning where to look for the pancreas,"*arXiv preprint arXiv:1804.03999,* 2018.

[13] F. I. Diakogiannis, F. Waldner, P. Caccetta, and C. Wu, "Resunet-a: a deep learning framework for semantic segmentation of remotely sensed data,"*ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 162, pp. 94-114, 2020.

[14] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1125-1134.

[15] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223-2232.

[16] F. Müller, D. Schug, P. Hallen, J. Grahe, and V. Schulz, "A novel DOI positioning algorithm for monolithic scintillator crystals in PET based on gradient tree boosting,"*IEEE Transactions on Radiation and Plasma Medical Sciences,* vol. 3, no. 4, pp. 465-474, 2018.

[17] P. Peng, M. S. Judenhofer, A. Q. Jones, and S. R. Cherry, "Compton PET: A simulation study for a PET module with novel geometry and machine learning for position decoding,"*Biomedical Physics & Engineering Express,* vol. 5, no. 1, p. 015018, 2018.

[18] A. Sanaat and H. Zaidi, "Depth of interaction estimation in a preclinical PET scanner equipped with monolithic crystals coupled to SiPMs using a deep neural network.,"*Appl Sci,* vol. 10, no. 14, p. 4753, 2020.

[19] F. Müller, D. Schug, P. Hallen, J. Grahe, and V. Schulz, "Gradient tree boosting-based positioning method for monolithic scintillator crystals in positron emission tomography,"*IEEE Transactions on Radiation and Plasma Medical Sciences,* vol. 2, no. 5, pp. 411-421, 2018.

[20] F. Hashimoto, K. Ote, R. Ota, and T. Hasegawa, "A feasibility study on 3D interaction position estimation using deep neural network in Cherenkov-based detector: a Monte Carlo simulation study,"*Biomedical Physics & Engineering Express,* vol. 5, no. 3, p. 035001, 2019.

[21] E. Berg and S. R. Cherry, "Using convolutional neural networks to estimate time-of-flight from PET detector waveforms,"*Physics in Medicine & Biology,* vol. 63, no. 2, p. 02LT01, 2018.

[22] R. Gladen, V. Chirayath, A. Fairchild, M. Manry, A. Koymen, and A. Weiss, "Efficient Machine Learning Approach for Optimizing the Timing Resolution of a High Purity Germanium Detector,"*arXiv preprint arXiv:2004.00008,* 2020.

[23] A. Shawahna, S. M. Sait, and A. El-Maleh, "FPGA-based accelerators of deep learning networks for learning and classification: A review,"*IEEE Access,* vol. 7, pp. 7823-7859, 2018.

[24] I. Shiri *et al.*, "Ultra-low-dose chest CT imaging of COVID-19 patients using a deep residual neural network,"*European radiology,* pp. 1-12, 2020.

[25] I. Häggström, C. R. Schmidtlein, G. Campanella, and T. J. Fuchs, "DeepPET: A deep encoder–decoder network for directly solving the PET image reconstruction inverse problem,"*Medical image analysis,* vol. 54, pp. 253-262, 2019.
[26] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen, "Image reconstruction by domain-transform manifold learning.,"*Nature,* vol. 555, no. 7697, pp. 487-492, 2018.
[27] S. Ravishankar, J. C. Ye, and J. A. Fessler, "Image reconstruction: From sparsity to data-adaptive methods and machine learning,"*Proceedings of the IEEE,* vol. 108, no. 1, pp. 86-109, 2019.
[28] A. Reader, G. Corda, A. Mehranian, C. da Costa-Luis, S. Ellis, and J. A. Schnabel, "Deep learning for PET image reconstruction,"*IEEE Trans. Radiat. Plasma Med. Sci.*, 2020.
[29] FDA. "510k Premarket Notification of AiCE Deep Learning Reconstruction (Canon)." https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K183046 (accessed.
[30] FDA. "510k Premarket Notification of Deep Learning Image Reconstruction (GE Medical Systems)." https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K183202 (accessed.
[31] W. Whiteley, V. Panin, C. Zhou, J. Cabello, D. Bharkhada, and J. Gregor, "FastPET: Near Real-Time Reconstruction of PET Histo-Image Data Using a Neural Network,"*IEEE Transactions on Radiation and Plasma Medical Sciences,* 2020.
[32] H. Arabi and H. Zaidi, "Non-local mean denoising using multiple PET reconstructions,"*Annals of nuclear medicine,* 2020.
[33] H. Arabi and H. Zaidi, "Improvement of image quality in PET using post-reconstruction hybrid spatial-frequency domain filtering.,"*Phys Med Biol,* vol. 63, no. 21, p. 215010, 2018.
[34] H. Arabi and H. Zaidi, "Spatially guided nonlocal mean approach for denoising of PET images,"*Medical physics,* vol. 47, no. 4, pp. 1656-1669, 2020.
[35] C. Chan, R. Fulton, R. Barnett, D. D. Feng, and S. Meikle, "Postreconstruction nonlocal means filtering of whole-body PET with an anatomical prior,"*IEEE transactions on medical imaging,* vol. 33, no. 3, pp. 636-50, 2014.
[36] A. J. Reader and H. Zaidi, "Advances in PET image reconstruction.,"*PET Clinics,* vol. 2, no. 2, pp. 173-190, 2007.
[37] J. Yan, J. C.-S. Lim, and D. W. Townsend, "MRI-guided brain PET image filtering and partial volume correction,"*Physics in Medicine & Biology,* vol. 60, no. 3, p. 961, 2015.
[38] Y. Wang *et al.*, "Semisupervised tripled dictionary learning for standard-dose PET image prediction using low-dose PET and multimodal MRI,"*IEEE Transactions on Biomedical Engineering,* vol. 64, no. 3, pp. 569-579, 2016.
[39] L. An *et al.*, "Multi-Level Canonical Correlation Analysis for Standard-Dose PET Image Estimation,"*Ieee Transactions on Image Processing,* vol. 25, no. 7, pp. 3303-3315, 2016.
[40] W. Zhang *et al.*, "Image reconstruction for positron emission tomography based on patch-based regularization and dictionary learning,"*Medical physics,* vol. 46, no. 11, pp. 5014-5026, 2019.
[41] J. Bland *et al.*, "MR-guided kernel EM reconstruction for reduced dose PET imaging,"*IEEE transactions on radiation and plasma medical sciences,* vol. 2, no. 3, pp. 235-243, 2017.
[42] G. Litjens *et al.*, "A survey on deep learning in medical image analysis.,"*Med Image Anal,* vol. 42, pp. 60-88, 2017.
[43] K. T. Chen *et al.*, "Ultra-low-dose (18)F-Florbetaben amyloid PET imaging using deep learning with multi-contrast MRI inputs.,"*Radiology,* vol. 290, no. 3, pp. 649-656, 2019.
[44] L. Xiang, Y. Qiao, D. Nie, L. An, Q. Wang, and D. Shen, "Deep auto-context convolutional neural networks for standard-dose PET image estimation from low-dose PET/MRI.,"*Neurocomputing,* vol. 267, pp. 406-416, 2017.
[45] Y. Wang *et al.*, "3D conditional generative adversarial networks for high-quality PET image estimation at low dose.,"*Neuroimage,* vol. 174, pp. 550-562, 2018.
[46] J. Ouyang, K. T. Chen, E. Gong, J. Pauly, and G. Zaharchuk, "Ultra-low-dose PET reconstruction using generative adversarial network with feature matching and task-specific perceptual loss.,"*Med Phys,* vol. 46, no. 8, pp. 3555-3564, 2019.
[47] T.-A. Song, S. R. Chowdhury, F. Yang, and J. Dutta, "Super-resolution PET imaging using convolutional neural networks,"*IEEE Transactions on Computational Imaging,* vol. 6, pp. 518-528, 2020.
[48] A. Sanaat, H. Arabi, I. Mainta, V. Garibotto, and H. Zaidi, "Projection-space implementation of deep learning-guided low-dose brain PET imaging improves performance over implementation in image-space.,"*J Nucl Med,* vol. 61, no. 9, pp. 1388-1396, 2020.
[49] J. Xu, E. Gong, J. Pauly, and G. Zaharchuk, "200x Low-dose PET Reconstruction using Deep Learning,"*ARXIV,* p. eprint arXiv:1712.04119, 2017.
[50] C. C. Liu and J. Qi, "Higher SNR PET image prediction using a deep learning model and MRI image.,"*Phys Med Biol,* vol. 64, no. 11, p. 115004, 2019.

[51] J. Cui *et al.*, "PET image denoising using unsupervised deep learning.,"*Eur J Nucl Med Mol Imaging*, p. *in press*, 2019.
[52] W. Lu *et al.*, "An investigation of quantitative accuracy for deep learning based denoising in oncological PET,"*Physics in Medicine & Biology*, vol. 64, no. 16, p. 165019, 2019.
[53] S. Kaplan and Y.-M. Zhu, "Full-dose PET image estimation from low-dose PET image using deep learning: a pilot study,"*Journal of digital imaging*, vol. 32, no. 5, pp. 773-778, 2019.
[54] L. Zhou, J. D. Schaefferkoetter, I. W. Tham, G. Huang, and J. Yan, "Supervised learning with CycleGAN for low-dose FDG PET image denoising,"*Medical Image Analysis*, p. 101770, 2020.
[55] Y. Lei *et al.*, "Whole-body PET estimation from low count statistics using cycle-consistent generative adversarial networks,"*Physics in Medicine & Biology*, vol. 64, no. 21, p. 215017, 2019.
[56] X. Dong *et al.*, "Deep learning-based attenuation correction in the absence of structural information for whole-body positron emission tomography imaging,"*Physics in medicine and biology*, vol. 65, no. 5, p. 055011, 2020.
[57] S. Lu, J. Tan, Y. Gao, Y. Shi, and Z. Liang, "Prior knowledge driven machine learning approach for PET sinogram data denoising," in *Medical Imaging 2020: Physics of Medical Imaging*, 2020, vol. 11312: International Society for Optics and Photonics, p. 113124A.
[58] X. Hong, Y. Zan, F. Weng, W. Tao, Q. Peng, and Q. Huang, "Enhancing the image quality via transferred deep residual learning of coarse PET sinograms.,"*IEEE Trans Med Imaging*, vol. 37, no. 10, pp. 2322-2332, 2018.
[59] S. I. Sanaat A, Arabi H, Mainta I, Nkoulou R, and Zaidi H "Deep learning-assisted ultra-fast/low-dose whole-body PET/CT imaging,"*European journal of nuclear medicine and molecular imaging*, vol. in press, 2021.
[60] A. J. Ramon, Y. Yang, P. H. Pretorius, K. L. Johnson, M. A. King, and M. N. Wernick, "Improving Diagnostic Accuracy in Low-Dose SPECT Myocardial Perfusion Imaging with Convolutional Denoising Networks,"*IEEE Transactions on Medical Imaging*, 2020.
[61] I. Shiri *et al.*, "Standard SPECT myocardial perfusion estimation from half-time acquisitions using deep convolutional residual neural networks,"*Journal of nuclear cardiology : official publication of the American Society of Nuclear Cardiology*, 2020.
[62] M. P. Reymann *et al.*, "U-Net for SPECT Image Denoising," in *2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*: IEEE, pp. 1-2.
[63] H. Zaidi and N. Karakatsanis, "Towards enhanced PET quantification in clinical oncology,"*The British Journal of Radiology*, vol. 91, no. 1081, p. 20170508, 2017.
[64] H. Zaidi and K. F. Koral, "Scatter modelling and compensation in emission tomography,"*European journal of nuclear medicine and molecular imaging*, vol. 31, no. 5, pp. 761-782, 2004.
[65] A. Mehranian, H. Arabi, and H. Zaidi, "Vision 20/20: Magnetic resonance imaging-guided attenuation correction in PET/MRI: Challenges, solutions, and opportunities,"*Medical physics*, vol. 43, no. 3, pp. 1130-1155, 2016.
[66] J. Teuho *et al.*, "Magnetic resonance-based attenuation correction and scatter correction in neurological positron emission tomography/magnetic resonance imaging—current status with emerging applications,"*Frontiers in Physics*, vol. 7, p. 243, 2020.
[67] Y. Berker and Y. Li, "Attenuation correction in emission tomography using the emission data--A review.,"*Med Phys*, vol. 43, no. 2, pp. 807-832, 2016.
[68] A. Mehranian, H. Zaidi, and A. J. Reader, "MR-guided joint reconstruction of activity and attenuation in brain PET-MR.,"*NeuroImage*, vol. 162, no. 15, pp. 276-288, 2017.
[69] A. Rezaei, C. M. Deroose, T. Vahle, F. Boada, and J. Nuyts, "Joint reconstruction of activity and attenuation in fime-of-flight PET: A quantitative analysis.,"*J Nucl Med*, vol. 59, no. 10, pp. 1624-1629, 2018.
[70] A. Mehranian, H. Arabi, and H. Zaidi, "Quantitative analysis of MRI-guided attenuation correction techniques in time-of-flight brain PET/MRI,"*NeuroImage*, vol. 130, pp. 123-133, 2016.
[71] H. Arabi and H. Zaidi, "One registration multi-atlas-based pseudo-CT generation for attenuation correction in PET/MRI,"*European journal of nuclear medicine and molecular imaging*, vol. 43, no. 11, pp. 2021-35, 2016.
[72] H. Arabi, N. Koutsouvelis, M. Rouzaud, R. Miralbell, and H. Zaidi, "Atlas-guided generation of pseudo-CT images for MRI-only and hybrid PET-MRI-guided radiotherapy treatment planning,"*Physics in medicine and biology*, vol. 61, no. 17, pp. 6531-52, 2016.
[73] H. Arabi and H. Zaidi, "Comparison of atlas-based techniques for whole-body bone segmentation,"*Medical image analysis*, vol. 36, pp. 98-112, 2017.
[74] H. Arabi and H. Zaidi, "Truncation compensation and metallic dental implant artefact reduction in PET/MRI attenuation correction using deep learning-based object completion,"*Physics in medicine and biology*, vol. 65, no. 19, p. 195002, 2020.

[75] F. Liu, H. Jang, R. Kijowski, G. Zhao, T. Bradshaw, and A. B. McMillan, "A deep learning approach for 18 F-FDG PET attenuation correction,"*EJNMMI physics*, vol. 5, no. 1, pp. 1-15, 2018.
[76] X. Dong *et al.*, "Synthetic CT generation from non-attenuation corrected PET images for whole-body PET imaging,"*Physics in medicine and biology*, vol. 64, no. 21, p. 215016, 2019.
[77] L. Shi, J. A. Onofrey, H. Liu, Y. H. Liu, and C. Liu, "Deep learning-based attenuation map generation for myocardial perfusion SPECT,"*Eur J Nucl Med Mol Imaging*, vol. 47, no. 10, pp. 2383-2395, 2020.
[78] D. Hwang *et al.*, "Improving the Accuracy of Simultaneously Reconstructed Activity and Attenuation Maps Using Deep Learning,"*J Nucl Med*, vol. 59, no. 10, pp. 1624-1629, 2018.
[79] D. Hwang *et al.*, "Generation of PET Attenuation Map for Whole-Body Time-of-Flight (18)F-FDG PET/MRI Using a Deep Neural Network Trained with Simultaneously Reconstructed Activity and Attenuation Maps,"*Journal of nuclear medicine : official publication, Society of Nuclear Medicine*, vol. 60, no. 8, pp. 1183-1189, 2019.
[80] H. Arabi and H. Zaidi, "Deep learning-guided estimation of attenuation correction factors from time-of-flight PET emission data,"*Medical image analysis*, vol. 64, p. 101718, 2020.
[81] I. Shiri *et al.*, "Direct attenuation correction of brain PET images using only emission data via a deep convolutional encoder-decoder (Deep-DAC),"*Eur Radiol*, vol. 29, no. 12, pp. 6867-6879, 2019.
[82] J. Yang, D. Park, G. T. Gullberg, and Y. Seo, "Joint correction of attenuation and scatter in image space using deep convolutional neural networks for dedicated brain (18)F-FDG PET,"*Physics in medicine and biology*, vol. 64, no. 7, p. 075019, 2019.
[83] H. Arabi, K. Bortolin, N. Ginovart, V. Garibotto, and H. Zaidi, "Deep learning-guided joint attenuation and scatter correction in multitracer neuroimaging studies,"*Human brain mapping*, vol. 41, no. 13, pp. 3667-3679, 2020.
[84] I. Shiri *et al.*, "Deep-JASC: joint attenuation and scatter correction in whole-body (18)F-FDG PET using a deep residual network,"*European journal of nuclear medicine and molecular imaging*, vol. 47, no. 11, pp. 2533-2548, 2020.
[85] F. Liu, H. Jang, R. Kijowski, T. Bradshaw, and A. B. McMillan, "Deep Learning MR Imaging-based Attenuation Correction for PET/MR Imaging,"*Radiology*, vol. 286, no. 2, pp. 676-684, 2018.
[86] H. Arabi, G. Zeng, G. Zheng, and H. Zaidi, "Novel adversarial semantic structure deep learning for MRI-guided attenuation correction in brain PET/MRI,"*European journal of nuclear medicine and molecular imaging*, vol. 46, no. 13, pp. 2746-2759, 2019.
[87] A. P. Leynes *et al.*, "Zero-Echo-Time and Dixon Deep Pseudo-CT (ZeDD CT): Direct Generation of Pseudo-CT Images for Pelvic PET/MRI Attenuation Correction Using Deep Convolutional Neural Networks with Multiparametric MRI,"*J Nucl Med*, vol. 59, no. 5, pp. 852-858, 2018.
[88] C. N. Ladefoged, L. Marner, A. Hindsholm, I. Law, L. Højgaard, and F. L. Andersen, "Deep Learning Based Attenuation Correction of PET/MRI in Pediatric Brain Tumor Patients: Evaluation in a Clinical Setting,"*Frontiers in neuroscience*, vol. 12, p. 1005, 2018.
[89] C. N. Ladefoged *et al.*, "AI-driven attenuation correction for brain PET/MRI: Clinical evaluation of a dementia cohort and importance of the training group size,"*NeuroImage*, vol. 222, p. 117221, 2020.
[90] H. Xiang, H. Lim, J. A. Fessler, and Y. K. Dewaraja, "A deep neural network for fast and accurate scatter estimation in quantitative SPECT/CT under challenging scatter conditions,"*Eur J Nucl Med Mol Imaging*, vol. 47, no. 13, pp. 2956-2967, 2020.
[91] K. Armanious *et al.*, "Independent brain (18)F-FDG PET attenuation correction using a deep learning approach with Generative Adversarial Networks,"*Hellenic journal of nuclear medicine*, vol. 22, no. 3, pp. 179-186, 2019.
[92] R. R. Colmeiro, C. Verrastro, D. Minsky, and T. Grosges, "Whole Body Positron Emission Tomography Attenuation Correction Map Synthesizing using 3D Deep Generative Adversarial Networks," 2020.
[93] L. Shi *et al.*, "A novel loss function incorporating imaging acquisition physics for PET attenuation map generation using deep learning," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2019: Springer, pp. 723-731.
[94] T. J. Bradshaw, G. Zhao, H. Jang, F. Liu, and A. B. McMillan, "Feasibility of Deep Learning-Based PET/MR Attenuation Correction in the Pelvis Using Only Diagnostic MR Images,"*Tomography*, vol. 4, no. 3, pp. 138-147, 2018.
[95] H. Jang, F. Liu, G. Zhao, T. Bradshaw, and A. B. McMillan, "Technical Note: Deep learning based MRAC using rapid ultrashort echo time imaging,"*Med Phys*, 2018.
[96] I. Mecheter, A. Amira, M. Abbod, and H. Zaidi, "Brain MR Imaging Segmentation Using Convolutional Auto Encoder Network for PET Attenuation Correction," in *Proceedings of SAI Intelligent Systems Conference*, 2020: Springer, pp. 430-440.
[97] K. Gong, J. Yang, K. Kim, G. El Fakhri, Y. Seo, and Q. Li, "Attenuation correction for brain PET imaging using deep neural network based on Dixon and ZTE MR images,"*Phys Med Biol*, vol. 63, no. 12, p. 125011, 2018.

[98] P. Blanc-Durand *et al.*, "Attenuation correction using 3D deep convolutional neural network for brain 18F-FDG PET/MR: Comparison with Atlas, ZTE and CT based attenuation correction,"*PloS one,* vol. 14, no. 10, p. e0223141, 2019.
[99] K. D. Spuhler, J. Gardus, 3rd, Y. Gao, C. DeLorenzo, R. Parsey, and C. Huang, "Synthesis of Patient-Specific Transmission Data for PET Attenuation Correction for PET/MRI Neuroimaging Using a Convolutional Neural Network,"*J Nucl Med,* vol. 60, no. 4, pp. 555-560, 2019.
[100] A. Torrado-Carvajal *et al.*, "Dixon-VIBE Deep Learning (DIVIDE) Pseudo-CT Synthesis for Pelvis PET/MR Attenuation Correction,"*J Nucl Med,* vol. 60, no. 3, pp. 429-435, 2019.
[101] K. Gong, P. K. Han, K. A. Johnson, G. El Fakhri, C. Ma, and Q. Li, "Attenuation correction using deep Learning and integrated UTE/multi-echo Dixon sequence: evaluation in amyloid and tau PET imaging,"*Eur J Nucl Med Mol Imaging*, 2020.
[102] K. Gong *et al.*, "MR-based attenuation correction for brain PET using 3D cycle-consistent adversarial network,"*IEEE Transactions on Radiation and Plasma Medical Sciences*, 2020.
[103] A. P. Leynes *et al.*, "Bayesian deep learning Uncertainty estimation and pseudo-CT prior for robust Maximum Likelihood estimation of Activity and Attenuation (UpCT-MLAA) in the presence of metal implants for simultaneous PET/MRI in the pelvis,"*arXiv preprint arXiv:2001.03414*, 2020.
[104] A. Pozaruk *et al.*, "Augmented deep learning model for improved quantitative accuracy of MR-based PET attenuation correction in PSMA PET-MRI prostate imaging,"*European journal of nuclear medicine and molecular imaging*, 2020.
[105] L. Tao, J. Fisher, E. Anaya, X. Li, and C. S. Levin, "Pseudo CT Image Synthesis and Bone Segmentation from MR Images Using Adversarial Networks with Residual Blocks for MR-Based Attenuation Correction of Brain PET Data,"*IEEE Transactions on Radiation and Plasma Medical Sciences*, 2020.
[106] B. Rajalingam and R. Priya, "Multimodal medical image fusion based on deep learning neural network for clinical treatment analysis,"*International Journal of ChemTech Research,* vol. 11, no. 06, pp. 160-176, 2018.
[107] S. K. Kang *et al.*, "Adaptive template generation for amyloid PET using a deep learning approach,"*Human brain mapping,* vol. 39, no. 9, pp. 3769-3778, 2018.
[108] B. Huang *et al.*, "Fully automated delineation of gross tumor volume for head and neck cancer on PET-CT using deep learning: a dual-center study,"*Contrast media & molecular imaging,* vol. 2018, 2018.
[109] C. Lian, S. Ruan, T. Denoeux, H. Li, and P. Vera, "Joint tumor segmentation in PET-CT images using co-clustering and fusion based on belief functions,"*IEEE Transactions on Image Processing,* vol. 28, no. 2, pp. 755-766, 2018.
[110] X. Zhao, L. Li, W. Lu, and S. Tan, "Tumor co-segmentation in PET/CT using multi-modality fully convolutional neural network,"*Physics in Medicine & Biology,* vol. 64, no. 1, p. 015011, 2018.
[111] L. Zhao, Z. Lu, J. Jiang, Y. Zhou, Y. Wu, and Q. Feng, "Automatic nasopharyngeal carcinoma segmentation using fully convolutional networks with auxiliary paths on dual-modality PET-CT images,"*Journal of digital imaging,* vol. 32, no. 3, pp. 462-470, 2019.
[112] P. Blanc-Durand, A. Van Der Gucht, N. Schaefer, E. Itti, and J. O. Prior, "Automatic lesion detection and segmentation of 18F-FET PET in gliomas: a full 3D U-Net convolutional neural network study,"*PLoS One,* vol. 13, no. 4, p. e0195798, 2018.
[113] K. H. Leung *et al.*, "A physics-guided modular deep-learning based automated framework for tumor segmentation in PET,"*Physics in Medicine & Biology*, 2020.
[114] T. Wang *et al.*, "A learning-based automatic segmentation method on left ventricle in SPECT imaging," in *Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging*, 2019, vol. 10953: International Society for Optics and Photonics, p. 109531M.
[115] E. Roccia *et al.*, "Quantifying brain [18 F] FDG uptake noninvasively by combining medical health records and dynamic PET imaging data,"*IEEE Journal of Biomedical and Health Informatics,* vol. 23, no. 6, pp. 2576-2582, 2019.
[116] J. Park *et al.*, "Measurement of glomerular filtration rate using quantitative SPECT/CT and deep-learning-based kidney segmentation,"*Scientific reports,* vol. 9, no. 1, pp. 1-8, 2019.
[117] D. Visvikis, C. C. Le Rest, V. Jaouen, and M. Hatt, "Artificial intelligence, machine (deep) learning and radio (geno) mics: definitions and nuclear medicine imaging applications,"*European Journal of Nuclear Medicine and Molecular Imaging*, pp. 1-8, 2019.
[118] H. Wang *et al.*, "Comparison of machine learning methods for classifying mediastinal lymph node metastasis of non-small cell lung cancer from 18 F-FDG PET/CT images,"*EJNMMI research,* vol. 7, no. 1, p. 11, 2017.
[119] R. Seifert, M. Weber, E. Kocakavuk, C. Rischpler, and D. Kersting, "AI and Machine Learning in Nuclear Medicine: Future Perspectives," in *Seminars in Nuclear Medicine*, 2020: Elsevier.
[120] L. Xu *et al.*, "Automated whole-body bone lesion detection for multiple myeloma on 68Ga-Pentixafor PET/CT imaging using deep learning methods,"*Contrast media & molecular imaging,* vol. 2018, 2018.

[121] R. Togo *et al.*, "Cardiac sarcoidosis classification with deep convolutional neural network-based features using polar maps,"*Computers in biology and medicine*, vol. 104, pp. 81-86, 2019.
[122] L. Ma, C. Ma, Y. Liu, and X. Wang, "Thyroid diagnosis from SPECT images using convolutional neural network with optimization,"*Computational intelligence and neuroscience*, vol. 2019, 2019.
[123] I. R. Duffy, A. J. Boyle, and N. Vasdev, "Improving PET Imaging Acquisition and Analysis With Machine Learning: A Narrative Review With Focus on Alzheimer's Disease and Oncology,"*Molecular imaging*, vol. 18, p. 1536012119869070, 2019.
[124] D. Lu, K. Popuri, G. W. Ding, R. Balachandar, and M. F. Beg, "Multimodal and multiscale deep neural networks for the early diagnosis of Alzheimer's disease using structural MR and FDG-PET images,"*Scientific reports*, vol. 8, no. 1, pp. 1-13, 2018.
[125] H. Choi, K. H. Jin, and A. s. D. N. Initiative, "Predicting cognitive decline with deep learning of brain metabolism and amyloid imaging,"*Behavioural brain research*, vol. 344, pp. 103-109, 2018.
[126] M. Liu, D. Cheng, W. Yan, and A. s. D. N. Initiative, "Classification of Alzheimer's disease by combination of convolutional and recurrent neural networks using FDG-PET images,"*Frontiers in neuroinformatics*, vol. 12, p. 35, 2018.
[127] M. Liu, D. Cheng, K. Wang, Y. Wang, and A. s. D. N. Initiative, "Multi-modality cascaded convolutional neural networks for Alzheimer's disease diagnosis,"*Neuroinformatics*, vol. 16, no. 3-4, pp. 295-308, 2018.
[128] T. Zhou, K. H. Thung, X. Zhu, and D. Shen, "Effective feature learning and fusion of multimodality data using stage-wise deep neural network for dementia diagnosis,"*Human brain mapping*, vol. 40, no. 3, pp. 1001-1016, 2019.
[129] P. Afshar, A. Mohammadi, K. N. Plataniotis, A. Oikonomou, and H. Benali, "From handcrafted to deep-learning-based cancer radiomics: challenges and opportunities,"*IEEE Signal Processing Magazine*, vol. 36, no. 4, pp. 132-160, 2019.
[130] W. A. Noortman, D. Vriens, W. Grootjans, Q. Tao, L.-F. de Geus-Oei, and F. H. Van Velden, "Nuclear medicine radiomics in precision medicine: why we can't do without artificial intelligence,"*The Quarterly Journal of Nuclear Medicine and Molecular Imaging*, 2020.
[131] W. Xu, "Predictive power of a radiomic signature based on 18F-FDG PET/CT images for EGFR mutational status in NSCLC,"*Frontiers in oncology*, vol. 9, p. 1062, 2019.
[132] I. Shiri *et al.*, "Next-generation radiogenomics sequencing for prediction of EGFR and KRAS mutation status in NSCLC patients using multimodal imaging and machine learning algorithms,"*Molecular Imaging and Biology*, pp. 1-17, 2020.
[133] M. Edalat-Javid *et al.*, "Cardiac SPECT radiomic features repeatability and reproducibility: A multi-scanner phantom study,"*Journal of nuclear cardiology : official publication of the American Society of Nuclear Cardiology*, 2020.
[134] X. Fave *et al.*, "Delta-radiomics features for the prediction of patient outcomes in non–small cell lung cancer,"*Scientific reports*, vol. 7, no. 1, pp. 1-11, 2017.
[135] G. Dissaux *et al.*, "Pretreatment 18F-FDG PET/CT Radiomics Predict Local Recurrence in Patients Treated with Stereotactic Body Radiotherapy for Early-Stage Non–Small Cell Lung Cancer: A Multicentric Study,"*Journal of Nuclear Medicine*, vol. 61, no. 6, pp. 814-820, 2020.
[136] S. A. Mattonen *et al.*, "Bone Marrow and Tumor Radiomics at 18F-FDG PET/CT: Impact on Outcome Prediction in Non–Small Cell Lung Cancer,"*Radiology*, vol. 293, no. 2, pp. 451-459, 2019.
[137] F. Tixier *et al.*, "Comparison of Tumor Uptake Heterogeneity Characterization Between Static and Parametric 18F-FDG PET Images in Non–Small Cell Lung Cancer,"*Journal of Nuclear Medicine*, vol. 57, no. 7, pp. 1033-1039, 2016.
[138] S. Ashrafinia *et al.*, "Radiomics analysis of clinical myocardial perfusion SPECT to predict coronary artery calcification,"*Journal of Nuclear Medicine*, vol. 59, no. supplement 1, pp. 512-512, 2018.
[139] A. Rahmim *et al.*, "Improved prediction of outcome in Parkinson's disease using radiomics analysis of longitudinal DAT SPECT images,"*NeuroImage: Clinical*, vol. 16, pp. 539-544, 2017.
[140] P.-P. Ypsilantis *et al.*, "Predicting response to neoadjuvant chemotherapy with PET imaging using convolutional neural networks,"*PloS one*, vol. 10, no. 9, p. e0137036, 2015.
[141] T. Li, E. C. Ao, B. Lambert, B. Brans, S. Vandenberghe, and G. S. Mok, "Quantitative imaging for targeted radionuclide therapy dosimetry-technical review,"*Theranostics*, vol. 7, no. 18, p. 4551, 2017.
[142] Z. Peng *et al.*, "A method of rapid quantification of patient-specific organ doses for CT using deep-learning-based multi-organ segmentation and GPU-accelerated Monte Carlo dose computing,"*Medical Physics*, 2020.
[143] T. Xie and H. Zaidi, "Estimation of the radiation dose in pregnancy: an automated patient-specific model using convolutional neural networks,"*European Radiology*, vol. 29, no. 12, pp. 6805-6815, 2019.
[144] O. Schoppe *et al.*, "Deep learning-enabled multi-organ segmentation in whole-body mouse scans,"*Nature communications*, vol. 11, no. 1, pp. 1-14, 2020.

[145] W. Fu *et al.*, "iPhantom: a framework for automated creation of individualized computational phantoms and its application to CT organ dosimetry,"*arXiv preprint arXiv:2008.08730*, 2020.
[146] P. Jackson, N. Hardcastle, N. Dawe, T. Kron, M. S. Hofman, and R. J. Hicks, "Deep learning renal segmentation for fully automated radiation dose estimation in unsealed source therapy,"*Frontiers in oncology*, vol. 8, p. 215, 2018.
[147] X. Tang *et al.*, "Whole liver segmentation based on deep learning and manual adjustment for clinical use in SIRT,"*European Journal of Nuclear Medicine and Molecular Imaging*, pp. 1-11, 2020.
[148] T. Ryden, M. van Essen, J. Svensson, and P. Bernhardt, "Deep learning-based SPECT/CT quantification of 177Lu uptake in the kidneys,"*Journal of Nuclear Medicine*, vol. 61, no. supplement 1, pp. 1401-1401, 2020.
[149] A. Akhavanallaf, I. Shiri, H. Arabi, and H. Zaidi, "Whole-body voxel-based internal dosimetry using deep learning,"*European Journal of Nuclear Medicine and Molecular Imaging*, pp. 1-13, 2020.
[150] M. S. Lee, D. Hwang, J. H. Kim, and J. S. Lee, "Deep-dose: a voxel dose estimation method using deep convolutional neural network for personalized internal dosimetry,"*Scientific reports*, vol. 9, no. 1, pp. 1-9, 2019.
[151] T. I. Götz, C. Schmidkonz, S. Chen, S. Al-Baddai, T. Kuwert, and E. Lang, "A deep learning approach to radiation dose estimation,"*Physics in Medicine & Biology*, vol. 65, no. 3, p. 035007, 2020.
[152] T. I. Götz, E. Lang, C. Schmidkonz, T. Kuwert, and B. Ludwig, "Dose voxel kernel prediction with neural networks for radiation dose estimation,"*Zeitschrift für Medizinische Physik*, 2020.
[153] S. Xue, A. Gafita, A. Afshar-Oromieh, M. Eiber, A. Rominger, and K. Shi, "Voxel-wise Prediction of Post-therapy Dosimetry for 177Lu-PSMA I&T Therapy using Deep Learning,"*Journal of Nuclear Medicine*, vol. 61, no. supplement 1, pp. 1424-1424, 2020.
[154] M. Del Prete, F.-A. Buteau, and J.-M. Beauregard, "Personalized 177 Lu-octreotate peptide receptor radionuclide therapy of neuroendocrine tumours: a simulation study,"*European Journal of Nuclear Medicine and Molecular Imaging*, vol. 44, no. 9, pp. 1490-1500, 2017.
[155] M. Perez-Liva *et al.*, "Ultrafast Ultrasound Imaging for Super-Resolution Preclinical Cardiac PET.,"*Mol Imaging Biol*, vol. 22, no. 5, pp. 1342-1352, 2020.

Chapter 7

Ultra-low dose chest CT imaging of Covid-19 patients using deep neural networks

Isaac Shiri* , Azadeh Akhavanallaf*, Amirhossein Sanaat, et al.

Contribution: Study conception and design, methodological development, computer programming, analysis and interpretation of results, manuscript preparation.

Eur Radiol, 31(3), 1420–1431 (2021)

* Authors contributed equally

Abstract

Objectives The current study aimed to design an ultra-low-dose CT examination protocol using a deep learning approach suitable for clinical diagnosis of COVID-19 patients.

Methods In this study, 800, 170, and 171 pairs of ultra-low-dose and full-dose CT images were used as input/output as training, test and external validation set, respectively, to implement the full-dose prediction technique. A residual convolutional neural network was applied to generate full-dose from ultra-low-dose CT images. The quality of predicted CT images was assessed using root mean square error (RMSE), structural similarity index (SSIM) and peak signal-noise ratio (PSNR). Scores ranging from 1 to 5 were assigned reflecting subjective assessment of image quality and related COVID-19 features, including Ground-glass opacities (GGO), Crazy Paving (CP), Consolidation (CS), Nodular Infiltrates (NI), Bronchovascular thickening (BVT) and Pleural effusion (PE).

Results The radiation dose in terms of CT dose index ($CTDI_{vol}$) was reduced by up to 89%. The RMSE decreased from 0.16±0.05 to 0.09±0.02 and from 0.16±0.06 to 0.08±0.02 for the predicted compared to ultra-low-dose CT images in the test and external validation set, respectively. The overall scoring assigned by radiologists showed an acceptance rate of 4.72±0.57 out of 5 for reference full-dose CT images, while ultra-low-dose CT images rated 2.78±0.9. The predicted CT images using the deep learning algorithm achieved a score of 4.42±0.8.

Conclusions The results demonstrated that the deep learning algorithm is capable of predicting standard full-dose CT images with acceptable quality for the clinical diagnosis of COVID-19 positive patients with substantial radiation dose reduction.

Abbreviations

SARS: Severe acute respiratory syndrome
COVID-19: Coronavirus disease 2019
CT: Computed tomography
GGO: Ground glass opacities
RT-PCR: Real-time reverse transcription-polymerase chain reaction
SARS-CoV-2: Severe acute respiratory syndrome coronavirus 2
WHO: World Health Organization
DLP: Dose-length product
ED: Effective dose
SNR: Signal-to-noise ratio
CNR: Contrast-to-noise ratio
GAN: Generative adversarial network
CNN: Convolutional neural network
AEC: Automatic exposure control
FBP: Filtered backprojection
ADMIRE: Advanced modeled iterative reconstruction
CP: Crazy Paving
CS: Consolidation
NI: Nodular Infiltrates
BVT: Bronchovascular thickening
PE: Pleural effusion
CTDI: CT dose index

I. Introduction

The emergence of novel coronavirus in December 2019 in Wuhan, China, known as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) was recognized as a global public health concern by the World Health Organization (WHO) [1]. SARS-CoV-2 disease 2019 or COVID-19 is an infectious disease that affects the upper and lower respiratory tract and induces mild to severe respiratory syndromes, including pneumonia [2]. Real-time reverse transcription-polymerase chain reaction (RT-PCR) is considered the standard method for COVID-19 diagnosis but is prone to a number of limitations, including the time of preparation and false-positive and false-negative rates in different clinical samples [3]. Conversely, early studies confirmed that computed tomography (CT) is a feasible approach for COVID-19 diagnosis [4]. Until recently, a wide range of clinical studies have been conducted on the feasibility of CT findings in the early detection and management of COVID-19 patients. However, there are still considerable knowledge gaps in the recognition of CT features linked to COVID-19 [4, 5].

As CT examinations account for the major cause of radiation exposure to the general public from diagnostic medical imaging procedures, the development of low-dose CT imaging protocols is highly desirable. A recent study demonstrated that DNA double-strand breaks, and chromosome aberrations increased in patients undergoing a standard-dose CT examination while no effect on human DNA was detected in patients undergoing low-dose CT scans [6]. Although a plethora of hardware and software technological advances in CT dose reduction have been reported, including high-sensitivity detectors, new automatic exposure control (AEC) systems, adaptive x-ray tube voltage and new image reconstruction algorithms, CT is still not a low-dose imaging modality [7]. Therefore, the level of radiation exposure from this modality is still a matter of concern [8]. Task-specific low-dose imaging protocols devised in both academic and corporate settings were adopted in clinical setting [9]. Zhou *et al.* [10] suggested a low-dose CT protocol enabling to significantly reduce the dose-length product (DLP) and effective dose (ED) without sacrificing signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR). Nevertheless, converting from conventional full-dose to low-dose CT imaging protocols is not a simple task owing to the fear of increasing the false positive rate due to the elevated level of noise and missing anatomical structures.

A number of professional societies, scientists and clinicians proposed appropriate low-dose CT protocols for COVID-19 [11-14]. However, these protocols are not widely deployed in clinical centers for the same above mentioned reasons. Clinicians and radiologists often tend to use established protocols employing full dose CT imaging and often lack time or are reluctant to develop or adopt new protocols, especially during emergency situations, such as during the COVID-19 outbreak.

In addition to conventional denoising approaches [15, 16], a number of deep learning algorithms have been proposed for medical image analysis [17-19], PET [20] and SPECT [21] denoising as well as CT image denoising and enhancement of image quality [10, 22-25]. Yang *et al.* [22] applied a generative adversarial network (GAN) with Wasserstein distance and perceptual loss to denoise low-dose CT images. In another study, Kim *et al.* [23] investigated the effect of different loss functions on convolutional neural network (CNN)-based image denoising performance using task-based image quality assessment for various signals and dose levels. Shin *et al.* [24] compared image quality of low-dose CT images obtained using a deep learning-based denoising algorithm with low-dose CT images reconstructed using filtered-backprojection (FBP) and advanced modeled iterative reconstruction (ADMIRE). They reported that deep learning techniques achieved better noise properties compared to FBP and ADMIRE reconstructions of low-dose CT images. In this work, we aimed to use deep learning algorithms on ultra-low-dose COVID-19 CT images to generate high quality images for a comparable diagnostic accuracy with full-dose CT images.

II. Materials and Methods

Data acquisition

This retrospective study was approved by the ethics committees of the participating centers. Written consent was waived with approval. We included 1141 volumetric chest CT exams from 9 medical centers, among which 312 volumetric CT images were from PCR-positive COVID-19 patients. COVID-19 patients were collected from three centers and various scanner models, including Emotion 16 (Siemens Healthcare), NeuViz Dual (Neusoft Medical

Systems) and Optima CT580 (GE Healthcare). All CT images were acquired in each center using the same protocol and were reconstructed using a filtered back-projection (FBP) algorithm (Table 1).

Table 1. Acquisition parameters of full-dose and low-dose chest CT protocols.

Parameters	Full-dose CT	Low-dose CT
$CTDI_{vol}$ (mGy)	6.5 (4.16-10.5)	0.72 (0.66-1.03)
Voltage (kVp)	100-120	90
Tube current (mA)	100-150	20-45
Pitch factor	1.3-1.8	0.75

Ultra-low-dose CT simulation

Based on Beer-Lambert law ($I = I_0 \exp(-\int \mu(e,x)dx)$), the incident flux level of the ultra-low-dose scan (I_0) can be calculated by adequately scaling the incident flux level of the corresponding full dose scan. According to the physics of CT transmission data (Eq. 1), we simulated ultra-low-dose CT projection data from full-dose projections in the sinogram domain by adding a statistically independent Poisson noise distribution and a Gaussian noise distribution.

$$\hat{I} = Poisson(I_0) + \text{Gaussian}(m_e, \sigma_e^2) \quad (1)$$

where $\hat{I}$ is the measured noisy signal recorded in the detector channels, I_0 is the mean number of photons passing through the patient determined based on a linear relationship with tube current (mAs). m_e and σ_e^2 are the mean and variance of the electronic noise, respectively. The whole procedure is as follows:

1. Converting Hounsfield Units (HUs) to linear attenuation coefficients according to tube voltage in the full-dose image ($\mu_{tissue} = \frac{HU \times (\mu_{water} - \mu_{air})}{1000} + \mu_{water}$),
2. Generating projection data (p_{sd}) from the attenuation map (μ_{tissue}) using the Radon transform on the full-dose image with the following setups: parallel beam geometry and 1080 projection angles in one rotation,
3. Converting projection data to the transmission data, i.e. $T_{sd} = \exp(-p_{sd})$,
4. Generating ultra-low-dose transmission data by multiplying ultra-low-dose scan incident flux by full-dose transmission data, i.e. $T_{uld} = I_0^{uld} \times T_{sd}$,
5. Simulating the noise in ultra-low-dose scan by adding Poisson noise and Gaussian noise to the transmission data, i.e. $I_{uld} = Poisson(T_{uld}) + \text{Gaussian}(m_e, \sigma_e^2)$,
6. Calculating ultra-low-dose projection data in the sinogram domain, i.e. $p_{uld} = \log(\frac{I_0^{uld}}{I_{uld}})$,
7. Reconstruction of the ultra-low-dose images using FBP algorithm,
8. Converting the reconstructed attenuation map to HU using the equation in step 1.

In the above-mentioned steps for simulating ultra-low-dose scan, three parameters should be determined, namely, the ultra-low-dose scan incident flux (I_0^{uld}), the mean (m_e) and the variance (σ_e^2) of electronic noise. In modern CT scanners, these parameters can be determined during routine calibration procedures. However, this is not practical for multi-centric clinical database. Hence, these parameters were set based on fitting noise level of the simulated ultra-low-dose CT images with a real ultra-low-dose CT image-set serving as reference. The reference ultra-low-dose CT images were acquired under a task-specific ultra-low-dose protocol for the diagnosis of COVID-19 on the MX 16-slice CT scanner (Philips Healthcare) with a reduced CT dose index ($CTDI_{vol}$) of about 0.72 mGy. The acquisition parameters of the protocol were as follows: tube potential of 90 kVp, tube current range of 20-45 mA, 0.5 sec rotation time, and pitch factor of 0.75 with the FBP image reconstruction procedure. To quantify the noise level of the simulated ultra-low-dose CT images, the noise index was produced based on the method proposed by Christianson *et al.* [26]. The incident flux level (I_0^{uld}) was determined when the magnitudes of noise levels in soft-tissue and lungs between simulated ultra-low-dose images were within 10% interval

compared to that in the reference images. In electronic systems, m_e is usually calibrated to be zero whereas the variance of electronic noise was initialized based on the method proposed by Zeng *et al.* [27] for the Definition, Edge CT scanner (Siemens Healthcare). Subsequently, an observer study was performed to evaluate the quality of simulated ultra-low-dose images against the full-dose images. Three physicists took part in this study to visually score the apparent Poisson noise and streak artifacts owing to statistical errors originating from low photon scanning and Gaussian noise. We categorized our dataset into multiple groups according to the scanner model and imaging protocol used. Consequently, three image-sets were randomly selected from each group for the evaluation process. Two ROIs (5×5 cm^2) were drawn in the soft-tissue (upper part of the liver) and lung regions without including adjacent anatomic structures. The average standard deviation (STD) across the ROIs was calculated. The simulation parameters were updated to obtain the same STD in two ROIs drawn on soft-tissue and lungs, while the visual similarity between simulated ultra-low-dose image and full dose image was preserved.

Deep learning algorithm

Network architecture

We applied a deep residual neural network (ResNet) for image to image transformation in an attempt to predict full-dose from ultra-low-dose CT images [28]. The residual model proposed by Wenqi et al. [28] for image classificaltion was modified for regression application in this study. Figure 1 presents the architecture of ResNet employed in the current study. This network combines 20 convolutional layers, including two seven and one six convolutional layers for low, medium, and high-level features extraction. For effective feature extraction, the ResNet architecture adopts a dilated convolution with factors of 2 and 4 for seven intermediate and six last layers. In this combination, every two convolutional layers are linked to a residual connection where a leaky rectified linear unit (LReLU) acts as an activation function. The ResNet implemented in TensorFlow (version 1.12.1) was utilized to transform ultra-low-dose to full dose chest CT images.

In this work, a 3×3×1 kernel was applied for all convolutions .The ResNet network has residual connections that bypass the parameterized layers through combining the input and output of a block to render a smooth information propagation, thus enhancing the training speed/quality. The ResNet architecture benefits from 9 residual blocks that proved efficient for improving the feature extraction process. This work avoids a large number of trainable parameters. More detail of ResNet architecture is presented in Figure1.

Implementation details

In this study, 800 (112 COVID-19), 170 (100 COVID-19), and 171 (100 COVID-19) pairs of ultra-low-dose and full-dose CT studies were used as input/output as training, test and external validation set, respectively, to implement the full-dose prediction technique. The ResNet model with an architecture of a 2D spatial window equal to 512×512voxels was employed (CT images were cropped to eliminate the bed and background air). To train the network, Adam optimizer and L2norm loss function were adopted. The training of the network for full-dose prediction took about 50 hours using a 2080TI GPU, Intel(R) Xeon 2.30 GHz 7i CUP, and 64 GB RAM. After ten epochs, the training loss reached its plateau.

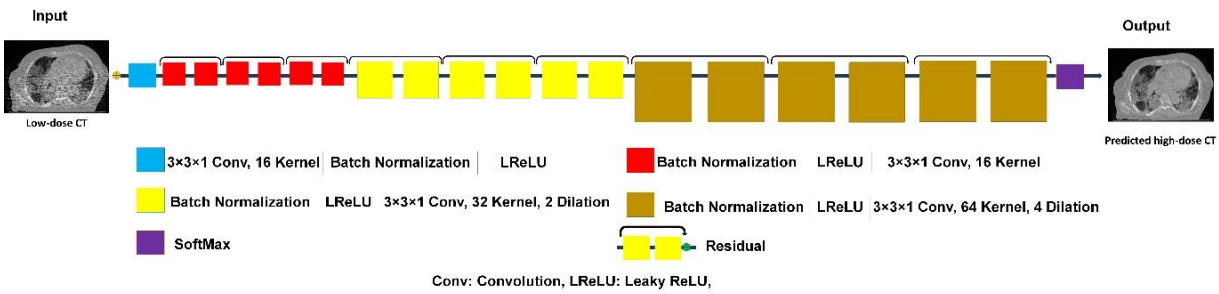


Figure 1. Architecture of the deep residual neural network (ResNet) along with details of the associated layers. Red color layer: layer with dilation 1, yellow color layer: layer with dilation 2, brown color layer: layer with dilation 4. Conv: convolutional kernel; LReLu: leaky rectified linear unit; SoftMax: Softmax function; Residual: residual connection.

Quantitative evaluation

Our qualitative and quantitative evaluation of the framework was performed on 170 tests and 171 external validation set. To this end, ultra-low-dose and predicted images were compared to reference full-dose images. The quality of CT images was assessed using voxel-wise root mean square error (RMSE). Moreover, the structural similarity index (SSIM) and peak signal-to-noise ratio (PSNR) were used as quantitative measures of the quality of the predicted CT images.

Clinical evaluation

All patient chest CT images were categorized into three groups, including full-dose, ultra-low-dose and predicted by lung windowing. Blind qualitative assessment of CT images was performed by a radiologist with ten years of experience. The radiologists' clinical evaluations were based on qualitative assessment, including appraisal of lesion density, shape, position, and margin in addition to the analysis of lesion type. For the qualitative assessment, scores ranging from 1 to 5 were assigned to each image as follows: excellent: 5, good: 4, adequate, 3, poor: 2 and uninterpretable: 1. This scoring scheme was separately used for overall assessment of image quality, i.e.margin, shape, and density as well as for lesion type. Lesion types included Ground-glass opacities (GGO), Crazy Paving (CP), Consolidation (CS), Nodular Infiltrates (NI), Bronchovascular thickening (BVT) and Pleural effusion (PE). To categorize lesions based on their location, they were attributed to any of the following anatomical regions in the lung: left lung, right lung, upper zone, lower zone, middle zone, superior segment, posterior segment, central and peripheral areas.

III. Results

The mean value of $CTDI_{vol}$ for the ultra-low-dose protocol based on which the simulation parameters are determined is about 0.72 mGy (range 0.66-1.02 mGy) (Table 1). In contrast, this index ranges from 4.16 to 10.5 mGy with an average of 6.5 mGy for the full-dose protocol. According to the adopted methodology, the incident flux was determined in the range 3.5- 4×10^3 for different scanner models.

The quantitative metrics, including RMSE, PSNR and SSIM for predicted full-dose and ultra-low-dose CT images in the test and external validation sets are plotted as box plots in Figure 2 and summarized in Table 2. The RMSE in units of normalized HU decreased from 0.16±0.05 to 0.09±0.02 and from 0.16±0.06 to 0.08±0.02 for predicted full-dose images from ultra-low-dose CT images in test and external validation set, respectively. The SIMM and PSNR increased from 0.89±0.07 to 0.97±0.01 and from 29.40±4.94 to to 33.60±2.70 for predicted full-dose images in external validation set, respectively.

Table 2. Mean and STD of peak signal-to-noise ratio (PSNR), structural similarity index (SSIM) and root mean square error (RMSE) for the predicted and ultra-low-dose CT images in the test and external validation sets and statistical difference between preicted and ultra-low-dose images.

Parameters	Images	Test	External Validation
RMSE	Predicted	0.09±0.02	0.08±0.02
	Ultra-Low-Dose	0.16±0.05	0.16±0.06
P-Value		P < 0.0001	P < 0.0001
PSNR	Predicted	32.97±2.60	33.60±2.70
	Ultra-Low-Dose	28.44±3.87	29.40±4.94
P-Value		P < 0.0001	P < 0.0001
SSIM	Predicted	0.97±0.02	0.97±0.01
	Ultra-Low-Dose	0.89±0.07	0.89±0.07
P-Value		P < 0.0001	P < 0.0001

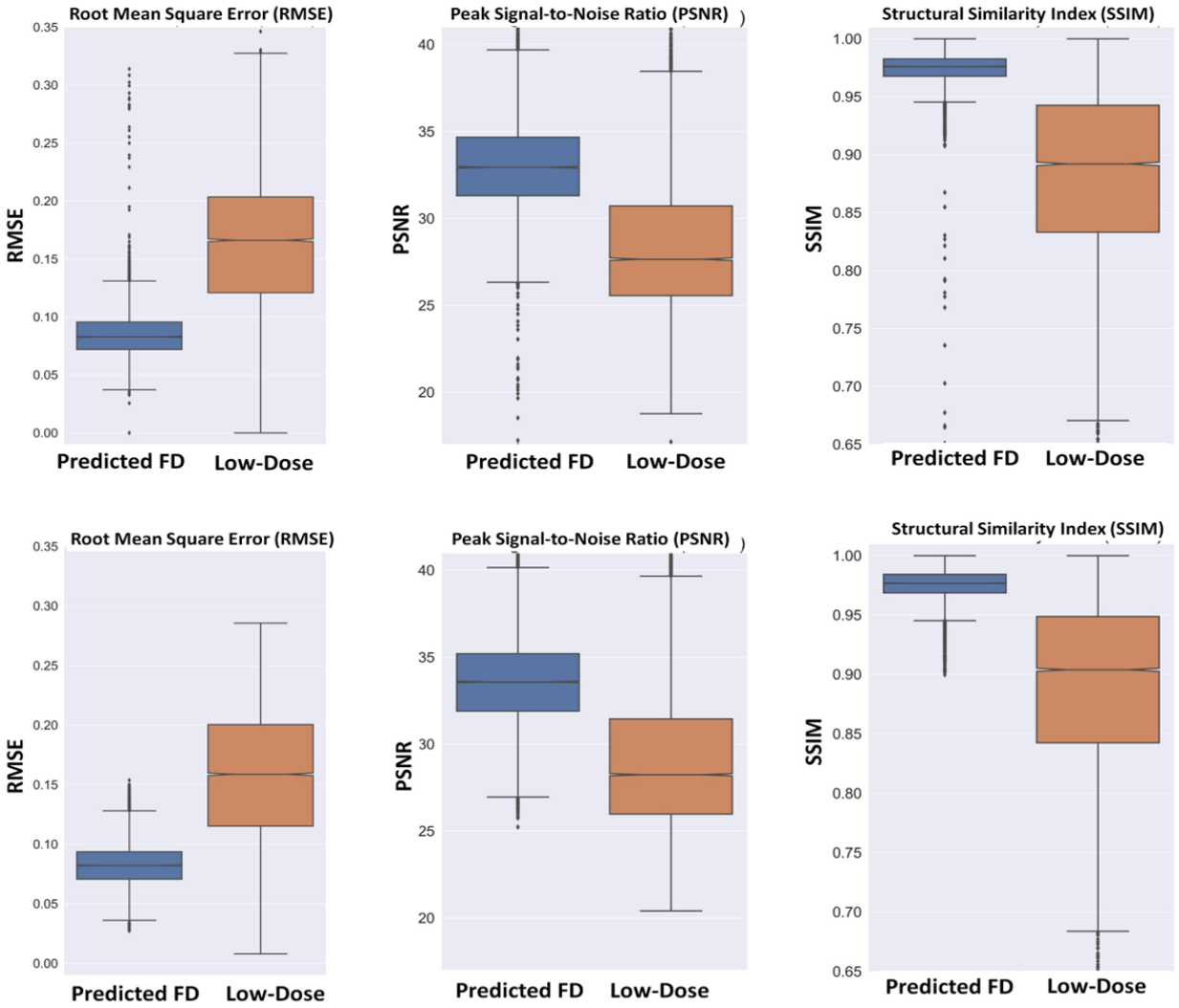


Figure 2. Mean and STD of peak signal-to-noise ratio (PSNR), structural similarity index (SSIM) and root mean square error (RMSE) for the predicted and ultra-low-dose CT images in the test (top) and external validation (bottom) sets.

Overall results associated with assessment of image quality are shown in figure 3a wherein high image quality variations can be observed in ultra-low-dose scans, while the predicted full-dose images are mostly scored good or excellent. Overall scoring shows that the full-dose images received the highest score (4.72±0.57) whereas the ultra-low-dose images were rated with the lowest scores (2.78±0.9). In figure 3b, the frequency of occurrence of each lesion type in the different series of images is shown. As can be seen, GGO has the highest occurence in all images, whereas mixed (all) had the same occurence for all images. Changes in the essence of features are as follows: in the ultra-low-dose group, GGO is shifted to normal feature whereas consolidation is turned to GGO.

Lesion detectability scoring results are shown in figure 4. The excellent score (score=5) for CS in full-dose images is in about 60% of the cases while it exceeds 90% in predicted full-dose CT images. CP, NI, and PE achieved an excellent score (100%) in predicted images is more than 40%, 70%, and 40% of the cases, respectively. The overall image quality scores assigned by human observers for different lesions are summarized in Table 3. Table 4 presents the visual scoring of different images for different aspects of CT findings, including lesion status, margin, shape, and density.

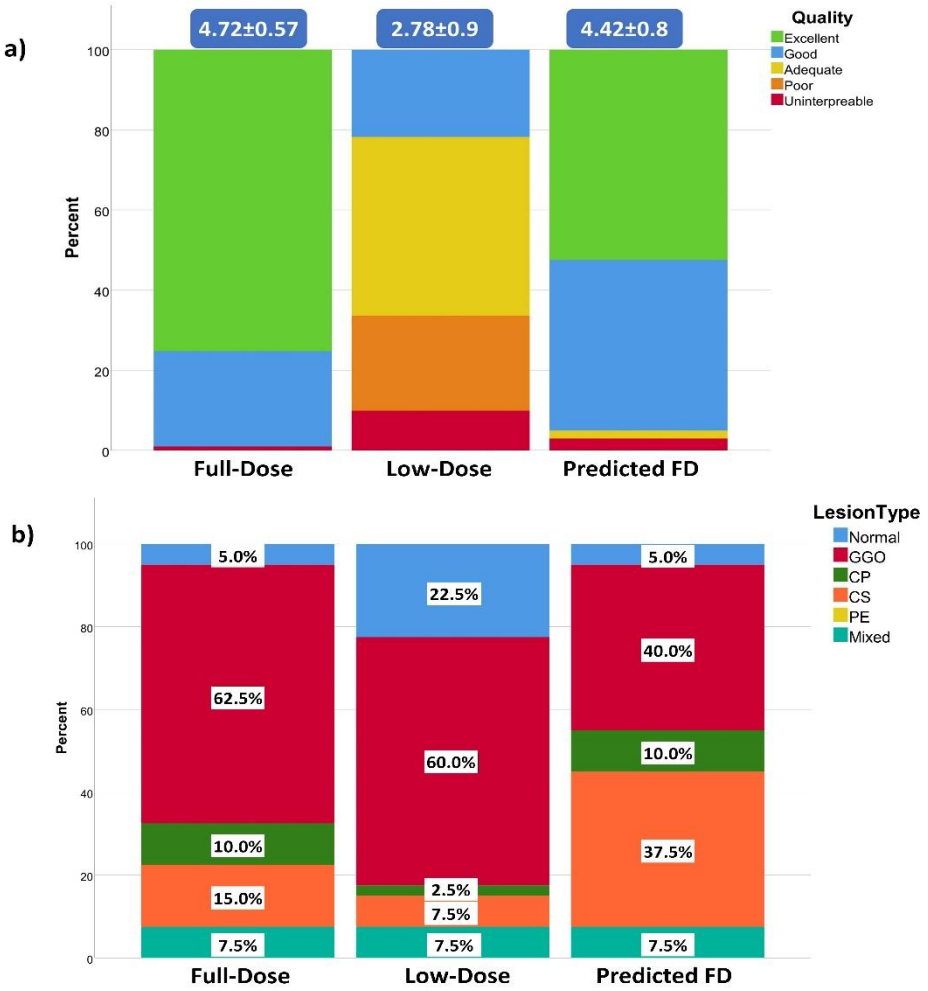


Figure 3. Image quality scoring of different images (a), Lesion type frequency in different images (b). Ground-glass opacities (GGO), Crazy Paving (CP), Consolidation (CS), Nodular Infiltrates (NI), Bronchovascular thickening (BVT), and Pleural effusion (PE). Scores (excellent: 5, good: 4, adequate, 3, poor: 2 and uninterpretable: 1)

Figure 5 and supplemental figures 1 and 2 presents a representative example of a full-dose, ultra-low-dose, and predicted full-dose CT images. The predicted CT images improved image quality, thus enabling most lesions to be easily classified. Figure 6 and supplemental figures 3 and 4 shows an example of an outlier in which image quality was improved; however, some relevant anatomical details were missing. Hence, the network failed to recover the full detail of images and GGO lesion converted to CS. For an outlier in the predicted group, GGO was shifted to consolidation.

Table 3. Image quality scores assigned by human observers for different lesions. Ground-glass opacities (GGO), Consolidation (CS), Crazy Paving (CP), Nodular Infiltrates (NI), Bronchovascular thickening (BVT) and Pleural effusion (PE). Scores (excellent: 5, good: 4, adequate, 3, poor: 2 and uninterpretable: 1).

Lesions	Full-dose	Ultra-low-dose	Predicted
GGO	4.70±0.47	2.67±0.61	3.90±1.09
CS	4.52±0.87	3.36±0.64	4.92±0.28
CP	5.00±0.00	3.00±0.00	4.50±0.71
NI	5.00±0.00	3.25±0.50	4.75±0.50
BVT	4.79±0.41	2.44±1.11	4.44±0.56
PE	5.00±0.00	2.50±1.05	4.50±0.55

Table 4. Image quality assessment through visual scoring of different images documenting different aspects of CT findings. Scores (excellent: 5, good: 4, adequate, 3, poor: 2 and uninterpretable: 1).

CT findings			Full-dose	Low-dose	Predicted
Lesion Status	Laterality	Left Lung	4.66±0.55	3.14±0.69	4.52±0.51
		Right Lung	4.70±0.53	3.12±0.65	4.52±0.51
	Cephalocaudal distribution	Upper	4.44±0.63	2.94±0.44	4.25±0.45
		Lower	4.68±0.54	3.10±0.60	4.48±0.51
		Middle	4.71±0.53	3.23±0.56	4.48±0.51
	Location	Central	4.67±0.58	3.33±1.15	5.00±0.00
		Peripheral	4.76±0.44	3.12±0.70	4.71±0.47
		Superior	4.65±0.59	3.25±0.64	4.60±0.50
		Posterior	4.68±0.54	3.23±0.62	4.65±0.49
		Central & Peripheral	4.63±0.62	3.19±0.54	4.44±0.51
Margin		Ill Defined	4.48±0.75	2.30±0.91	4.19±0.56
		Well Defined	4.67±0.55	3.15±0.60	4.93±0.27
Shape	Nodular		5.00±0.00	4.00±0.00	5.00±0.00
	Wedged		5.00±0.00	3.33±0.82	5.00±0.00
	Elongated		4.00±1.41	2.00±1.41	4.50±0.71
	Confluent		4.54±0.66	3.00±0.66	4.54±0.51
Density	Part Solid		4.83±0.41	2.40±1.14	3.60±1.52
	Solid		4.60±1.26	3.40±0.70	4.80±0.42
	Pure GGO		4.63±0.49	2.79±0.66	3.96±1.27
	GGO and CS		5.00±0.00	2.80±0.84	4.40±0.55

IV. Discussion

Despite the controversies and heated debates around the potential haphazardous effects of low-levels of ionizing radiation and the linear-no-threshold theory [29], concerns from radition exposure are still current [30]. Since CT imaging is widely used in clinical diagnosis, prognosis, and assessment of response to treatment and follow-up of a number of diseases, it is an incremental source of radiation dose to patients in modern healthcare [7, 31]. With respect to the current COVID-19 crisis, chest CT imaging is the fastest diagnostic approach. However, it remains a high dose imaging modality, and as such, developing a ultra-low-dose protocol enabling to maintain optimal image quality is clinically relevant in public health management. Therefore, as a reponse to this outbreak and the subsequent demand for CT imaging for mass population, an ultra-low-dose imaging approach was proposed to minimize radiation exposure of the population. This is achieved through a deep learning approach introduced for COVID-19 patients diagnosis by generating high-quality full-dose from ultra-low-dose CT images. It was shown that although the simulated ultra-low-dose CT images were diagnostically compromised, the generated full-dose images were appropriate for the task at hand. The proposed ultra-low-dose approach based on deep learning algorithms succeded to reduce the $CTDI_{vol}$ by up to 89%, reflecting a substantial reduction of the radiation dose associated with diagnostic CT examinations.

A number of studies have assessed the role of low-dose CT for COVID-19 management [11-14]. Agostini *et al.* [12], evaluated the feasibility and diagnostic reliability of a low-dose, long-pitch dual-source chest CT protocol for COVID-19 patients in terms of signal-to-noise and contrast-to-noise ratio and Likert scales. They reported that their low-dose CT protocol achieved significant dose reduction, lower motion artifacts with optimum signal and contrast-to-noise ratio. However, this protocol is only applicable on third-generation dual-source CT scanners, and as such, it not applicable on older CT imaging systems. Dangis *et al.* [14] examined the accuracy and reproducibility of low-dose sub-millisievert chest CT for the diagnosis of COVID-19. They demonstrated that low-dose CT has excellent sensitivity, specificity, positive predictive value, negative predictive value, and accuracy for diagnosis of COVID-19 with a mean effective dose of 0.56±0.25 mSv. In the current study, the simulated ultra-

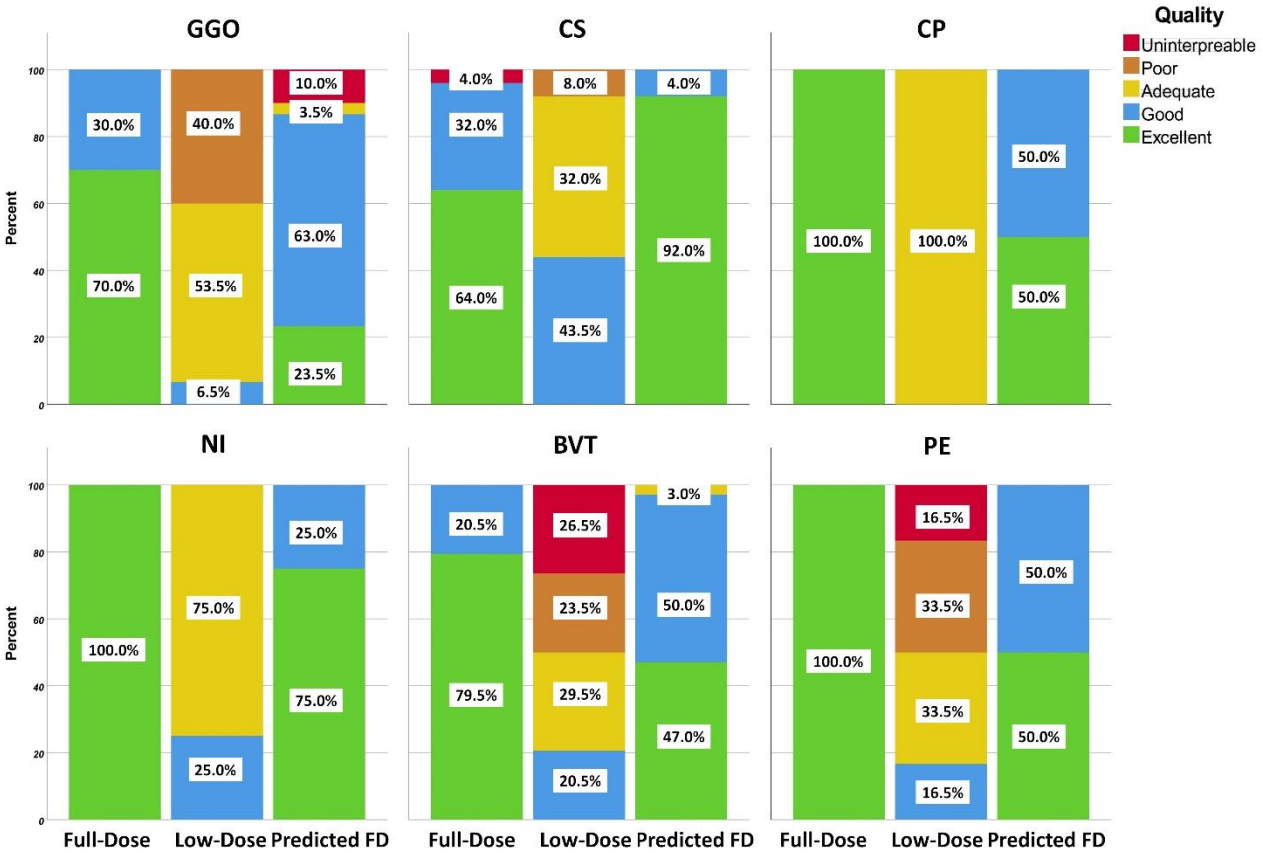


Figure 4. Image quality scoring of different images. Ground-glass opacities (GGO), Crazy Paving (CP), Consolidation (CS), Nodular Infiltrates (NI), Bronchovascular thickening (BVT), and Pleural effusion (PE). Scores (excellent: 5, good: 4, adequate, 3, poor: 2 and uninterpretable: 1).

low-dose CT images represent the outcome of a protocol with a significant reduction of $CTDI_{vol}$ (up to 89%) compared to the corresponding full dose CT images, which is a good metric for comparing patient effective dose and risks of ionizing radiation [32]. This is a commended effort in view of the current recommendations in radiation protection [33], particularly for the diagnosis and follow up of sensitive population, such as pediatric patients and pregnant women.

The results of this study demonstrated that by using CNNs, we could generate images with a significantly lower dose and acceptable image quality. Although image quality in the predicted images was not exactly identical to full-dose CT images, most COVID-19 features, including nodular infiltrate, consolidation, and crazy paving features obtained high scores, almost similar to full-dose CT images.

We also demonstrated that the texture of COVID-19 lesions could be erroneously altered in the predicted CT images, which would skew the diagnosis/scoring. We observed that in the ultra-low-dose group, GGO was shifted to normal feature, whereas consolidation was shifted to GGO. In the low-dose group, the shift of GGO to normal features might be due to closeness of mean HU value of GGO to normal. In addition, as the differences between the HU value of GGO and consolidation lesions are located in the normal neighborhood, they may be depicted and diagnosed as similar features. Likewise, in the predicted group, GGO was shifted to consolidation owing to the

local induced bias noise pattern in ultra-low dose images, heterogeneity of lesions, and smoothing effect of deep learning in some outlier cases. The low-dose simulation would result in overall zero bias (zero-mean noise signal) with elevated noise variance depending on the underlying signals/textures and level of simulated low dose scanning. Due to the fine texture as well as relatively low density (low CT numbers) of GGO lesions, the streak-like noise patterns led to mostly positive bias and rougher textures in these lesions. As such, the likelihood of misinterpretation of GGO with CP increased in the resulting synthetic standard-dose CT images. In addition, the minimum widely used learnable kernel employed in the current study is (3×3), which would slightly smooth the structures of the resulting synthetic images. The local positive noise-induced bias alomg with the smoothness of the structures in the resulting CT images led to the misidentification of some GGO lesions with CP.

Although ultra-low-dose CT can be equally effective in COVID-19 detection and diagnosis as the full-dose CT, it suffers from a number of limitations, particularly the increased noise level caused by photon deprivation. One of the limitations of the present study was that during the clinical assessment, the ultra-low-dose images could be easily identified by radiologists because of the high of noise present. This might have led them to be subconsciously biased, hence assigning lower scores to these images. We reported outliers originating mostly from the low quality of the simulated ultra-low-dose CT images (high noise level and/or noise-induced artifact) caused by photon starvation in simulated corpulent patients. Application of the current method in COVID-19 imaging warranted a thorough investigation of outliers owing to inter/intra-patient variation and noise variability.

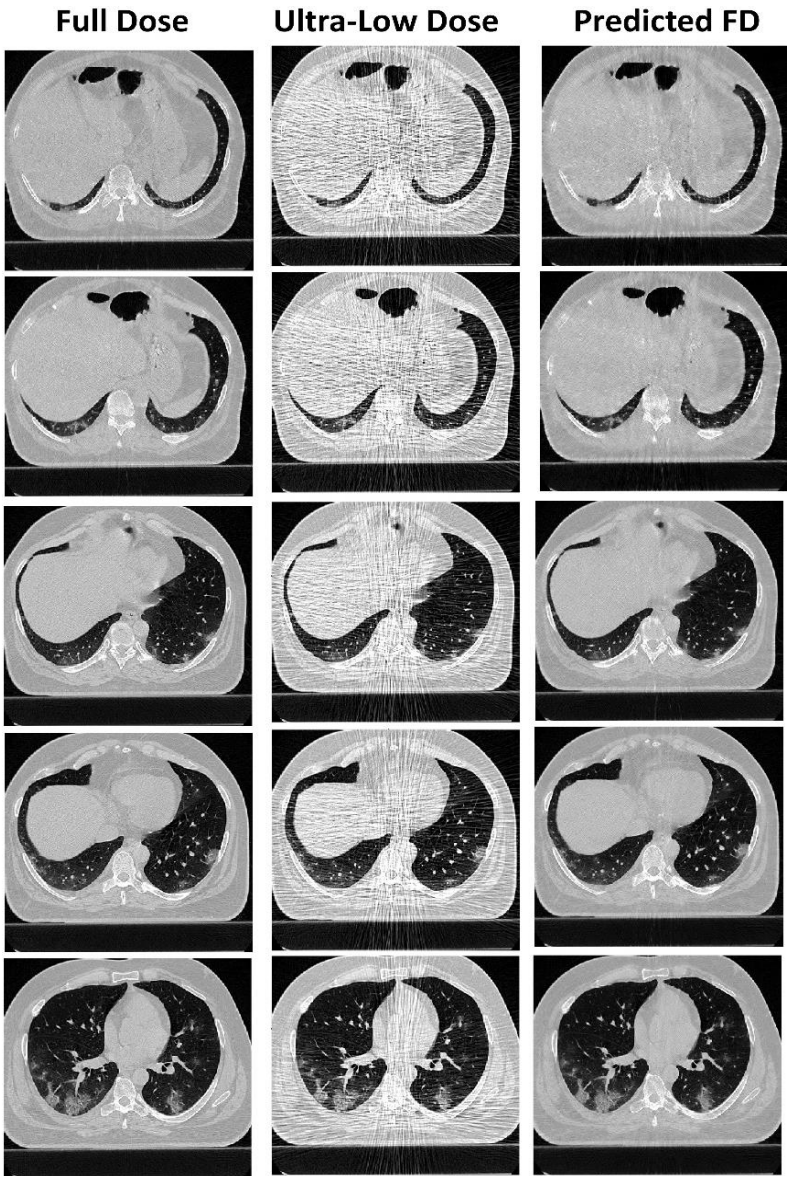


Figure 5. Representative full-dose image and corresponding ultra-low-dose and predicted full-dose images.

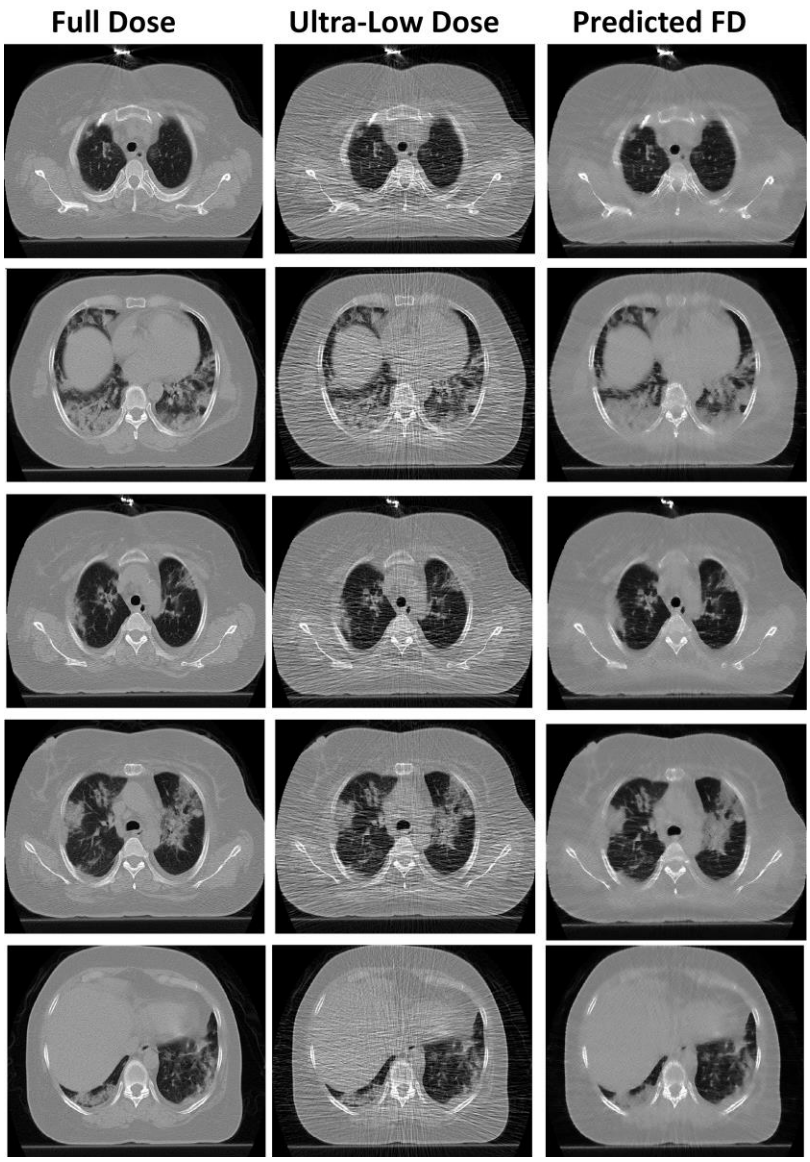


Figure 6. Outlier report: CT images of a patient where the deep learning algorithm improved image quality but changed the patchy lesion to consolidation in predicted images. The red arrows pinpoint changes in the identified lesions.

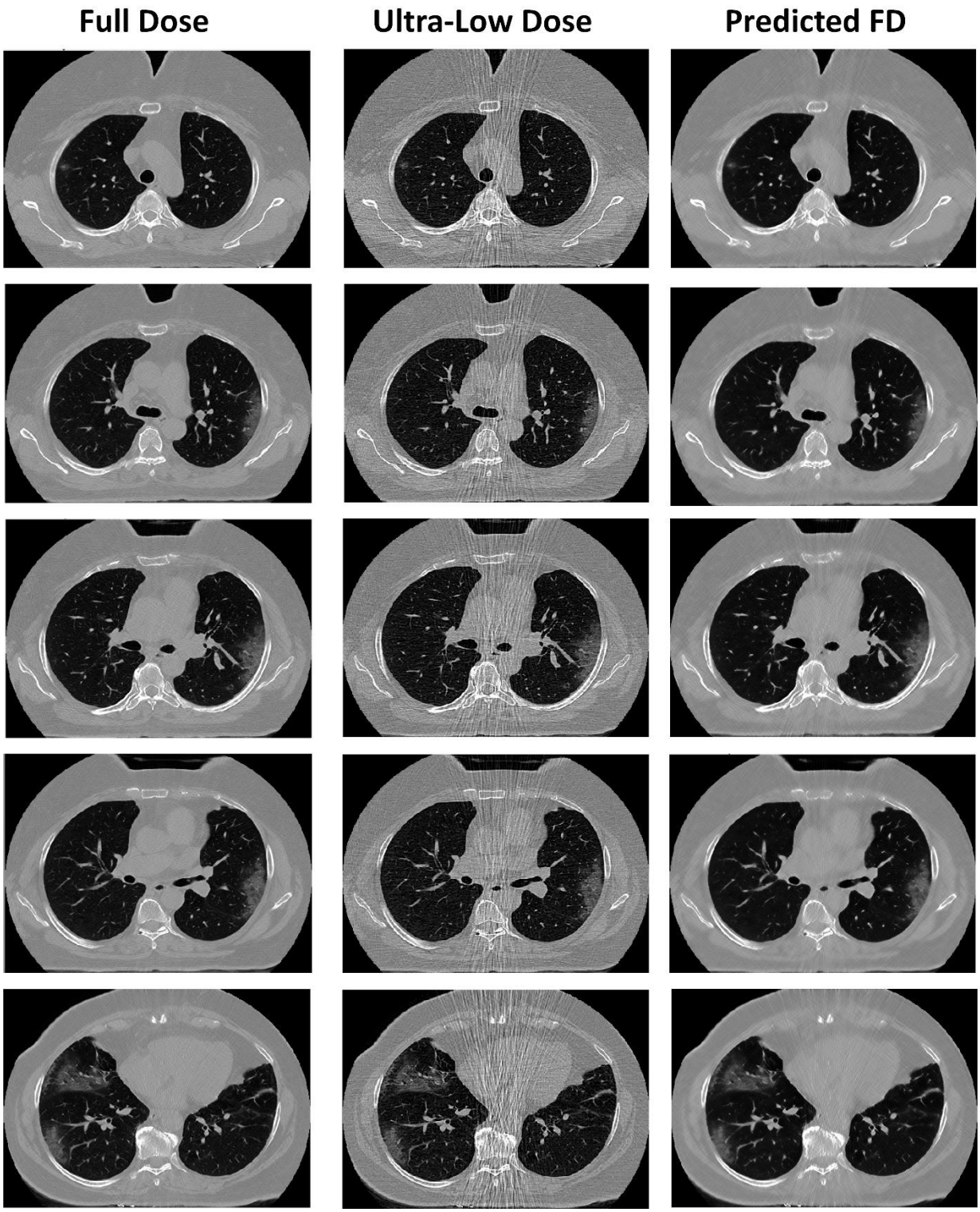
V. Conclusion

Ultra-low-dose CT imaging of COVID-19 patients would result in loss of critical information about lesion types. However, the results presented in this work indicated that ResNet is an optimal algorithm for generating ultra-low-dose CT images for COVID-19 diagnosis. Nevertheless, the deep learning solution failed to recover the correct lesion structure/density for a number of patients and as such, further research and development is warranted to address these limitations.

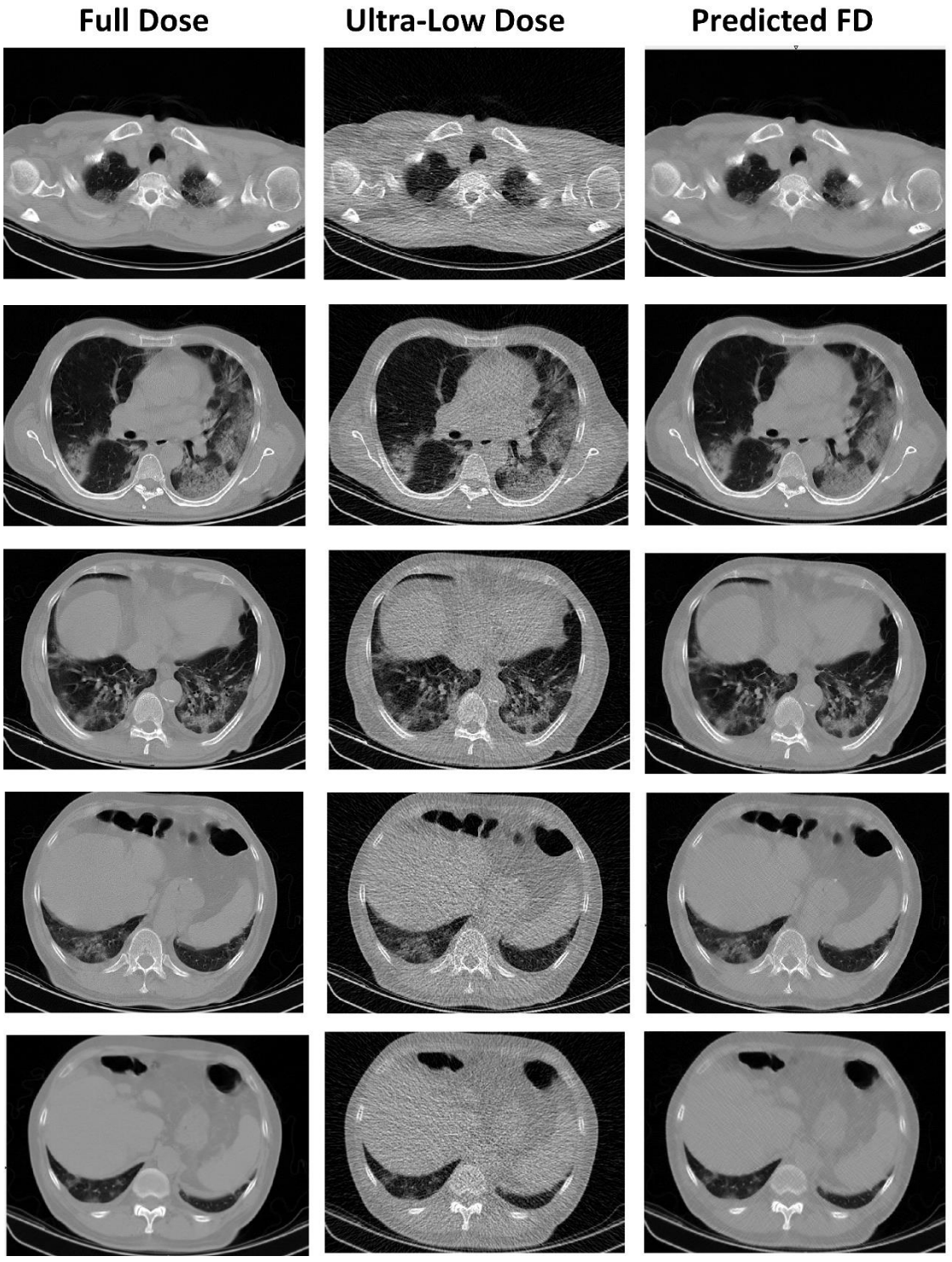
Acknowlegment

Open access funding provided by University of Geneva. This work was supported by the Swiss National Science Foundation under grant SNFN 320030_176052.

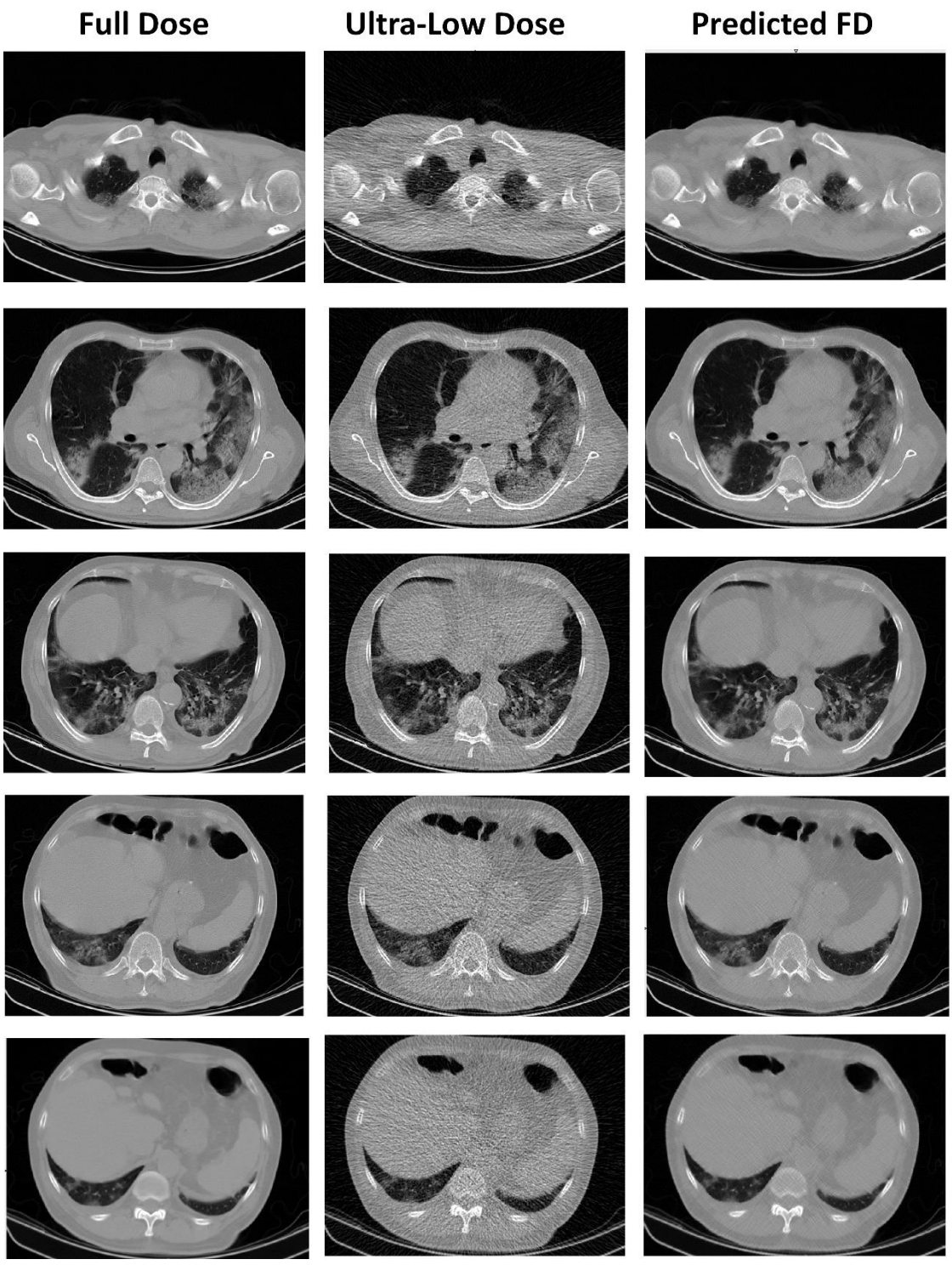
Supplemental material



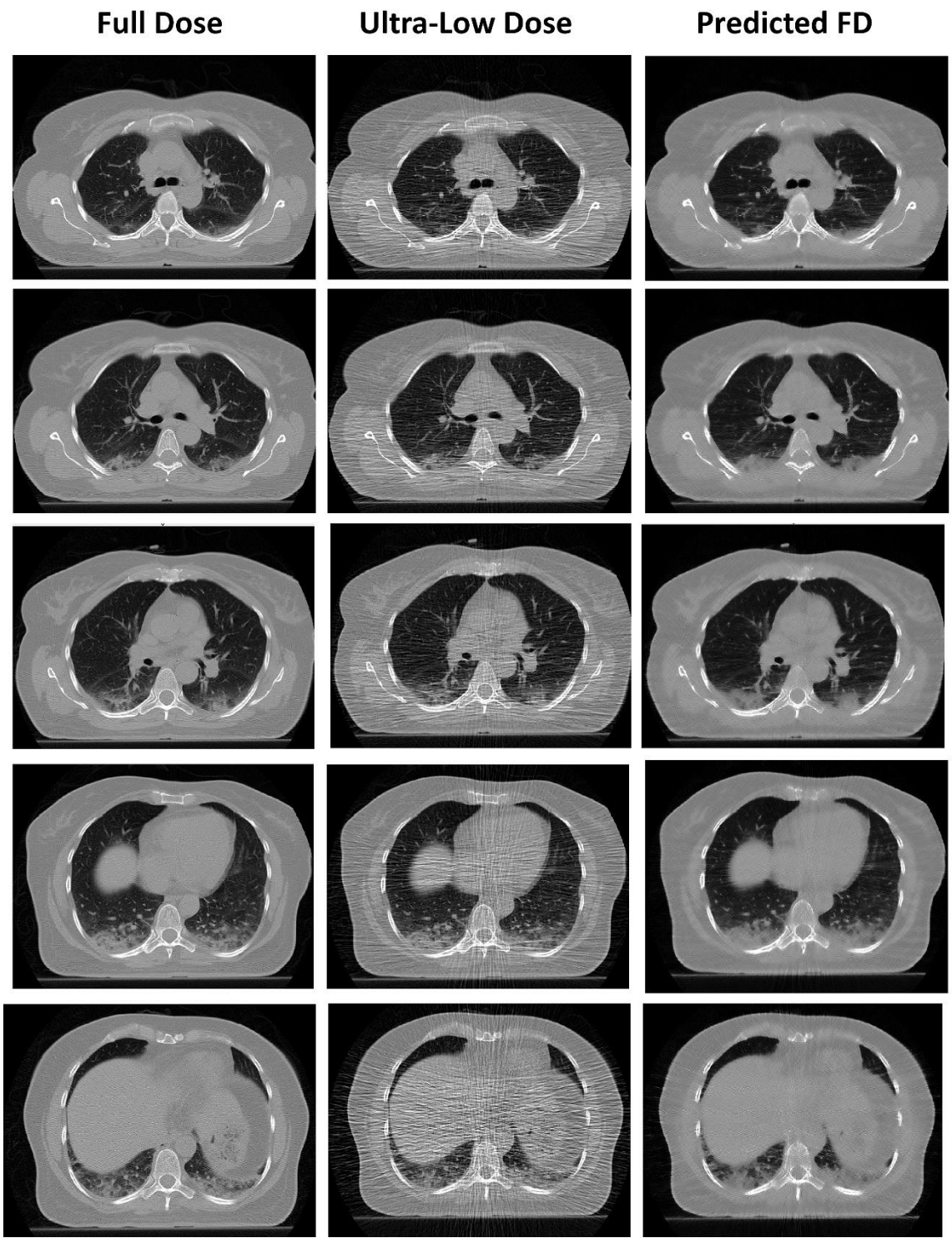
Supplemental Figure 1. Representative full-dose image and corresponding ultra-low-dose and predicted full-dose images case-2.



Supplemental Figure 2. Representative full-dose image and corresponding ultra-low-dose and predicted full-dose images case-3.



Supplemental Figure 3. Outlier report: CT images of a patient where the deep learning algorithm improved image quality but changed the patchy lesion to consolidation in predicted images case 2.



Supplemental Figure 4. Outlier report: CT images of a patient where the deep learning algorithm improved image quality but changed the patchy lesion to consolidation in predicted images case 3.

References

[1] Z. Wu and J. M. McGoogan, "Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: Summary of a report of 72314 cases from the Chinese Center for Disease Control and Prevention.,"*JAMA*, p. *in press*, 2020.

[2] C.-C. Lai, T.-P. Shih, W.-C. Ko, H.-J. Tang, and P.-R. Hsueh, "Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and corona virus disease-2019 (COVID-19): the epidemic and the challenges,"*International journal of antimicrobial agents*, p. 105924, 2020.

[3] Y.-W. Tang, J. E. Schmitz, D. H. Persing, and C. W. Stratton, "The laboratory diagnosis of COVID-19 infection: current issues and challenges,"*Journal of clinical microbiology*, 2020.

[4] H. Kim, H. Hong, and S. H. Yoon, "Diagnostic performance of CT and reverse transcriptase-polymerase chain reaction for coronavirus disease 2019: a meta-analysis,"*Radiology*, p. 201343, 2020.

[5] C. Bao, X. Liu, H. Zhang, Y. Li, and J. Liu, "COVID-19 Computed Tomography Findings: A Systematic Review and Meta-Analysis,"*Journal of the American College of Radiology*, 2020.

[6] H. Sakane *et al.*, "Biological Effects of Low-Dose Chest CT on Chromosomal DNA,"*Radiology*, vol. 295, no. 2, pp. 439-445, 2020.

[7] M. M. Rehani, T. P. Szczykutowicz, and H. Zaidi, "CT is still not a low-dose imaging modality,"*Med Phys*, vol. 47, no. 2, pp. 293-296, 2020.

[8] J. Hsieh, B. Nett, Z. Yu, K. Sauer, J.-B. Thibault, and C. A. Bouman, "Recent advances in CT image reconstruction,"*Current Radiology Reports*, vol. 1, no. 1, pp. 39-51, 2013.

[9] H. W. Goo, "CT radiation dose optimization and estimation: an update for radiologists,"*Korean journal of radiology*, vol. 13, no. 1, pp. 1-11, 2012.

[10] T. Zhao, M. McNitt-Gray, and D. Ruan, "A convolutional neural network for ultra-low-dose CT denoising and emphysema screening,"*Medical Physics*, vol. 46, no. 9, pp. 3941-3950, 2019.

[11] A. Radpour *et al.*, "COVID-19 evaluation by low-dose high resolution CT scans protocol.,"*Acad Radiol*, vol. 27, no. 6, p. 901, 2020.

[12] A. Agostini *et al.*, "Proposal of a low-dose, long-pitch, dual-source chest CT protocol on third-generation dual-source CT using a tin filter for spectral shaping at 100 kVp for CoronaVirus Disease 2019 (COVID-19) patients: a feasibility study.,"*Radiol Med*, vol. 125, no. 4, pp. 365-373, 2020.

[13] Z. Kang, X. Li, and S. Zhou, "Recommendation of low-dose CT in the detection and management of COVID-2019.,"*Eur Radiol*, p. *in press*, 2020.

[14] A. Dangis *et al.*, "Accuracy and reproducibility of low-dose submillisievert chest CT for the diagnosis of COVID-19,"*Radiology: Cardiothoracic Imaging*, vol. 2, no. 2, p. e200196, 2020.

[15] U. Baum *et al.*, "Improvement of image quality of multislice spiral CT scans of the head and neck region using a raw data-based multidimensional adaptive filtering (MAF) technique.,"*Eur Radiol*, vol. 14, no. 10, pp. 1873-1881, 2004.

[16] J. Wessling *et al.*, "The effect of dose reduction and feasibility of edge-preserving noise reduction on the detection of liver lesions using MSCT.,"*Eur Radiol*, vol. 17, no. 7, pp. 1885-1891, 2007.

[17] G. Litjens *et al.*, "A survey on deep learning in medical image analysis.,"*Med Image Anal*, vol. 42, pp. 60-88, 2017.

[18] B. Sahiner *et al.*, "Deep learning in medical imaging and radiation therapy.,"*Med Phys*, vol. 46, no. 1, pp. e1-e36, 2019.

[19] I. Shiri *et al.*, "Deep-JASC: joint attenuation and scatter correction in whole-body (18)F-FDG PET using a deep residual network,"*Eur J Nucl Med Mol Imaging*, 2020.

[20] A. Sanaat, H. Arabi, I. Mainta, V. Garibotto, and H. Zaidi, "Projection-space implementation of deep learning-guided low-dose brain PET imaging improves performance over implementation in image-space,"*J Nucl Med*, 2020.

[21] I. Shiri *et al.*, "Standard SPECT myocardial perfusion estimation from half-time acquisitions using deep convolutional residual neural networks,"*J Nucl Cardiol*, 2020.

[22] Q. Yang *et al.*, "Low-Dose CT Image Denoising Using a Generative Adversarial Network With Wasserstein Distance and Perceptual Loss,"*IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1348-1357, 2018.

[23] B. Kim, M. Han, H. Shim, and J. Baek, "A performance comparison of convolutional neural network-based image denoising methods: The effect of loss functions on low-dose CT images,"*Med Phys*, vol. 46, no. 9, pp. 3906-3923, 2019.

[24] Y. J. Shin *et al.*, "Low-Dose Abdominal CT Using a Deep Learning-Based Denoising Algorithm: A Comparison with CT Reconstructed with Filtered Back Projection or Iterative Reconstruction Algorithm,"*Korean J Radiol*, vol. 21, no. 3, pp. 356-364, 2020.

[25] M. Gholizadeh-Ansari, J. Alirezaie, and P. Babyn, "Deep learning for low-dose CT denoising using perceptual loss and edge detection layer.,"*Journal of Digital Imaging*, p. *in press*, 2020.

[26] O. Christianson, J. Winslow, D. P. Frush, and E. Samei, "Automated Technique to Measure Noise in Clinical CT Examinations,"*AJR Am J Roentgenol*, vol. 205, no. 1, pp. W93-9, 2015.
[27] D. Zeng *et al.*, "A simple low-dose x-ray CT simulation from high-dose scan,"*IEEE Trans. Nucl. Sci.*, vol. 62, no. 5, pp. 2226-2233, 2015.
[28] W. Li, G. Wang, L. Fidon, S. Ourselin, M. J. Cardoso, and T. Vercauteren, "On the Compactness, Efficiency, and Representation of 3D Convolutional Networks: Brain Parcellation as a Pretext Task," in *Information Processing in Medical Imaging*, Cham, M. Niethammer *et al.*, Eds., 2017// 2017: Springer International Publishing, pp. 348-360.
[29] A. K. Jones, M. K. O'Connor, and H. Zaidi, "The eventual rejection of the linear no-threshold theory will lead to a drastic reduction in the demand for diagnostic medical physics services.,"*Med Phys*, vol. 46, no. 8, pp. 3325-3328, 2019.
[30] D. J. Brenner and H. Hricak, "Radiation exposure from medical imaging: time to regulate?,"*JAMA*, vol. 304, no. 2, pp. 208-209, 2010.
[31] H. Abdollahi *et al.*, "Radiation dose in cardiac SPECT/CT: An estimation of SSDE and effective dose,"*Eur J Radiol*, vol. 85, no. 12, pp. 2257-2261, 2016.
[32] NAS/NRC (National Academy of Sciences/National Research Council), "Health risks from exposure to low levels of ionizing radiation: BEIR VII – Phase 2," in "BEIR VII phase," National Research Council, Washington DC, 2006, vol. 2. [Online]. Available: http://www.nap.edu/catalog.php?record_id=11340
[33] ICRP, "The 2007 Recommendations of the International Commission on Radiological Protection. ICRP publication 103,"*Annals of the ICRP*, vol. 37, no. 2-4, pp. 1-332, 2007.

Chapter 8

Real-time, acquisition parameter-free, voxelwise patient-specific Monte Carlo dose construction in total-body CT scanning using deep neural networks

Yazdan Salimi, Azadeh Akhavanallaf, Isaac Shiri, Hossein Arabi, and Habib Zaidi

Contribution: Study conception and design, methodological development, computer programming, analysis and interpretation of results, manuscript preparation.

Eur Radiol (2022) submitted

Abstract

Purpose: Computed tomography (CT) is among the most widely used medical imaging modalities in clinical setting. Estimating patient-specific radiation dose and associated radiation risks is critical in optimization procedures. We propose a deep learning-guided approach to generate voxel-based dose maps from whole-body CT acquisitions.

Methods: Sixty-three whole-body CT images were converted to density maps with 5 mm^3 isotropic voxel size. The voxel-wise dose maps corresponding to each source position/angle were calculated using Monte Carlo (MC) simulations considering patient- and scanner-specific characteristics (SP_MC). The dose distribution in a uniform cylinder was computed through MC calculations (SP_uniform). The density map and SP_uniform dose maps were fed into a residual deep neural network (DNN) to predict SP_MC through as an image regression task. The whole-body dose maps reconstructed by the DNN and MC were compared in the 11 test cases scanned with two tube voltages through transfer learning with/without tube current modulation (TCM). The voxel-wise and organ-wise dose evaluations, such as mean error (ME, mGy), mean absolute error (MAE, mGy), relative error (RE, %), and relative absolute error (RAE, %), were performed.

Results: The model performance for the 120 kVp and TCM test set in terms of ME, MAE, RE, and RAE voxel-wise parameters was -0.0302 ± 0.0244 mGy, 0.0854 ± 0.0279 mGy, -1.13 ± 1.41 %, and 7.17 ± 0.44 %, respectively. The organ-wise errors for 120 kVp and TCM scenario averaged over all segmented organs in terms of ME, MAE, RE, and RAE were -0.144 ± 0.342 mGy, and 0.23 ± 0.28 mGy, -1.11 ± 2.90 %, 2.34 ± 2.03 %, respectively. The performance of our model was almost similar in the external test set, considering the fixed tube current and tube current modulation scenarios.

Conclusion: Our proposed deep learning model is able to generate voxel-level dose maps from a whole-body CT scan with reasonable accuracy suitable for organ-level dose estimation. Through the generation of a dose distribution from a single source position, our model can generate accurate and personalized dose maps in few seconds for a wide range of acquisition parameters.

Abbreviations

CT: Computed Tomography
TCM: Tube Current Modulation
DL: Deep Learning
ICRP: International Commission on Radiological Protection
MC: Monte Carlo
SP_MC: Single Point Monte Carlo Dose
SP_DL: Single Point Deep Learning Dose
SP_uniform: Single Point Monte Carlo Dose in a Uniform Material
WBCT: Whole-Body Computed Tomography
WBDM: Whole-body Dose Map
ME: Mean Error
MAE: Mean Absolute Error
RE: Relative Error
RAE: Relative Absolute Error

I. Introduction

The capability of visualizing inside the human body through non-invasive medical imaging examinations is a tremendous opportunity to diagnose various pathologies. X-ray Computed Tomography (CT) is one of the prevalent imaging modalities used in the initial clinical diagnosis, follow-up, staging, radiation therapy planning and in emergency departments to provide valuable information for a wide range of indications [1]. In addition, CT is also commonly attached to nuclear medicine instrumentation, such as single-photon emission computed tomography (SPECT) or positron emission tomography (PET), for concurrent SPECT/CT [2] or PET/CT [3] imaging on hybrid imaging devices. At the same time, CT, one of the high-dose examinations, is responsible for a significant part of ionizing radiation exposure of patients [4, 5]. The International Commission on Radiological Protection (ICRP) [6] suggested estimating the radiation dose delivered to patients from medical imaging procedures toward the optimization rule known as ALARA in order to minimize the risks through the appropriate use of ionizing radiation.

The recent emphasis on personalized medicine and patient-specific justification/optimization substantiates the critical demand to calculate specific parameters related to radiation risks [7-10]. The organ dose is a requirement for patient-specific dose calculation and has a good correlation with radiation risks [9]. On the other hand, it has been shown that the radiation dose delivered to specific organs can reach the deterministic dose levels, especially in serial CT examinations, which is common practice in patients follow-up, e.g., in the recent Covid-19 pandemic [11-13].

The estimation of organ doses can be performed using multiple methodologies. The most straightforward approach uses conversion factors specific to the scanning protocols. An alternative option is to use dedicated software tools, such as ImpactDose[1] and Radimetrics [14]. Both above mentioned software packages proved to have a low correlation with organ doses calculated by more accurate Monte Carlo (MC) simulation tools using patient-specific or reference computational models [15, 16], particularly when the Tube Current Modulation (TCM) system is activated [17-19]. While MC calculations using patient-specific computational models is deemed to be the most accurate approach and is often regarded as the gold standard technique, its downsides, including computational time, high computational burden, and required expertise in computer programming, limit its adoption in clinical setting. Exploiting the parallel computational power of GPUs enabled MC calculations to be faster and more suited for adoption in clinical setting [20, 21]. Yet, the complexity associated with the technique remains a significant hurdle. Deep learning-based algorithms are currently used in various medical imaging applications, including image regression [22], registration [23], segmentation [24], radiation dosimetry calculation [25, 26], and optimization [27, 28]. This study aimed to develop a fully automated method to estimate patient-specific MC-based dose maps associated with whole-body (WB) CT examinations in real-time using deep neural network algorithms.

II. Materials and methods

Study population

This study included 63 patients (35 male and 28 females) who underwent whole-body PET/CT imaging on a Biograph mCT scanner (Siemens Healthineers, Erlangen, Germany). All CT scans were performed in helical mode using 120 kVp tube potential, and Siemens CareDose4D TCM was activated. Images were reconstructed with the extended 70 cm field-of-view option, voxel size equal to 1.523 mm in the axial plane, and 1.4 mm slice thickness using filtered-back projection algorithm. Figure 1 shows the flowchart of the different steps followed in this study protocol.

[1] https://impactdose.software.informer.com/

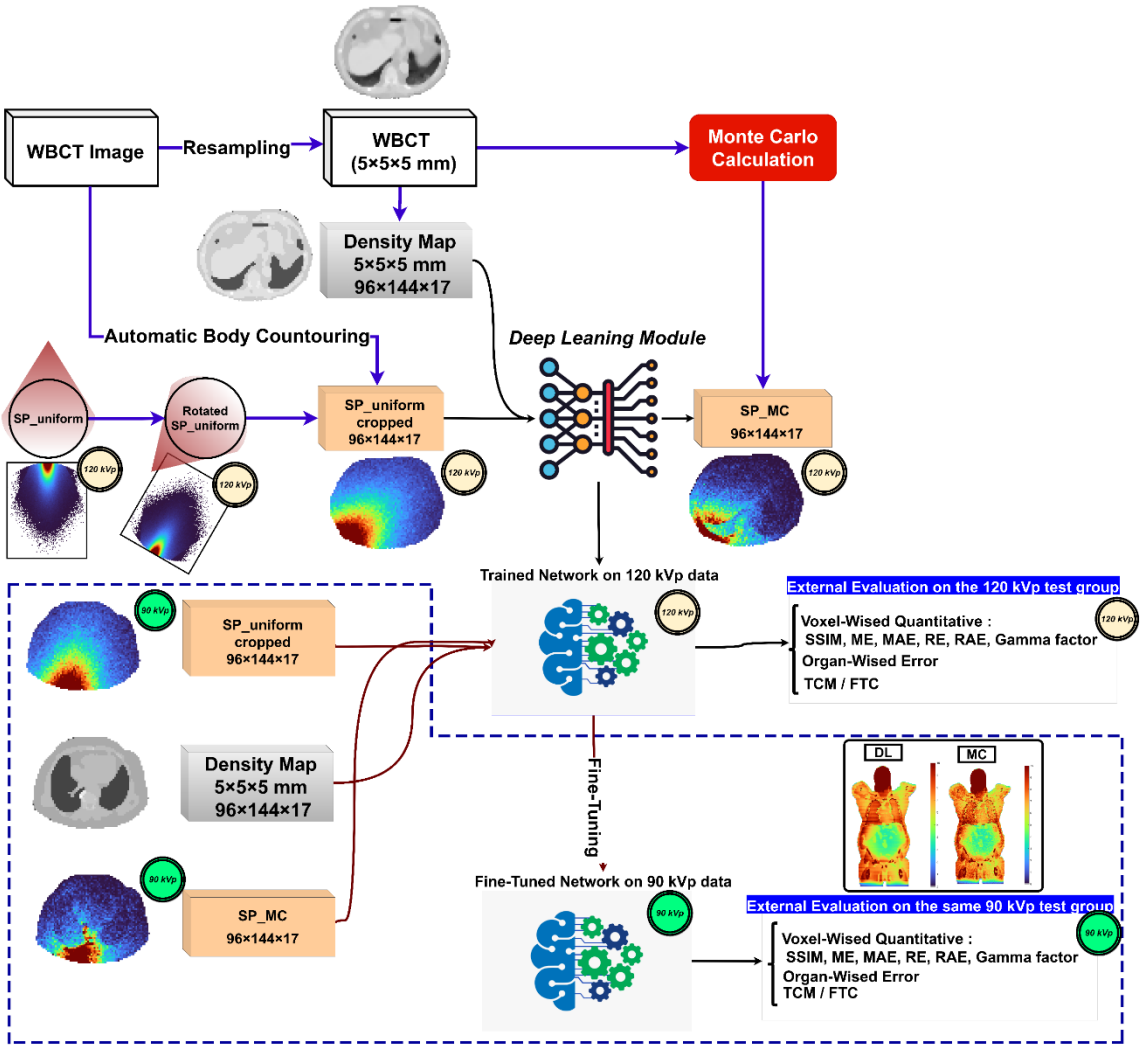


Figure 1. Flowchart summarizing the different steps involved in the implementation of the whole process. The blue dashed line shows the 90 kVp generalizability test. DL: Deep learning. MC: Monte Carlo.

Monte Carlo simulations

CT HU values were converted to density maps using linear multi-regression models for the segmentation of CT images into different tissue densities, as proposed by Schneider et al. [29]. Subsequently, the resulting density maps were resampled to 5 mm^3 cubic isotropic voxel size. The essential components incorporated into MC simulations, including accurate source model and protocol-related parameters, were adopted from our in-house MC simulation code developed and validated in a previous study [30]. The acquisition parameters, including tube voltage, collimation width, table speed, rotation time, pitch, and tube current modulation, were implemented in this simulation. This simulator is based on the MCNPX general-purpose Monte Carlo radiation transport code (version 2.6) [31].

The output of MC simulations is a 3D dose map for a single source position (SP_MC) with limited axial coverage. Monte Carlo simulations were run for multiple discrete source positions to simulate helical whole-body CT scanning. A random starting location was generated for the source owing to the lack of information about the tube start angle in the DICOM header. Accordingly, a spiral motion of the source position in 2 mm axial intervals along the Z-axis (craniocaudal axis) was modeled. Finally, considering the longitudinal tube current modulation (extracted from the DICOM header for TCM), simulated dose maps for each source position was multiplied by the corresponding tube current and were superimposed to construct the complete voxel dose distribution.

Data preparation

MC calculations were performed for a total number of 63 patients with 120 kVp tube voltage. Then, by keeping all parameters similar, except kVp, MC calculations were repeated with 90 kVp tube voltage for patients in the test group (11 cases) plus 20 cases randomly selected from the train and validation groups to perform the fine-tuning process described later in the text. The cases from the train and validation were used for performing transfer learning and fine-tuning.

Monte Carlo calculation of radiation dose in a uniform cylinder at 90 and 120 kVp

A uniform water-filled cylinder with a 715 mm diameter located within the CT gantry was simulated, and the dose map for a single source position (zero degrees, located at the anterior point) was calculated for a large number of simulated events (4×10^{10} particles) tracked by the MC simulator. This dose map, referred to as the single-source position uniform map (SP_uniform), was calculated for two tube voltages, namely 90 and 120 kVp for a single source. It should be mentioned that the 90 kVp uniform dose maps were used for testing the network generalizability through fine-tuning.

Generation of single-source position images and corresponding density maps

The body contour was automatically segmented on all CT images utilizing analytical image processing methods. All body contour segmentations were reviewed and confirmed visually. The MC output images (SP_MC) having a size of 96×144×17 voxels were saved, and the density map for the same axial coverage range cropped to the same size. The SP_uniform images were cropped to the same axial coverage body contour and normalized to a conversion factor (CF) calculated by Eq. (1) to compensate for the effect of attenuation taking place in the SP_uniform dose calculation on the large cylinder.

$$\mathrm{CF} = e^{(d_{SP\_MC} - d_{SP\_uniform})}. \qquad \text{Eq. (1)}$$

where e is Euler's number, d_{SP_MC} is the distance from the edge of the body contour to the x-ray tube source in a specific source position. $d_{SP\_uniform}$ is the distance from the edge of a large cylinder simulated to the source in a specific position. Since the cylinder size was larger than the size of our largest patient, the CF was always greater than 1. Figure 2 shows the examples of SP_MC, SP_Uniform, and the corresponding CT slices, when the source is in the right lateral position. The two images of SP_uniform and SP_MC were normalized by all voxel intensities by a fixed value. Each source position was saved in a separate image and used for training the neural network.

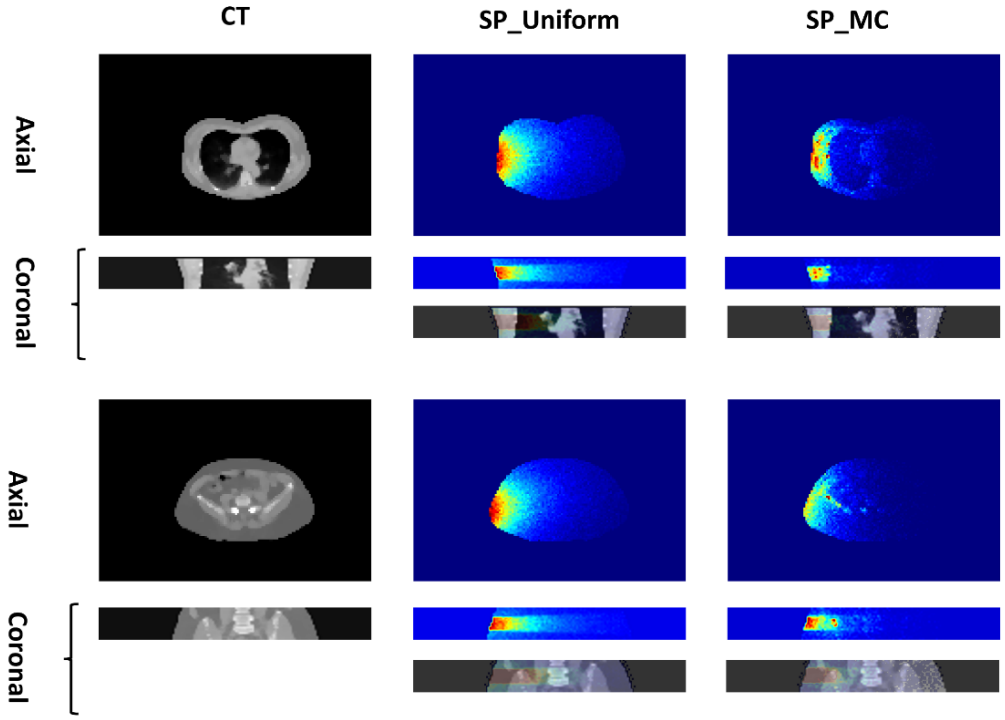


Figure 2. Examples of axial and coronal slices of CT, SP_Uniform, and SP_MC dose maps corresponding to a single source position/angle. In these cases, the x-ray tube is in the right lateral position.

Network architecture and training details

From all 63 WB CT images (27,632 source positions), 11 cases (4792 source positions) were used as the untouched test set. Figure 1 shows the steps performed in this study and examples of mentioned three images of SP_uniform, SP_MC and density maps. The SP_unifrom in a unique source position/angle and the density map images were fed as input to the neural network to predict the SP_MC image as the output in the corresponding source position/angle. A deep residual network (ResNet) was trained in Python (TensorFlow) to generate the SP_MC images from the two mentioned inputs. The ResNet is composed of 20 convolutional layers (19 layers with kernel size 3×3×3 and the last layer with kernel size 1×1×1) where the image size is kept constant through the different layers (no down or up pooling was applied). Different feature levels, including low, medium, and high, were extracted by using dilation of 0 (first seven layers), 2 (six middle layers) and 4 (six last layers), respectively, in a convolutional kernel. Every two layers were connected using a residual connection to avoid gradient vanishing or exploding. The training was continued for 100 epochs using the "Adam" optimizer and L2 loss function. The initial learning rate of 10^{-3} was reduced in a piecewise method every five epochs. The trained network was tested on the external group datasets, and the deep neural network output was named SP_DL.

Generalizability evaluation on 90 kVp data (Fine-tuning)

To test the generalizability of the proposed model for kVps other than 120 kVp, we performed MC simulations to calculate the voxel dose maps by considering the 90 kVp spectrum on the untouched test group (11 cases) and 20 patients selected from the train and validation group. The same pre-processing steps mentioned earlier were followed to derive SP_MC, SP_uniform, and density maps at 90 kVp. SP_uniform and density maps were fed to the trained network on 120 kVp data, and the process of fine-tuning continued for 50 epochs by body fine-tuning approach, i.e., the weights from the trained network on 120 kVp data served as the initial weights for 90 kVp fine-tuning. SP_uniform and density maps were fed to the fine-tuned network on 90 kVp training datasets, and SP_DL images at 90 kVp were generated for the same test group (11 cases). These SP_DL images were compared to SP_MC images at 90 kVp.

Dose map reconstruction from single source positions

The dose maps from the single source position were corrected by factors related to the tube calibration described in a previous study [30]. The tube current was extracted from the DICOM header. Then, the dose maps corresponding to a single source position/angle were superimposed to reconstruct the whole-body dose maps (WBDM) using both SP_MC and SP_DL dose maps, referred to as WBDM_MC and WBDM_DL, respectively. The final WBDM was a matrix of 96×144×Z voxels, where Z is the image size along the Z-axis, and the voxel value is the absorbed dose in that voxel in units of milli-gray (mGy). We have considered two strategies for WBDM calculation, fixed 100mA tube current (FTC) and TCM activated according to the actual tube current recorded from the DICOM images.

Evaluation metrics

1) Voxel-wise quantitative dose evaluation

The WBDM_DL images were compared with WBDM_MC images serving as the standard of reference (ground truth) at the voxel level. Voxel-wise parameters, including structural similarity index (SSIM), mean error (ME, mGy), mean absolute error (MAE, mGy), relative error (RE, %), relative absolute error (RAE, %), and gamma pass rate were calculated. Gamma pass rate, as described earlier by Low et al. [32] with 1 mm and 1% criterion, was considered.

2) Organ-level dose evaluation

In addition to voxel-wise errors, eight organs, including the Liver, Heart, Bone, Kidneys (both), Spleen, Bladder, Lungs (both), and brain, were segmented manually on the test WBCT images. The organ doses were estimated by calculating the mean voxel value on WBDM images inside the organ segmentations. The organ absorbed doses calculated on WBDM_DL and WBDM_MC were compared for each organ in terms of mean error (ME, mGy), mean absolute error (MAE, mGy), relative error (RE, %), and relative absolute error (RAE, %). These voxel-wise and organ-wised metrics were calculated for both 90 kVp and 120 kVp external datasets, considering both FTC and TCM scenarios.

3) Statistical analysis

The Kolmogorov-Smirnov test was used to check the normality of distributions. The mentioned organ-wise evaluation metrics were compared between the two groups of 90 and 120 kVps using the Mann-Whitney test. P-values less than 0.05 were considered statistically significant.

III. Results

The age of included patients was 58.9 ± 17.2 years. The average patients' water equivalent diameter was 26.6±2.7 (range 16.45 - 32.95) cm. The average tube current implemented by TCM was 140.7 ±48.71 (56 to 306) mA. Table 1 summarizes the demographic information of patients.

Table 1. Demographic description of the test and train & validation groups

metric	Train & Validation	Test
sex	29 male, 23 female	6 male, 5 female
age	60.1 ± 16.9	53.2 ± 17.9
kVp	120	120
Pitch	0.8	0.8
CTDIvol	5.74±2.70	8.33 ± 4.00
Patient Height	169 ± 12	167 ± 12
Patient Weight	75.1 ± 15.8	76.4 ± 17.4
Tube Current	135.5 ± 45.4	167.6 ± 67.8

Voxel-wise error metrics

Table 2 summarizes the results of voxel-wise metrics for two external validation groups acquired at 90 and 120 kVp. The model performance for the 120 kVp and TCM test set in terms of voxel-wise parameters, including SSIM, PSNR, Gamma, ME, MAE, RE, and RAE, was 0.997 ± 0.002, 46.69 ± 1.98, 98.47 ± 0.81, -0.0359 ± 0.0244 mGy, 0.1091 ± 0.0279 mGy, -1.16 ± 1.41 %, and 7.13 ± 0.44 %, respectively. All voxel-wise parameters were in the same range for 120 kVp, TCM and FTC test sets. The voxel-wise evaluation results after performing transfer learning and fine-tuning on 90 kVp data were also comparable to 120 kVp, except RAE, which was almost 1.5% higher in the 90 kVp test group compared with 120 kVp results (8.63 vs. 7.17). Considering the FTC and TCM scenarios, the performance of our model was almost similar in the 90 kVp test set.

Table 2. Voxel-wise metrics for two external validation groups acquired at 90 and 120 kVp.

	120 kVp		**90 kVp**	
Metrics	FTC	TCM	FTC	TCM
SSIM	0.997 ± 0.002 (0.993 to 0.998)	0.997 ± 0.002 (0.993 to 0.998)	0.994 ± 0.005 (0.981 to 0.998)	0.994 ± 0.005 (0.981 to 0.998)
PSNR	46.69 ± 1.98 (44.95 to 50.17)	47.68 ± 1.98 (44.95 to 50.17)	45.11 ± 3.85 (37.51 to 48.77)	46.18 ±5.08 (37.48 to 51.66)
Gamma Value	98.47 ± 0.81 (96.73 to 99.72)	98.91 ±0.81 (96.73 to 99.72)	98.26 ± 1.29 (95.28 to 99.08)	98.64 ± 1.41 (95.28 to 99.68)
ME (mGy)	-0.0359 ± 0.0244 (-0.0826 to 0.0025)	-0.0302 ± 0.0244 (-0.0826 to 0.0025)	-0.0167 ± 0.0149 (-0.0372 to 0.0161)	-0.0126 ± 0.0124 (0.0326 to 0.0133)
MAE (mGy)	0.1091 ± 0.0279 (0.0513 to 0.1401)	0.0854 ± 0.0279 (0.0513 to 0.1401)	0.1088 ± 0.0308 (0.0776 to 0.1626)	0.0892 ± 0.0462 (0.0471 to 0.1713)
RE (%)	-1.16 ± 1.41 (-3.72 to 1.39)	-1.13 ± 1.41 (-3.72 to 1.39)	0.27 ± 1.33 (-1.99 to 2.00)	0.28 ± 1.33 (-2.00 to 1.98)
RAE (%)	7.13 ± 0.44 (6.57 to 7.89)	7.17 ± 0.44 (6.57 to 7.89)	8.58 ± 1.83 (6.15 to 10.80)	8.63 ± 1.82 (6.19 to 10.82)

Figure 3 shows the joint histogram comparing WBDM_DL and WBDM_MC doses at the voxel level. The high correlation depicted in figure 3 (R^2>0.999) indicates the excellent agreement between MC and DL results.

Figure 4 shows two examples of WBDM_DL and WBDM_MC and their corresponding bias maps displayed in a coronal view for a combination of two kVps (90 and 120) and two TCM and FTC scenarios. The bias map shows excellent agreement between MC and DL results. The highest differences in terms of RAE (%) are depicted in the lung/chest wall interval and soft tissue/skull (bony tissue), where there is a gradient in density and chemical composition characteristics of biological tissues and, consequently, radiation interaction properties with tissues. The average RAE for all organs was always less than 4.5 % for both kVps and TCM and FTC scenarios.

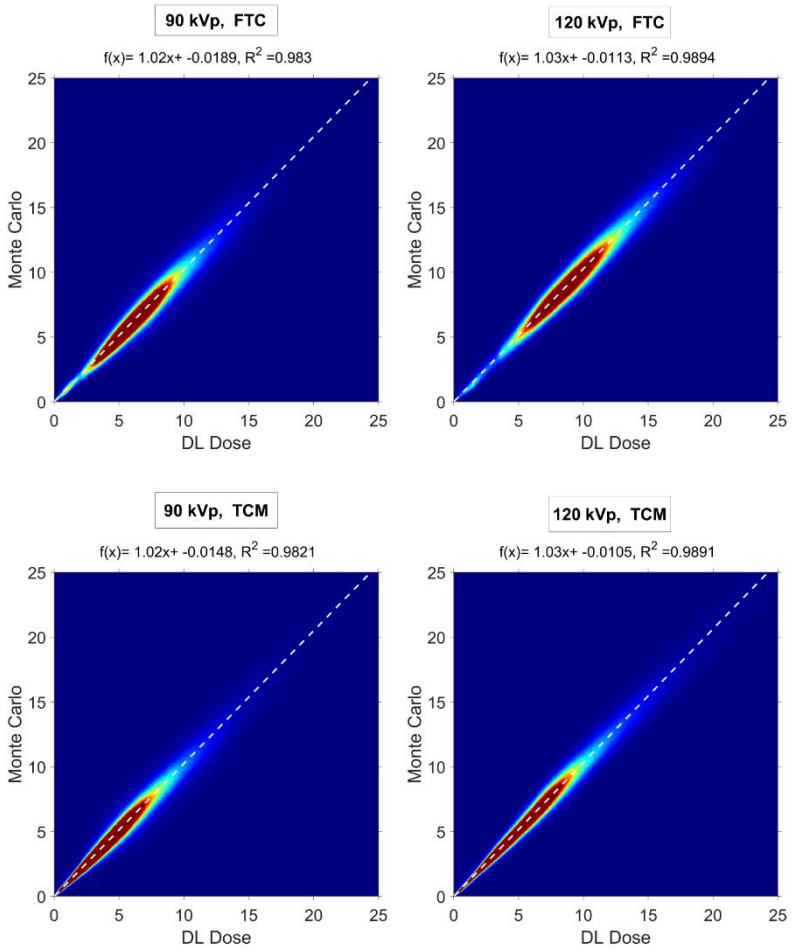


Figure 3. Joint-histograms comparing the voxel-wise doses of WBDM_DL and WBDM_MC at 90 kVp, FTC (top left), 120 kVp, FTC (top right), 90 kVp, TCM (bottom left), and 120 kVp, TCM (bottom right). The white dashed line shows the fitted line and the regression line equation. The correlation coefficient (R^2) is also shown for each histogram.

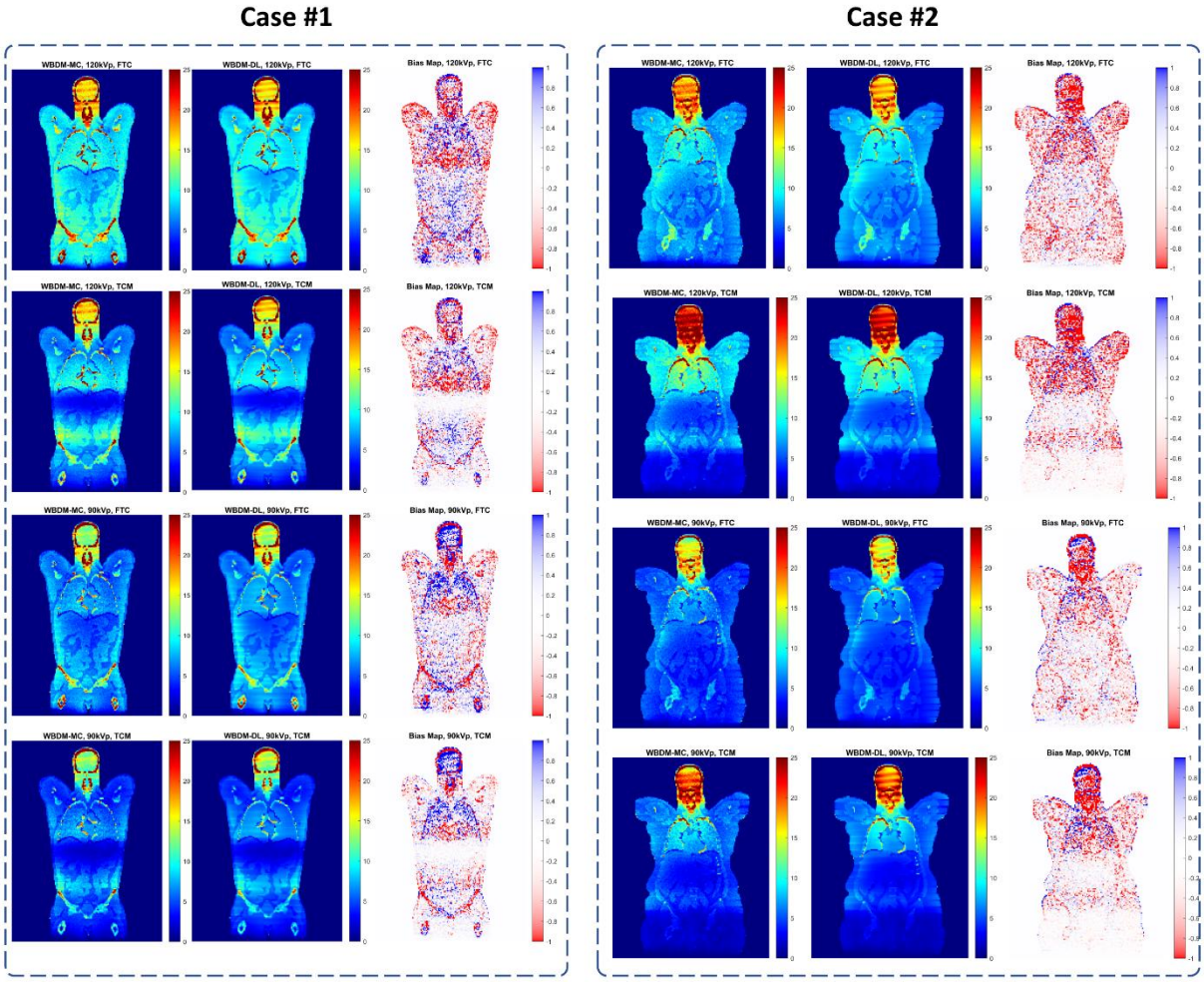


Figure 4. Coronal slices of WBDM_DL, WBDM_MC and the corresponding bias maps for two cases from the external test sets. The caption for each study displays the kVp and tube current scenario (TCM or FTC). Case #1: 74 y/o male, patient height = 172 cm, patient weight = 85 kg, average water equivalent diameter [1] = 28.4 cm, the effective diameter at the largest slice = 32.9 cm. case #2: 65 y/o female, patient height = 158 cm, patient weight = 87 kg, water equivalent diameter = 29.5 cm, the effective diameter at the largest slice = 35.6 cm. The voxel value here is mGy, and the color bar is shown beside each image.

Organ-wise error metrics

The organ-wise error for 120 kVp and TCM scenario averaged over all segmented organs in terms of RE (%), RAE (%), ME (mGy), and MAE (mGy) was -1.11 ± 2.90, 2.34 ± 2.03, -0.144 ± 0.342, and 0.23 ± 0.28 respectively. Supplemental-Table 3 and 4 summarize organ-wise metrics calculated on 120 kVp and 90 kVp test sets, respectively. There was no statistically significant difference between the metrics calculated in organ-wise evaluations between the FTC and TCM performance in either the 90 and 120 kVp test sets (Mann-Whitney, $p >> 0.05$). The highest average errors were observed in the heart, bone, and brain regions, where there is a higher gradient in density and surrounding tissues. Figure 5 compares organ doses measured on DL and MC reconstructed dose maps. The violin plots show overall good agreement between the distributions of DL and MC organ doses. Figure 6 shows the boxplot of RE and RAE (%) between the calculated organ doses.

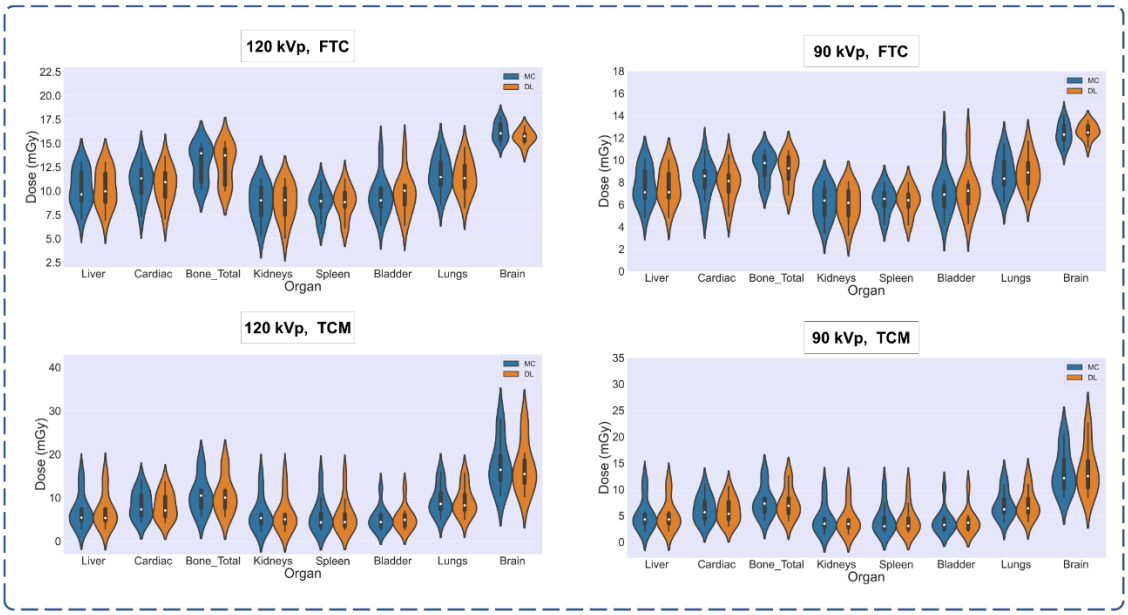


Figure 5. Violin plots of organ dose distributions calculated by MC (orange) and DL (blue) at 90 and 120 kVp for both FTC and TCM scenarios.

IV. Discussion

In this work, we proposed a novel method for dose map calculation using deep neural networks through two input channels. The model estimates the radiation voxel dose map by combining the attenuation and source angle/position information from the SP_uniform image with the attenuation characteristics from the density map image (Figure 2). This model predicts the dose distribution corresponding to a single source position/angle around the patient's body, which can be an excellent option to calculate the absorbed doses with a lower interval in the source position movements, which proved to be more realistic [30]. Other acquisition parameters, such as pitch, scan mode (spiral, sequential), rotation time, and other parameters, such as tube current, could be modeled by providing single-source position dose maps. This capability of calculating single source position and angle enables the calculation of whole-body dose maps in more complicated acquisition settings, such as organ-based TCM algorithms where the tube current is reduced in anterior arcs. The model's generalizability was examined through transfer learning to a different kVp dataset and fine-tuning the model. The results in terms of organ-wise dose metrics demonstrated the robustness of the developed model. Our model’s performance was similar when considering FTC and TCM techniques in both voxel-wise and organ-wise metrics. Besides, the dose map calculation is feasible for an acquisition performed using dual-source CT scanners or single source dual-energy mode by considering the source from each kVp (x-ray tube) as a single source position/angle.

The patients included in the training and test datasets covered a wide range of body shapes and BMIs. As shown in figure 4, the model is robust against patient size and composition changes. Wang *et al.* [33] proposed analytic linear Boltzmann modeling of the radiation dose in an anthropometric phantom. They reported errors of less than 3 %, but their model was specific for a single phantom and didn't consider variability in the human body. Although performed independently, our study bears some similarities with the study published by Maier *et al.* [25] in the sense that we used two channel inputs to our model to predict the voxel dose maps. We used whole-body CT images covering a larger axial field-of-view range from the skull to mid-thigh and trained a single general model for the full coverage. The single general model applicable to all scan protocols is easier to implement. They introduced multiple models by changing the parameters, while the generalizability is more practical in daily clinical routine by accessing single source dose maps. Besides, our proposed methodology is capable of reconstructing the dose maps directly from CT images without additional time-consuming deterministic methods for solving the Boltzmann transport equation.

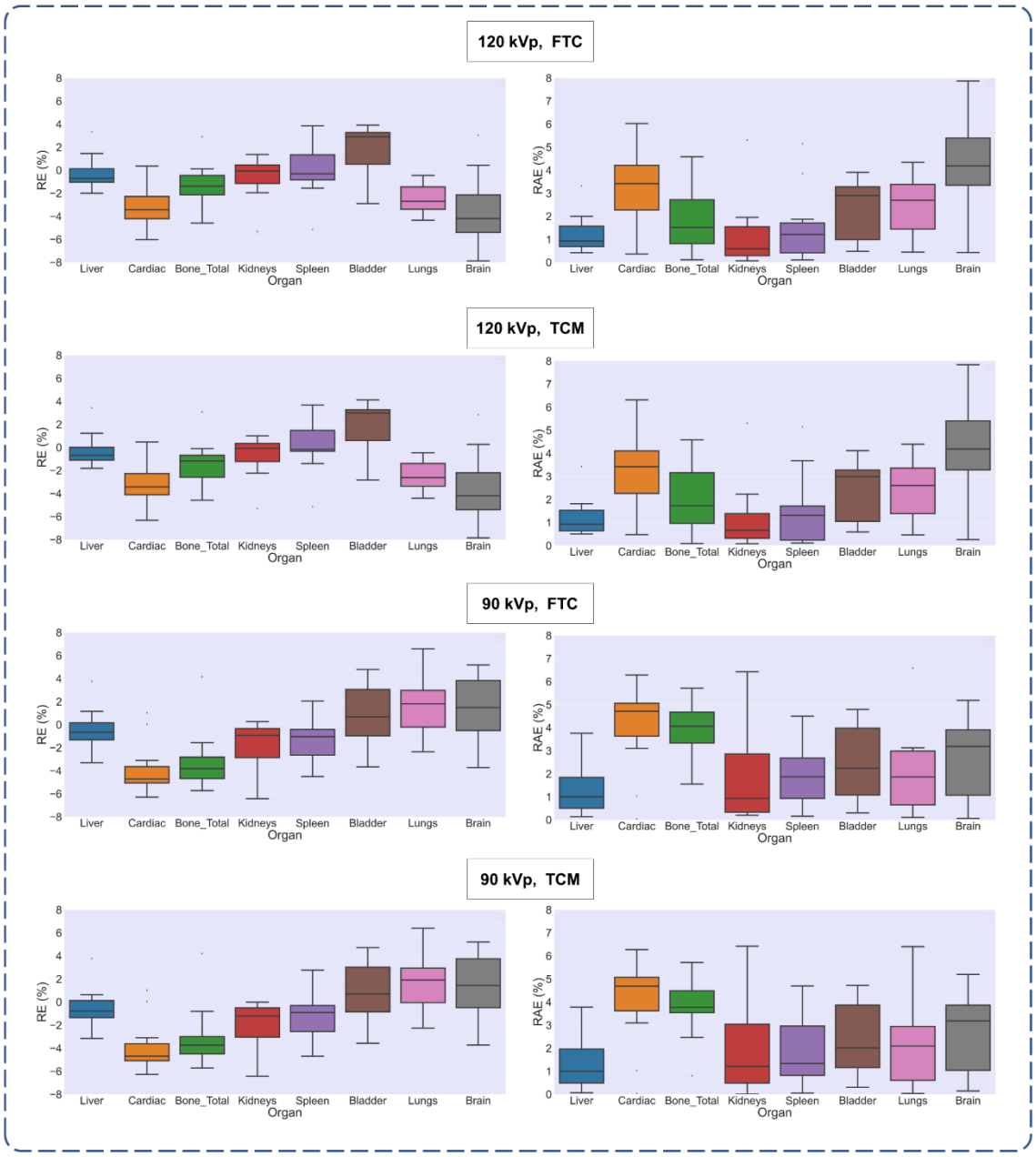


Figure 6. Boxplots of the RE and RAE (%) comparing organ doses calculated by DL and MC methods for eight segmented organs and a combination of 90 and 120 kVps and TCM/FTC scenarios.

Tzanis *et al.* [26] used DL to generate voxel dose maps calculated by Monte Carlo simulations by converting the image into a long vector and introducing scan parameters, such as tube current and scan mode, as additional columns. They included 343 head & neck scans and reported organ doses delivered to three tissues/organs, including the brain, cranial bones, and eye lens, with average errors less than 6% (range 0 -19%) in terms of organ RAE. Our proposed model provides more accurate results in terms of organ-wise RAE (average 2.74). Besides, they only used 120 kVp acquisitions and a scan range limited to the head & neck region.

Organ masks is a critical requirement for calculating organ doses. We segmented multiple organs to evaluate the performance of our model in organ-level dose calculation. The labor-intensive and time-consuming segmentations are important limitations of using dose maps in radiation risk estimation. Despite the presence of voxels with a higher error than the average in terms of voxel-wise RAE (%), the organ dose errors were negligible, especially for large organs, such as the liver. The slightly higher error in some organ doses and voxels could be attributed to methodological limitations, such as coarse image matrix size (voxel size of 5 mm), that we adopted to reduce the computational time. The large voxel size can also cause higher errors in voxel-wise metrics. The excellent performance achieved by our model in organ-level doses is much better than pre-tabulated software

outputs. Moreover, we used only a single scanner to train our DL network using a limited number of patients. Still, the main bottleneck was the high computational time required to generate the Monte Carlo dose maps as ground truth. In addition, in our method, the radiation dose delivered to organs out of the scan reconstruction axial range is missing.

V. Conclusion

Our proposed deep learning model can generate whole-body dose maps from a CT scan acquisition with reasonable accuracy at the voxel level and excellent performance at organ-level dose estimation. The whole process, including pre-processing and model inference on a new dataset, can be performed within seconds, which makes personalized dosimetry with an acceptable accuracy a possible option in clinical routine. Conversely, by generating a dose distribution from a single source position, our model can generate accurate and personalized dose maps for a wide range of acquisition parameters.

Acknowledgments

This work was supported by the Euratom research and training programme 2019-2020 Sinfonia project under grant agreement No 945196.

Supplemental material

Supplemental-Table 3. the results of the organ-wised evaluation of the results for the 120 kVp dataset.

	FTC				TCM			
Organ	RE %	RAE %	ME (mGy)	MAE (mGy)	RE %	RAE %	ME (mGy)	MAE (mGy)
Liver	-0.27 ± 1.54 (-2 to 3.32)	1.26 ± 0.84 (0.42 to 3.32)	-0.023 ± 0.158 (-0.244 to 0.32)	0.128 ± 0.086 (0.055 to 0.32)	-0.27 ± 1.54 (-2 to 3.32)	1.26 ± 0.84 (0.42 to 3.32)	-0.023 ± 0.16 (-0.244 to 0.32)	0.13 ± 0.09 (0.05 to 0.32)
Heart	-3.09 ± 1.77 (-6.02 to 0.36)	3.16 ± 1.63 (0.36 to 6.02)	-0.331 ± 0.19 (-0.658 to 0.041)	0.338 ± 0.175 (0.041 to 0.66)	-3.09 ± 1.77 (-6.02 to 0.36)	3.16 ± 1.63 (0.36 to 6.02)	-0.33 ± 0.19 (-0.66 to 0.04)	0.34 ± 0.17 (0.041 to 0.66)
Bone	-1.31 ± 1.98 (-4.58 to 2.9)	1.86 ± 1.41 (0.12 to 4.58)	-0.146 ± 0.24 (-0.482 to 0.427)	0.227 ± 0.156 (0.017 to 0.48)	-1.31 ± 1.98 (-4.58 to 2.9)	1.86 ± 1.41 (0.12 to 4.58)	-0.146 ± 0.24 (-0.48 to 0.43)	0.23 ± 0.16 (0.02 to 0.48)
Kidneys	-0.56 ± 1.88 (-5.31 to 1.37)	1.2 ± 1.51 (0.07 to 5.31)	-0.037 ± 0.132 (-0.284 to 0.109)	0.097 ± 0.093 (0.005 to 0.28)	-0.56 ± 1.88 (-5.31 to 1.37)	1.2 ± 1.51 (0.07 to 5.31)	-0.037 ± 0.132 (-0.284 to 0.11)	0.1± 0.09 (0.01 to 0.28)
Spleen	-0.05 ± 2.3 (-5.15 to 3.85)	1.6 ± 1.57 (0.11 to 5.15)	0.011 ± 0.188 (-0.337 to 0.356)	0.138 ± 0.12 (0.009 to 0.36)	-0.05 ± 2.3 (-5.15 to 3.85)	1.6 ± 1.57 (0.11 to 5.15)	0.011 ± 0.19 (-0.34 to 0.36)	0.138 ± 0.12 (0.009 to 0.36)
Bladder	2.39 ± 3.73 (-2.89 to 11.6)	3.11 ± 3.1 (0.48 to 11.6)	0.229 ± 0.328 (-0.221 to 1.05)	0.282 ± 0.28 (0.06 to 1.05)	2.39 ± 3.73 (-2.89 to 11.6)	3.11 ± 3.1 (0.48 to 11.6)	0.229 ± 0.328 (-0.221 to 1.05)	0.282 ± 0.28 (0.06 to 1.05)
Lungs	-2.42 ± 1.3 (-4.34 to -0.45)	2.42 ± 1.3 (0.45 to 4.34)	-0.275 ± 0.135 (-0.47 to -0.046)	0.275 ± 0.135 (0.046 to 0.47)	-2.42 ± 1.3 (-4.34to-0.45)	2.42 ± 1.3 (0.45 to 4.34)	-0.275 ± 0.135 (-0.47 to -0.046)	0.275 ± 0.135 (0.046 to 0.47)
Brain	-3.52 ± 3.25 (-7.87 to 3.05)	4.15 ± 2.29 (0.43 to 7.87)	-0.587 ± 0.525 (-1.198 to 0.45)	0.681 ± 0.38 (0.07 to 1.198)	-3.52 ± 3.25 (-7.87 to 3.05)	4.15 ± 2.29 (0.43 to 7.87)	-0.587 ± 0.525 (-1.198 to 0.45)	0.681 ± 0.38 (0.066 to 1.2)
All Organs	-1.10 ± 2.89 (-7.87 to 11.63)	2.34 ± 2.01 (0.07 to 11.6)	-0.144 ± 0.348 (-1.198 to 1.049)	0.270 ± 0.261 (0.005 to 1.2)	-1.11 ± 2.90 (-7.84 to 11.8)	2.34 ± 2.03 (0.08 to 11.8)	-0.144 ± 0.342 (-1.497 to 0.765)	0.23 ± 0.28 (0.01to 1.49)

Supplemental-Table 4. the results of the organ-wised evaluation of the results for the 90 kVp dataset after fine-tuning.

	FTC				TCM			
Organ	RE %	RAE %	ME (mGy)	MAE (mGy)	RE %	RAE %	ME (mGy)	MAE (mGy)
Liver	-0.41 ± 1.83 (-3.3 to 3.76)	1.37 ± 1.22 (0.14 to 3.76)	-0.04 ± 0.155 (-0.305 to 0.299)	0.112 ± 0.109 (0.009 to 0.305)	-0.5 ± 1.81 (-3.15 to 3.78)	1.37 ± 1.22 (0.07 to 3.78)	-0.036 ± 0.114 (-0.262 to 0.211)	0.08 ± 0.085 (0.001 to 0.262)
Heart	-3.81 ± 2.3 (-6.29 to 1.03)	4.01 ± 1.9 (0.04 to 6.29)	-0.325 ± 0.207 (-0.557 to 0.095)	0.343 ± 0.173 (0.002 to 0.557)	-3.85 ± 2.32 (-6.27 to 1.03)	4.04 ± 1.92 (0.04 to 6.27)	-0.24 ± 0.177 (-0.455 to 0.078)	0.255 ± 0.153 (0.002 to 0.455)
Bone	-3.16 ± 2.72 (-5.72 to 4.15)	3.92 ± 1.22 (1.56 to 5.72)	-0.282 ± 0.255 (-0.463 to 0.433)	0.361 ± 0.097 (0.167 to 0.463)	-3.06 ± 2.75 (-5.72 to 4.2)	3.82 ± 1.34 (0.8 to 5.72)	-0.212 ± 0.223 (-0.491 to 0.353)	0.276 ± 0.124 (0.101 to 0.491)
Kidneys	-1.81 ± 2.12 (-6.42 to 0.27)	1.86 ± 2.07 (0.2 to 6.42)	-0.101 ± 0.11 (-0.33 to 0.014)	0.104 ± 0.107 (0.01 to 0.33)	-2.03 ± 2.08 (-6.42 to 0)	2.03 ± 2.08 (0 to 6.42)	-0.078 ± 0.08 (-0.223 to 0)	0.078 ± 0.08 (0 to 0.223)
Spleen	-1.33 ± 2.08 (-4.5 to 2.05)	1.99 ± 1.39 (0.16 to 4.5)	-0.085 ± 0.134 (-0.314 to 0.12)	0.127 ± 0.091 (0.007 to 0.314)	-1.25 ± 2.22 (-4.7 to 2.77)	1.96 ± 1.56 (0.05 to 4.7)	-0.056 ± 0.079 (-0.203 to 0.083)	0.075 ± 0.059 (0.001 to 0.203)
Bladder	1.59 ± 4.3 (-3.6 to 11.86)	3.1 ± 3.27 (0.31 to 11.9)	0.109 ± 0.285 (-0.275 to 0.787)	0.211 ± 0.213 (0.013 to 0.787)	1.61 ± 4.22 (-3.57 to 11.7)	3.05 ± 3.23 (0.31 to 11.7)	0.067 ± 0.133 (-0.083 to 0.383)	0.101 ± 0.106 (0.013 to 0.383)
Lungs	1.42 ± 2.5 (-2.35 to 6.59)	2.12 ± 1.88 (0.11 to 6.59)	0.112 ± 0.211 (-0.236 to 0.549)	0.178 ± 0.154 (0.008 to 0.549)	1.46 ± 2.43 (-2.25 to 6.4)	2.09 ± 1.86 (0.04 to 6.4)	0.096 ± 0.165 (-0.133 to 0.425)	0.14 ± 0.125 (0.002 to 0.425)
Brain	2.19 ± 4.88 (-3.72 to 14.1)	3.57 ± 3.88 (0.06 to 14.1)	0.239 ± 0.56 (-0.522 to 1.528)	0.426 ± 0.421 (0.008 to 1.528)	2.11 ± 4.74 (-3.72 to 13.6)	3.5 ± 3.73 (0.15 to 13.6)	0.37 ± 0.884 (-0.522 to 2.702)	0.572 ± 0.756 (0.013 to 2.702)
All Organs	-0.67 ± 3.59 (-6.42 to 14.1)	2.74 ± 2.40 (0.04 to 14.1)	-0.047 ± 0.321 (-0.557 to 1.528)	0.233 ± 0.224 (0.002 to 1.528)	-0.69 ± 3.57 (-6.42 to 13.6)	2.73 ± 2.38 (0.01 to 13.6)	-0.011 ± 0.375 (-0.522 to 2.702)	0.197 ± 0.318 (0.001 to 2.701)

References

[1] H. Alkadhi and A. Euler, "The future of Computed Tomography: Personalized, functional, and precise.,"*Invest Radiol,* vol. 55, no. 9, pp. 545-555, 2020.

[2] P. Ritt, "Recent Developments in SPECT/CT.,"*Sem Nucl Med,* vol. 52, no. 3, pp. 276-285, 2022.

[3] N. Aide, C. Lasnon, C. Desmonts, I. S. Armstrong, M. D. Walker, and D. R. McGowan, "Advances in PET/CT Technology: An Update.,"*Sem Nucl Med,* vol. 52, no. 3, pp. 286-301, 2022.

[4] W. A. Kalender, "Dose in x-ray computed tomography.,"*Phys Med Biol,* vol. 59, no. 3, pp. R129-R150, 2014.

[5] Y. Salimi, M. R. Deevband, P. Ghafarian, and M. R. Ay, "Uncertainties in effective dose estimation for CT transmission scan in total body PET-CT imaging with Auto mA3D tube current modulation,"*International Journal of Radiation Research,* Original Research vol. 16, no. 4, pp. 465-472, 2018.

[6] H. G. Menzel, C. Clement, and P. DeLuca, "ICRP Publication 110. Realistic reference phantoms: an ICRP/ICRU joint effort. A report of adult reference computational phantoms,"*Ann ICRP,* vol. 39, no. 2, pp. 1-164, 2009.

[7] I. A. Tsalafoutas, M. Hassan Kharita, H. Al-Naemi, and M. K. Kalra, "Radiation dose monitoring in computed tomography: Status, options and limitations,"*Phys Med,* vol. 79, pp. 1-15, 2020.

[8] X. Li, D. Steigerwalt, and M. M. Rehani, "T-shirt size as a classification for body habitus in computed tomography (CT) and development of size-based dose reference levels for different indications,"*Eur J Radiol,* vol. 151, p. 110289, 2022.

[9] F. Ria, W. Fu, J. Hoye, W. P. Segars, A. J. Kapadia, and E. Samei, "Comparison of 12 surrogates to characterize CT radiation risk across a clinical population,"*Eur Radiol,* vol. 31, no. 9, pp. 7022-7030, 2021.

[10] L. Klein *et al.*, "Patient-specific radiation risk-based tube current modulation for diagnostic CT,"*Med Phys,* vol. 49, no. 7, pp. 4391-4403, 2022.

[11] C. J. Martin and M. Barnard, "How much should we be concerned about cumulative effective doses in medical imaging?,"*J Radiol Prot,* vol. 42, no. 1, 2022.

[12] N. Zewde, F. Ria, and M. M. Rehani, "Organ doses and cancer risk assessment in patients exposed to high doses from recurrent CT exams,"*Eur J Radiol,* vol. 149, p. 110224, 2022.

[13] R. S. Arellano, K. Yang, and M. M. Rehani, "Analysis of patients receiving >/= 100 mSv during a computed tomography intervention,"*Eur Radiol,* vol. 31, no. 5, pp. 3065-3070, 2021.

[14] Bayer HealthCare. "Radimetrics Enterprise Platform: Dose Management Solution." Bayer HealthCare. http://www.radiologysolutions.bayer.com/products/ct/dosemanagement/rep/. (accessed 2016).

[15] A. Akhavanallaf *et al.*, "An update on computational anthropomorphic anatomical models,"*Digit Health,* vol. 8, p. 20552076221111941, 2022.

[16] T. Xie, A. Akhavanallaf, and H. Zaidi, "Construction of patient-specific computational models for organ dose estimation in radiological imaging,"*Med Phys,* vol. 46, no. 5, pp. 2403-2411, 2019.

[17] N. Guberina *et al.*, "Verification of organ doses calculated by a dose monitoring software tool based on Monte Carlo Simulation in thoracic CT protocols,"*Acta Radiol,* vol. 59, no. 3, pp. 322-326, 2018.

[18] R. Al-Senan, K. Brown, M. Erdman, and S. King, "The uncertainty of thyroid dose estimate in chest CT,"*Biomed Phys Eng Express,* vol. 6, no. 6, 2020.

[19] M. Tahiri, M. Mkimel, Y. Benameur, R. El Baydaoui, M. R. Mesradi, and O. El Rhazouani, "Organ Dose Estimation for Adult Chest CT Examination Using GATE Monte Carlo Simulation,"*Physics of Particles and Nuclei Letters,* vol. 18, no. 4, pp. 502-509, 2021.

[20] S. Sharma, A. Kapadia, W. Fu, E. Abadi, W. P. Segars, and E. Samei, "A real-time Monte Carlo tool for individualized dose estimations in clinical CT,"*Phys Med Biol,* vol. 64, no. 21, p. 215020, 2019.

[21] B. De Man, M. Wu, P. FitzGerald, M. Kalra, and Z. Yin, "Dose reconstruction for real-time patient-specific dose estimation in CT,"*Med Phys,* vol. 42, no. 5, pp. 2740-51, 2015.

[22] B. H. M. van der Velden, H. J. Kuijf, K. G. A. Gilhuijs, and M. A. Viergever, "Explainable artificial intelligence (XAI) in deep learning-based medical image analysis.,"*Med Image Anal,* vol. 79, p. 102470, 2022.

[23] Y. Fu, Y. Lei, T. Wang, W. J. Curran, T. Liu, and X. Yang, "Deep learning in medical image registration: a review.,"*Phys Med Biol,* vol. 65, no. 20, p. 20TR01, 2020.

[24] Y. Fu, Y. Lei, T. Wang, W. J. Curran, T. Liu, and X. Yang, "A review of deep learning based methods for medical image multi-organ segmentation.,"*Phys Med,* vol. 85, pp. 107-122, 2021.

[25] J. Maier, L. Klein, E. Eulig, S. Sawall, and M. Kachelriess, "Real-time estimation of patient-specific dose distributions for medical CT using the deep dose estimation,"*Med Phys,* vol. 49, no. 4, pp. 2259-2269, 2022.

[26] E. Tzanis and J. Damilakis, "A novel methodology to train and deploy a machine learning model for personalized dose assessment in head CT,"*Eur Radiol,* vol. 32, no. 9, pp. 6418-6426, 2022.
[27] K. A. S. H. Kulathilake, N. A. Abdullah, A. Q. M. Sabri, and K. W. Lai, "A review on deep learning approaches for low-dose Computed Tomography restoration.,"*Complex Intell Syst,* p. *in press*, 2022.
[28] Y. Salimi *et al.*, "Deep learning-based fully automated Z-axis coverage range definition from scout scans to eliminate overscanning in chest CT imaging.,"*Insights Imaging,* vol. 12, no. 1, p. 162, 2021.
[29] W. Schneider, T. Bortfeld, and W. Schlegel, "Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions,"*Phys Med Biol,* vol. 45, no. 2, pp. 459-78, 2000.
[30] A. Akhavanallaf, T. Xie, and H. Zaidi, "Assessment of uncertainties associated with Monte Carlo-based personalized dosimetry in clinical CT examinations,"*Phys Med Biol,* vol. 65, no. 4, p. 045008, 2020.
[31] J. S. Hendricks, G. W. McKinney, M. L. Fensin, M. R. James, R. C. Johns, and et al, "MCNPX, Extensions 2.6.0.," Los Alamos National Laboratory, NM, Los Alamos, NM, LA-UR-08-2216, 2008.
[32] D. A. Low, W. B. Harms, S. Mutic, and J. A. Purdy, "A technique for the quantitative evaluation of dose distributions,"*Med Phys,* vol. 25, no. 5, pp. 656-61, 1998.
[33] A. Wang, A. Maslowski, T. Wareing, J. Star-Lack, and T. G. Schmidt, "A fast, linear Boltzmann transport equation solver for computed tomography dose calculation (Acuros CTD),"*Med Phys,* vol. 46, no. 2, pp. 925-933, 2019.
[34] AAPM, "Use of Water Equivalent Diameter for Calculating Patient Size and Size-Specific Dose Estimates (SSDE) in CT," AAPM, The Report of AAPM Task Group 220, 2014.

Chapter 9

Whole-body voxel-based internal dosimetry using deep learning

Azadeh Akhavanallaf, Isaac Shiri, Hossein Arabi, and Habib Zaidi

Contribution: Study conception and design, methodological development, computer programming, analysis and interpretation of results, manuscript preparation.

Eur J Nucl Med Mol Imaging Vol. 48, No. 3, pp 670-682 (2021)

Abstract

Purpose In the era of precision medicine, patient-specific dose calculations using Monte Carlo (MC) simulations is deemed the gold standard technique for risk-benefit analysis of radiation hazards and correlation with patient outcome. Hence, we propose a novel method to perform whole-body personalized organ-level dosimetry taking into account the heterogeneity of activity distribution, non-uniformity of surrounding medium and patient-specific anatomy using deep learning algorithms.

Methods We extended the voxel-scale MIRD approach from single S-value kernel to specific S-value kernels corresponding to patient-specific anatomy to construct 3D dose maps using hybrid emission/transmission image sets. In this context, we employed a Deep Neural Network (DNN) to predict the distribution of deposited energy, representing specific S-values, from a single source in the center of a 3D kernel composed of human body geometry. The training dataset consists of density maps obtained from CT images and the reference voxelwise S-values generated using Monte Carlo simulations. Accordingly, specific S-value kernels are inferred from the trained model and whole-body dose maps constructed in a manner analogous to the voxel-based MIRD formalism, i.e. convolving specific voxel S-values with the activity map. The dose map predicted using the DNN was compared to the reference generated using MC simulations and two MIRD-based methods, including single and multiple S-values and Olinda/EXM software package.

Results The predicted specific voxel S-value kernels exhibited good agreement with the MC-based kernels serving as reference with a Mean Relative Absolute Error (MRAE) of 4.5±1.8 (%). Bland & Altman analysis showed the lowest dose bias (2.6%) and smallest variance (CI: -6.6, +1.3) for DNN. The MRAE of estimated absorbed dose between DNN, MSV, and SSV with respect to the MC simulation reference were 2.6%, 3%, and 49%, respectively. In organ-level dosimetry, the MRAE between the proposed method and MSV, SSV, and Olinda/EXM were 5.1%, 21.8% and 23.5%, respectively

Conclusion The proposed DNN-based WB internal dosimetry exhibited comparable performance to the direct Monte Carlo approach while overcoming the limitations of conventional dosimetry techniques in nuclear medicine.

I. Introduction

Personalized medicine is a new paradigm aiming at improving healthcare while lowering the costs, thus offering great potential for patient-specific diagnosis and optimal treatment [1]. Precision medicine aims at shifting from the current one-size fits-all strategy to an individualized model. Dose calculation in nuclear medicine is tightly linked to this approach [2]. In this framework, personalized dose estimation is crucial for optimizing clinical procedures while minimizing the risk of radiation-induced toxicity [3].

In current clinical practice, patient dose monitoring is commonly based on simplified models, such as those derived by the Medical Internal Radiation Dose Committee (MIRD) formalism [4]. The traditional MIRD technique is based on organ-level dosimetry using time-integrated activity and radionuclide S-values, which represents the mean absorbed dose to a target organ per radioactive decay in a source organ. These quantitative parameters are modeled based on a reference computational model. This approach assumes a uniform activity distribution within each organ and ignores individual anatomical characteristics. To cope with inter-subject variability of anatomical features, the organ-level dosimetry approach was later extended by developing habitus-specific and patient-specific computational models [5-9]. Furthermore, voxel-based dosimetry techniques have been developed, including dose point kernel [10] and voxel S-value (VSV) [4] approaches. Unlike probabilistic methods, dose point kernel is a deterministic approach that calculates the radial absorbed dose distribution around an isotropic point source in a homogeneous water medium [11, 12]. Voxel-level MIRD schema is defined as a 3D voxel matrix representing the mean absorbed dose to a target voxel per unit activity in a source voxel embedded in an infinite homogeneous medium using Monte Carlo (MC) simulations. However, voxel-based dose calculation should in principle take into account non-uniform activity distribution of the radiotracer, the heterogeneity of the medium consisting of different material compositions, e.g. lung, soft tissue, and bone is ignored. In this regard, direct MC simulations, deemed the gold standard for implementation of a reliable dose calculation framework in clinical setting, enables accurate estimation of whole-body dose map [13, 14]. Though MC simulation takes into account the non-uniform activity distribution and heterogeneity of patient-specific anatomical features, it suffers from expensive computational burden. A number of previous works reported on the use of MC simulations in the context of personalized dosimetry in nuclear medicine [15-17]. Hybrid PET/CT or SPECT/CT images are fed into the MC simulator to model energy deposition of radiation emitted from the injected radiotracer considering the patient-specific anatomy and voxelwise activity distribution obtained from CT and PET/SPECT images, respectively. Several works focused on reaching an optimal compromise between accurate voxel-scale dosimetry and the computational burden [18, 19]. Khazaee Moghadam *et al.* proposed a tissue-specific dose point kernel approach implemented on a stylized phantom [20]. Lee *et al.* extended further this idea by applying this methodology on real patient data [21]. They considered multiple material densities for internal dose calculation by providing multiple voxelwise S-value kernels for various media with different densities according to human body tissues. This enabled to provide multiple voxel-scale dose maps in an analogous manner to the MIRD calculations. Consequently, each density-specific dose map was multiplied by the corresponding binary mask of the given density regions obtained from CT-based segmentation, thus enabling the calculation of the final dose map by superposition of the multiple density-specific dose maps. This method improves the accuracy of dosimetry calculations compared to the single voxel S-value approach, but relies on a basic assumption that energy depositions in each voxel arise mainly from self-absorption. This simplification introduces an extra error on the estimated dose distribution, particularly in the boundary of tissues with different densities.

Accurate patient-specific dosimetry is becoming a must taking advantage of advances in targeted radionuclide therapy and theranostic imaging [2]. In personalized dosimetry, MC simulation is still considered the most accurate technique and the de facto reference standard for research application. Yet, this approach is not employed in routine clinical procedures owing to the heavy computational burden. Deep learning emerged as a promising technique in the area of computer vision and image processing, exhibiting superior performance over conventional state-of-the-art methods in medical images analysis in PET and SPECT imaging, including attenuation and scatter correction [22-24], low-count image reconstruction [25-27], as well as automated image segmentation [9, 28]. More recently, deep learning approaches were employed for radiation dose estimation. Mardani *et al.* introduced a dose distribution prediction method in external beam radiation therapy using a multi-layer convolutional auto-encoder architecture [29]. Nguyen *et al.* used a U-Net architecture for clinical treatment plan optimization to improve the treatment plan quality and uniformity while reducing the computational time [30]. Ma *et al.* implemented a deep

learning method to provide isodose features for modulated arc therapy treatment plans [31]. Kearney *et al.* proposed a 3D fully-convolutional dose prediction algorithm for prostate stereotactic body radiotherapy patients [32].

For effective training of a deep learning algorithm, well-defined ground truth is an essential ingredient [33]. In the above-mentioned seminal works, the ground truth was obtained from a substitute of MC dosimetry for the training of the networks that may bear some inaccuracies owing to the simplifications in physical models [34]. To address this limitation, Lee et al. used a U-Net deep neural architecture for internal dosimetry where the training ground truths were obtained from direct MC simulation [35]. They fed CT images, representing patient structural features, and static PET images, representing the activity distribution, into the network as input to predict a 3D dose map rate. Gotz et al. set out to estimate dose maps of patients who received ^{177}Lu-PSMA using a modified U-Net network [36]. In this work, the training datasets consisted of a two-channel input, including CT images (i.e. patient-specific density map), MIRD-based voxel-scale dose map obtained from SPECT images as well as the ground truth obtained from direct MC simulations. In these two works, the deep learning networks were trained using whole-body dose maps obtained from direct MC simulations. However, generation of a comprehensive training dataset in this manner would be challenging owing to the prohibitive computational burden of MC calculations. Hence, these works either relied on a limited number of training samples or made some approximations that could affect the accuracy of the proposed models. Lee et al. reported that the time required for a single full whole-body MC simulation exceeds 4704.03 hours using a CPU with four cores and 16 GB RAM [35]. However, GPU-based MC simulations have been recently proposed to overcome this challenge [37-39]. In this regard, we proposed a novel methodology to estimate whole-body dose distributions using a deep convolutional neural network, wherein unlike previous studies, generation of training datasets is no longer a bottleneck. The proposed dose map generation framework consists of two steps. In the first step, a deep neural network (DNN) is employed to predict dose distribution kernels, wherein the training dataset consists of only density maps obtained from CT images as input and the corresponding dose distribution kernel for a point source with unit activity obtained from MC simulations as output. In this approach, the simulation time for generating a ground truth (dose distribution map around the central voxel source) covering the annihilation photon mean free path is about 8000 times less than that required for whole-body MC simulations. This strategy makes it possible to provide a diverse and extensive training dataset. In addition, this approach would reduce the complexity of the training process as the DNN model should learn simpler features corresponding to a point source distribution compared to direct translation from hybrid density/activity maps to absorbed dose map. In the second step, specific dose distribution kernels predicted by the trained model are convolved with the activity map (here time-integrated activity from dynamic PET images) to generate the final whole-body dose map, in a manner analogous to the voxel-based MIRD formalism.

II. Materials and Methods

Method description

Direct MC simulations, wherein the 3D hybrid PET/CT or SPECT/CT images are fed into a simulator to produce the whole-body dose distribution, are regarded as the gold standard approach. The computational burden of direct MC simulations for building a comprehensive and large training dataset is prohibitive. Hence, we split the direct process into two main parts as schematically illustrated in Figure 1. The idea is inspired from the MIRD-based voxel-scale dosimetry formalism [4] where a single voxel S-value kernel is convolved with the activity map (e.g PET images) to produce a whole-body dose map (Eq. 1). In the present study, we extended this idea through estimation of the specific kernels according to the density map obtained from patient's CT images. Analogous to the MIRD-based voxel-scale dose kernel, we generate specific kernels, i.e. $S(voxel_k \leftarrow voxel_h)$ in Eq. 1, in such a way that the central voxel contains the unit activity of given a radiotracer, where the surrounding medium is defined based on the patient density map.

The principle of the reciprocity theorem states the reversibility or bilateralism of the interactions upon location interchange of the source and target in a uniform isotropic model. Loevinger introduced this theorem to dose calculation problems in a uniform homogenous medium [40]. Cristy reported that the reciprocity theory is warranted in heterogeneous computational phantoms for photons [41]. We extended this theory to heterogeneous media by applying a source to target correction factor of the energy-absorption coefficient ratio [42]. Since the

deposited energy in each voxel depends on the energy fluence multiplied by the mass energy-absorption coefficient of the medium [43], we modified the conditions of location interchange of source and target tacking into account the ratio of the energy-absorption coefficient of target voxel to the source voxel. The reciprocal energy fluence is assumed to be approximately equal for annihilation photons with dominant Compton scattering interaction.

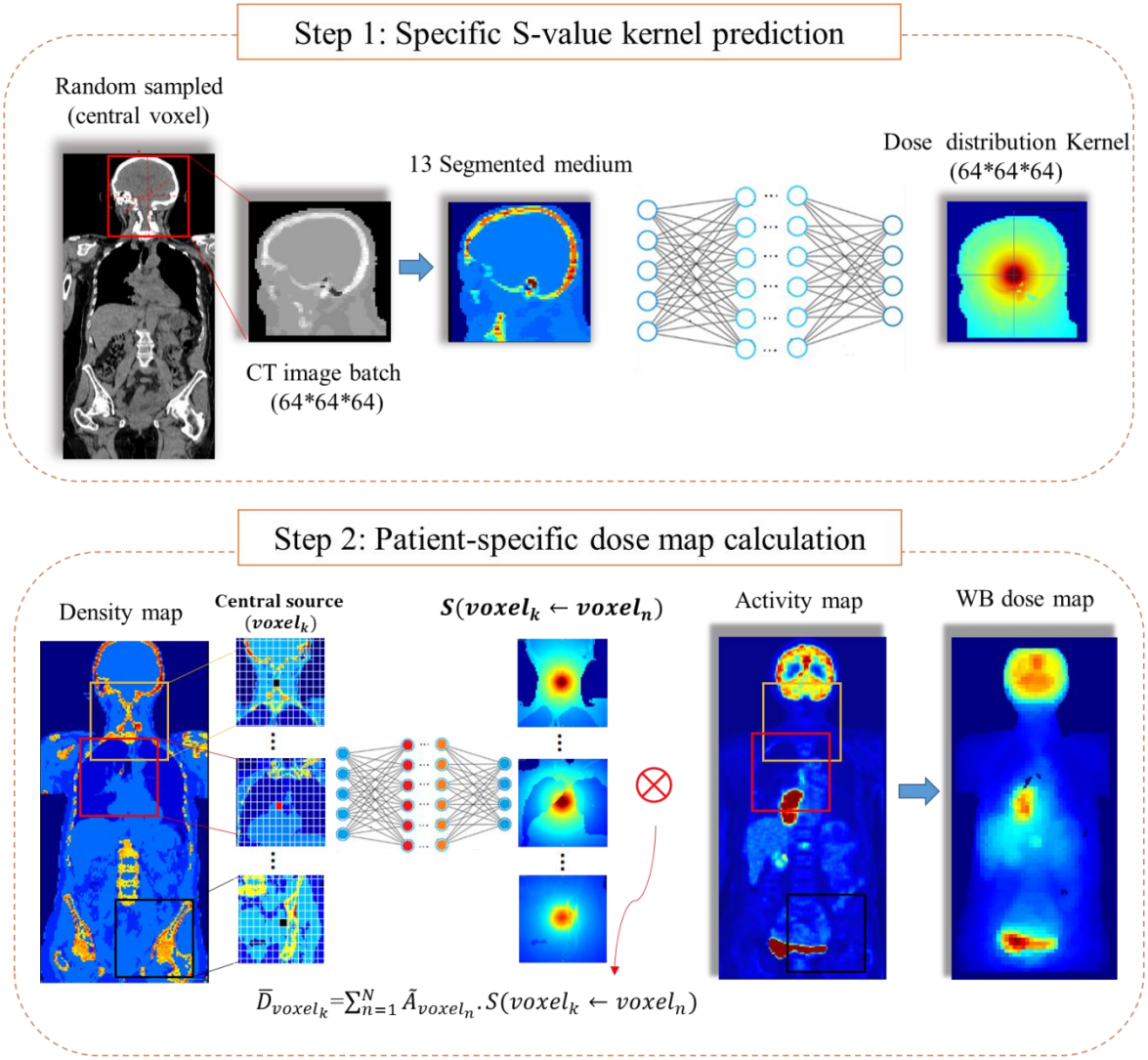


Figure 1. Schematic representation of the voxel-scale dosimetry procedure. The top and bottom panels show the deep learning-based specific S-value kernel prediction and MIRD-based voxel dosimetry formalism.

To generate the specific kernels, the distribution of deposited energy around the source voxel was calculated using MC simulations. The size of the kernel depends on the type of radiotracer, i.e. decay mode and energy spectrum. In this work, we defined the size of kernels as 19.2×19.2×19.2 cm^3 where the mean free path of annihilation photons in human tissue has been reported to be about 7 cm [44].

$$\bar{D}(voxel_k) = \sum_{h=0}^{N} \tilde{A}_{voxel_h} . S(voxel_k \leftarrow voxel_h) \qquad (1)$$

In the first step, we employed a DNN to predict the specific energy deposition kernel when the source voxel is located in the center of the kernel (Figure 1). The input data for the training is 3D volume density maps while the corresponding output is 3D volume dose map obtained from MC simulations. To prepare the input dataset for training, single voxels were randomly sampled from whole-body CT images and the surrounding volumes (19.2×19.2×19.2 cm^3) were extracted into 64×64×64 matrices to generate the input samples. Given the input matrices, the MC simulator was employed to produce the dose distribution kernel considering a unit activity at the center of each matrix. In other words, the training of the model was performed for single-point sources located in

various positions within the density volume map, i.e. whole-body CT images. Hence, to produce a comprehensive training dataset covering different anatomical sites, we randomly sampled voxels from different whole-body CT images and the surrounding volumes were extracted to generate the input samples. In the second step, the whole-body dose map was calculated by voxelwise convolution of the specific kernels with the activity map (Eq. 1). Hence, we inferred the specific dose distribution kernel for each source voxel, i.e. $S(voxel_k \leftarrow voxel_h)$, using our trained neural network model. We estimated the whole-body dose map in an analogous way to the MIRD voxel formalism, which convolves a single S-value kernel with each voxel in the activity map, yet using specific S-values kernel for each voxel.

Deep Neural Network architecture

In this work, the *ResNET* [45] architecture implemented on TensorFlow platform, composed of 20 convolutional layers with dilation convolution operations within different levels of feature extraction, was utilized. The dilation factor supports the expansion of the receptive field-of-view without resolution loss by increasing the space between original kernel elements. For low-level feature extraction, a dilation factor of zero was used within the first seven layers, a dilation factor of two within the second seven layers for medium-level feature extraction, and a dilation factor of four within the last six layers for high-level feature extraction. Leaky rectified linear unit (LReLU) was used as activation function. The *ResNET* architecture benefits from 9 residual blocks, which results in a large number of receptive fields and improves the process of feature extraction and network convergence (Figure 2).

For the training of the model, pairs of CT density images and deposited energy kernels were considered as input/target, respectively. The *ResNET* model with a 3D spatial window equal to 3×3×3 voxels were used. The following setting was used for the training of the network: learning rate = 0.001, sample per volume = 1, optimizer = Adam, and decay = 0.0001. The optimization of the network was carried out based on the L2 loss function.

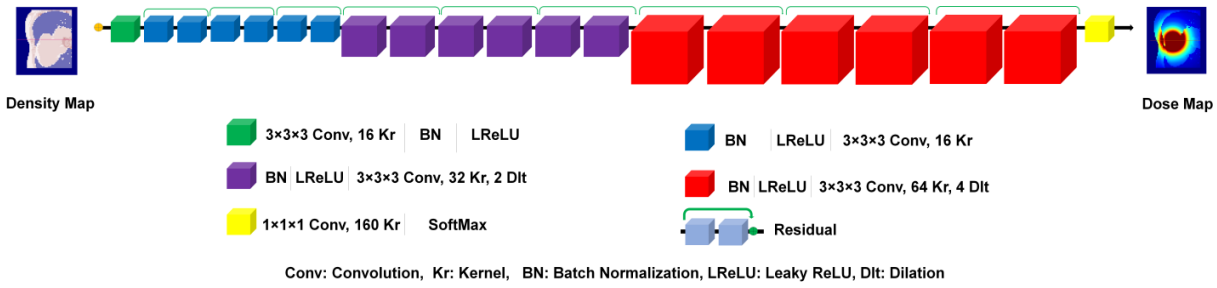


Figure 2. Schematic diagram of the ResNET architecture.

Data preparation

To prepare the training data set, density maps were extracted from CT images. CT Hounsfield Units (HUs) have a strong correlation with electron density, and consequently with the mass density of the medium. We converted HU values to mass density using the methodology proposed by Schneider *et al.* which established linear multi-regression models for the segmentation of CT images into different tissue densities [46]. We extracted density maps consisting of 13 tissue densities, including air, lung, fat, soft tissue, and bone where values higher than 100 HUs were divided into eight discrete density values. Afterward, the whole-body density maps were resampled to 3 mm voxel size in three-dimensions. To build the ground truth data, MC simulations served as standard of reference. The MCNP transport code [47] was employed for the generation of energy deposition kernels, i.e. specific voxel S-values. To this end, one voxel was randomly sampled from the whole-body density maps and a 3D matrix of 64×64×64 voxels around the central voxel was extracted. This matrix, representing a heterogeneous medium of patients' anatomical structures, was directly imported to the MCNP code. The material compositions of 13 segmented tissues were defined based on Schneider *et al.* [46]. The central voxels of the extracted 3D matrix were defined as source location with uniform distribution of Fluorodeoxyglucose (^{18}F-FDG). Since the resolution of the activity distribution (here PET images with an average resolution of 3 mm) determines the spatial accuracy of dosimetry estimations, we adopted the same resolution for the calculation of dose maps. The energy spectrum of emitted positrons was taken from [48], where the positron energy spectrum follows a Fermi distribution with an average of 242.8 KeV and maximum energy of 633.5 KeV. The output of MC simulations consist of 3D kernels

(64×64×64) with 3 mm resolution using energy deposition mesh tally in unit of MeV/cm^3 per particle. Three million particles were tracked to reach a statistical uncertainty less than 4% in the border voxels at about 10 cm away from the central voxel.

Clinical studies

To provide whole-body dose maps from an activity map based on Eq. 1, a specific S-value kernel is required for each single voxel of the activity map. Whole-body unenhanced CT images of 24 patients acquired on Siemens Definition Edge system were used for the training of the model (generation of the training dataset). The study protocol was approved by Geneva Ethics committee and all patients provided written informed consent. For evaluation of the model, hybrid PET/CT image sets consisting of a low-dose CT scan and dynamic whole-body PET scans were employed. The hybrid PET/CT image sets were acquired on a Siemens Biograph mCT scanner using a dynamic scanning protocol at 13-time points after intravenous injection of 18F-FDG [49, 50]. PET scanning was conducted using continuous bed motion scan at ever increasing time intervals. PET image reconstruction was performed using 3D iterative ordinary Poisson OSEM (3D-OP-OSEM) algorithm with a voxel size of 4.07 × 4.07 × 3 mm.

Dose map calculation

To estimate whole-body voxelwise absorbed dose, the trained model was fed with patient-specific density maps to generate the specific dose distribution kernels, $S(voxel_k \leftarrow voxel_h)$, for each single voxel (i.e. $voxel_k$) in the PET image, wherein the corresponding voxel in CT images and its surrounding 64×64×64 voxels were considered as the input density map. The predicted specific S-values were corrected by element-wise multiplication of the ratio of the energy-absorption coefficient of the target voxel to the source voxel obtained from [51]. Lastly, specific S-values underwent voxelwise convolution with the cumulated activity map to create the whole-body dose map (Eq. 1). The cumulated activity map was calculated by analytical integration of voxelwise time activity curves over 13-time points dynamic PET frames (Eq. 2).

$$\tilde{A}_{Total} = \frac{1}{2}\sum_{i=0}^{13}(A_i + A_{i+1}) \,.\, \Delta t_i + \int_T^{\infty} A_f e^{-\lambda t} dt \qquad (2)$$

In Eq. 2, $\tilde{A}_{Total}$ is the total number of disintegrations, A_i is the activity concentration in the source organ obtained from static images at the ith time frame, A_f is the activity concentration in the last time point of measurement, and λ is the decay factor of the radionuclide. Bladder voiding schedules were not taken into account. To conduct patient-specific whole-body voxelwise dose estimation, the results were converted in Gy after multiplication by a correction factor of 0.9673 corresponding to the fraction of positron emission for ^{18}F.

To evaluate the proposed method, the predicted absorbed dose from the current model was compared against direct MC dose estimation serving as standard of reference and different MIRD-based approaches, including the OLINDA/EXM software (organ-scale MIRD formalism) [52], single voxel S-value (SSV) and multiple voxel S-value (MSV). For organ-level dosimetry, regions-of-interest were manually drawn on CT images to delineate eight organs, namely brain, heart, kidneys, liver, lungs, spleen, bone, and bladder. Lesions identified on PET images were segmented using a fixed threshold of 42% of SUV_{max} and manually edited to remove the background and include necrotic regions. The kinetic data required by Olinda/EXM software were calculated from the cumulated activity using Eq. 2 and the masses of organs were modified based on organ masks defined from the segmentation of CT images. SSV and MSV voxel-scale dosimetry was designed based on the MIRD formalism (Eq. 1) where the voxel S-value kernels were generated from MCNP code with the same kernel size used in the previous step, i.e. 19.2 cm in 3D with 3mm resolution. Ten million particles were simulated to build a 64×64×64 kernel in an infinite homogenous medium considering a unit activity in the central voxel. In the MSV method [21], the S-value kernels of four different media consisting of soft-tissue, lung, and two different densities of bone (with different Calcium content) were simulated.

Quantitative analysis

Voxelwise mean absolute error (MAE), mean relative absolute error (MRAE %), and root mean square error (RMSE) were calculated between reference and predicted dose maps.

$$MAE = \frac{1}{vxl}\sum_{v=1}^{vxl}\left|Image_{predicted}(v) - Image_{ref}(v)\right| \quad (3)$$

$$MRAE(\%) = \frac{1}{vxl}\sum_{v=1}^{vxl}\left|\frac{(Image_{predicted})_v - (Image_{ref})_v}{(Image_{ref})_v}\right| \times 100\% \quad (4)$$

$$RMSE = \sqrt{\frac{1}{vxl}\sum_{v=1}^{vxl}(Image_{predicted}(i) - Image_{ref}(i))^2} \quad (5)$$

where $Image_{predicted}$ stands for dose map generated by the DNN and $Image_{ref}$ stands for the reference dose map. *vxl* and *v* denote the total number of voxels and voxel index, respectively.

III. Results

Network validation

The total number of training dataset consisted of 12'100 pairs of volumetric images of density maps and energy deposition kernels extracted from 24 different CT image sets. The specific voxelwise S-value kernels, obtained from the DNN, were in good agreement with the reference MC kernels. The axial profiles plotted over reference and predicted voxelwise S-value kernels in the lung region are shown in Figure 3. The mean relative voxel-wise difference between the two profiles is about 3.3%. Figure 4 illustrates the comparison of predicted voxel S-values (64×64×64) against MC simulations for the test case in the lung region with MRAE, RMSE and MAE of 4.5±1.8 (%), (1.8 ± 0.53)×10^{-5} (MeV/cm^3) and (1.8±0.71)×10^{-6} (MeV/cm^3), respectively. Furthermore, the voxelwise joint histogram plot depicting the correlation between the predicted kernels and MC simulations is presented, where a coefficient of determination (R^2) of 0.98 was achieved.

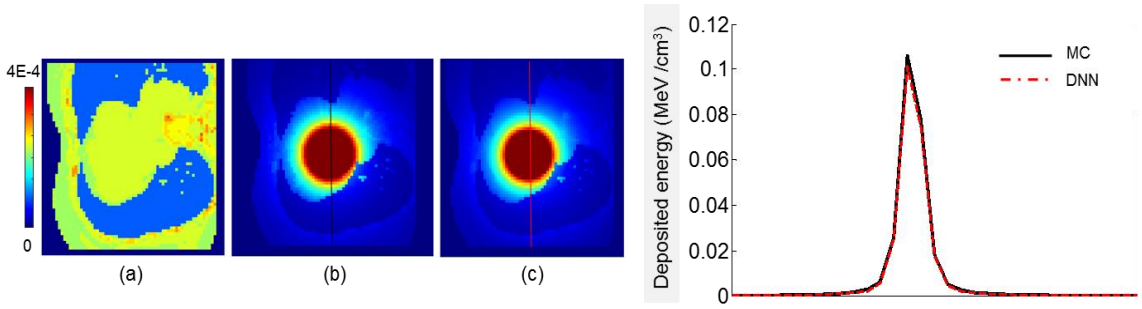


Figure 3. (a) CT-based density map, (b) reference kernel obtained from MC simulations, (c) predicted kernel by the DNN model. Line profiles across the S-value kernels (right panel) comparing kernels obtained from MC simulations of DNN model predictions.

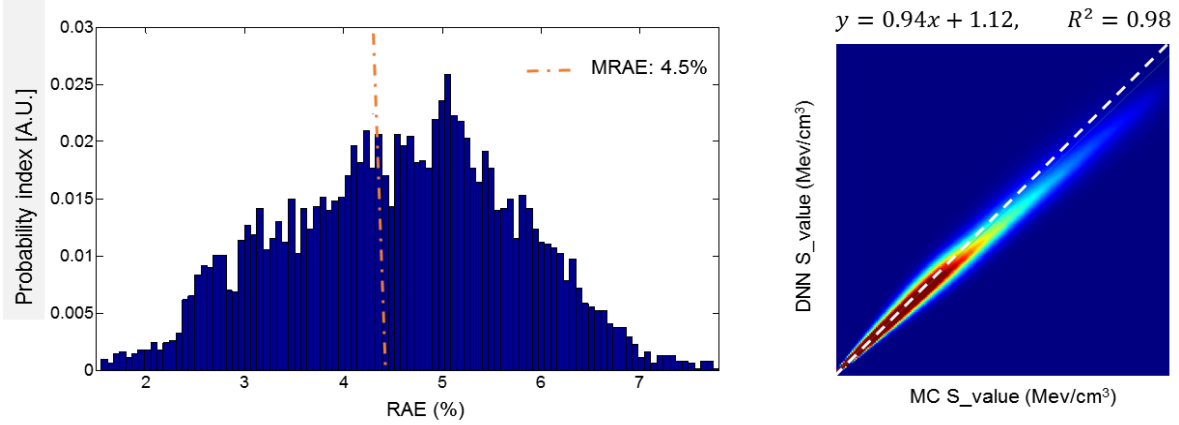


Figure 4. Probability distribution of Relative Absolute Error (RAE) for predicted voxelwise S-value kernels (64×64×64) with respect to MC simulations (left). A.U.= arbitrary units. Voxelwise joint histogram plot depicting the correlation of predicted kernels with respect to MC simulations (right).

Analysis of dose distributions

To assess the impact of medium heterogeneity on dosimetry results, a whole-body map of deposited energy was generated for a patient-specific computational phantom with unit activity distribution using three different methods, including DNN, MSV, and SSV. In this regard, calculation of patient-specific absorbed dose map (step 2 in Figure 1) involves filling the patient's body contour with a unit activity distribution instead of a time-integrated activity map obtained from a dynamic PET series. Dose profiles over axial and coronal slices are illustrated in Figure 5. It is expected that SSV in medium with densities lower than water overestimates the deposited energy while underestimating the deposited energy for higher density media. The deposited energies obtained from MSV confirms the limitation of this method in the heterogeneous boundaries in the spine area with an average density of about 1.12 g/cm^3 (b-b' line profile). The voxelwise dose maps predicted by DNN and estimated using MIRD-based methods, including SSV and MSV, were compared with the results obtained from MC simulations for a patient diagnosed with lung adenocarcinoma having a pulmonary tumor of about 120 g. Figure 6 displays a representative dose profile drawn on axial views comparing dose maps estimated from DNN, MSV, and SSV against MC simulations.

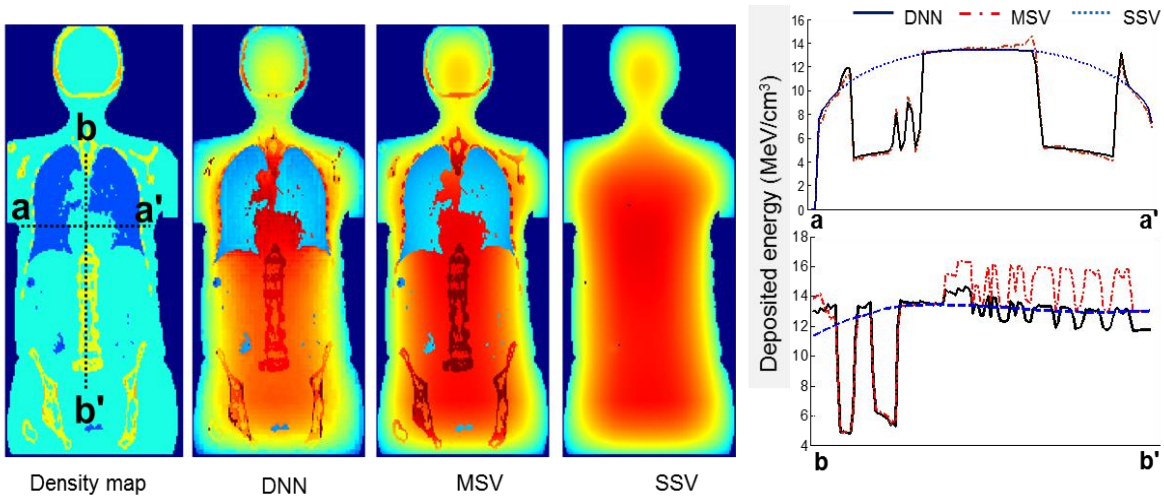


Figure 5. Voxelwise deposited energy (MeV/cm^3) in a patient-specific computational phantom with unit activity distribution estimated by DNN, MSV, and SSV.

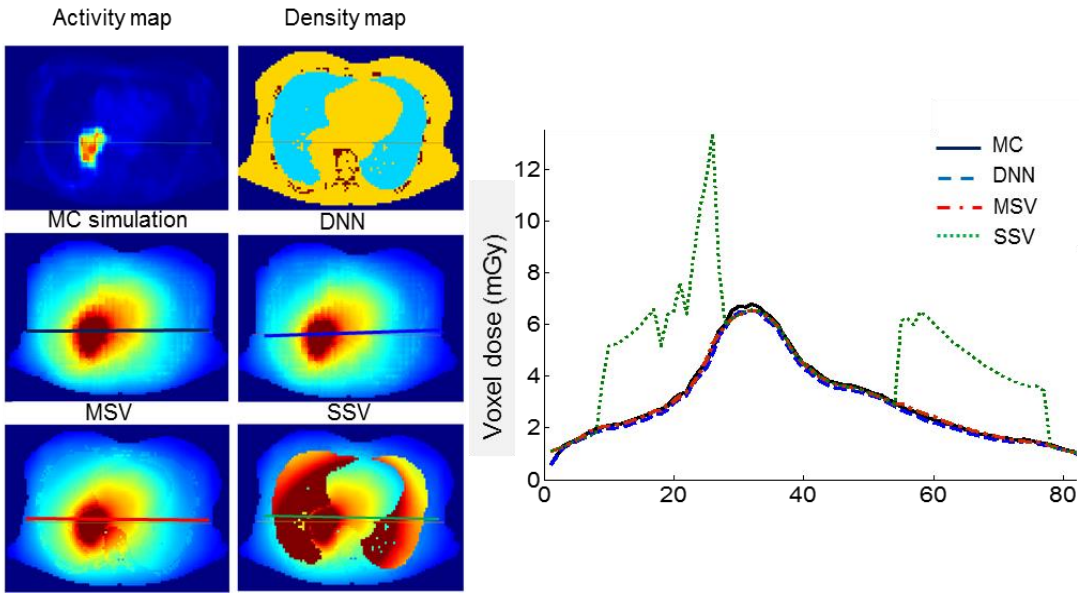


Figure 6. Dose distributions and profiles (right) drawn on axial views comparing dose maps estimated using DNN, MSV, and SSV methods against MC simulations.

To quantify the agreement between the different methods with respect to the standard of reference, Bland-Altman plots compare absorbed doses calculated using DNN, MSV and SSV with MC-based calculations. Figure 7 illustrates the bias and variance with 95% confidence interval (CI) of these methods against the standard of reference method, where the data points reflect the percent difference of voxelwise dose values. The results show that the lowest absorbed dose bias (2.6 %) and the smallest variance (CI: -6.6%, +1.3%) were achieved by the DNN approach. In addition, the results obtained using MSV demonstrated good agreement with the ground truth (absorbed dose bias of 2.9 % and variance of CI: -6.8%, +12.6%), except in some regions corresponding to heterogeneous boundaries. Conversely, SSV showed significant discrepancy compared to the reference in lung and bone regions. In the lung region illustrated in Figure 8 (top left), four VOIs over the heart, bone, lower lobes of the lungs, and pulmonary tumor were drawn on fused PET/CT images to perform quantitative analysis of absorbed doses within the VOIs. The mean absolute relative errors of estimated absorbed doses between DNN, MSV, and SSV against MC simulations were 2.6±0.94 %, 3±3.5 %, and 49±68 %, respectively.

Whole-body voxelwise absorbed dose estimations based on time-integrated activity and patient-specific anatomy obtained from a dynamic PET/CT scan are presented in Figure 9 along with two profiles plotted over axial and coronal views. Organ-level dosimetry was extracted from the dose maps obtained from DNN, MSV and SSV methods and compared against a commercial organ-based MIRD dosimetry software, i.e. Olinda/EXM (Figure 10). In most organs, Olinda/EXM underestimates the absorbed dose compared to other voxel-based methods except for lung and pulmonary tumor. The MRAE between organ doses estimated by DNN method and MSV, SSV and Olinda/EXM were 5.1%, 21.8% and 23.5 %, respectively.

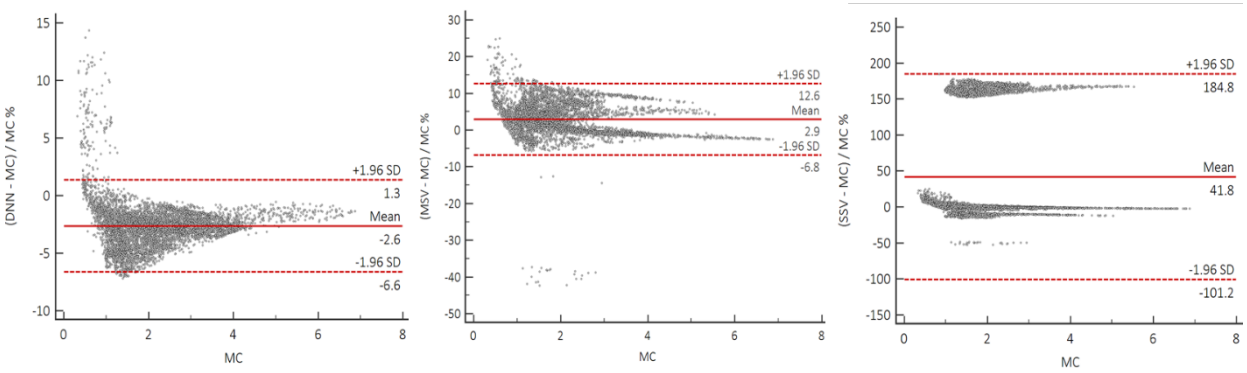


Figure 7. Bland & Altman plots of voxelwise dose differences in the lung region calculated using DNN (left), MSV (middle) and SSV (right) with respect to MC-based calculations serving as standard of reference. The solid and dashed lines denote the mean and 95% CI of the dose value differences, respectively.

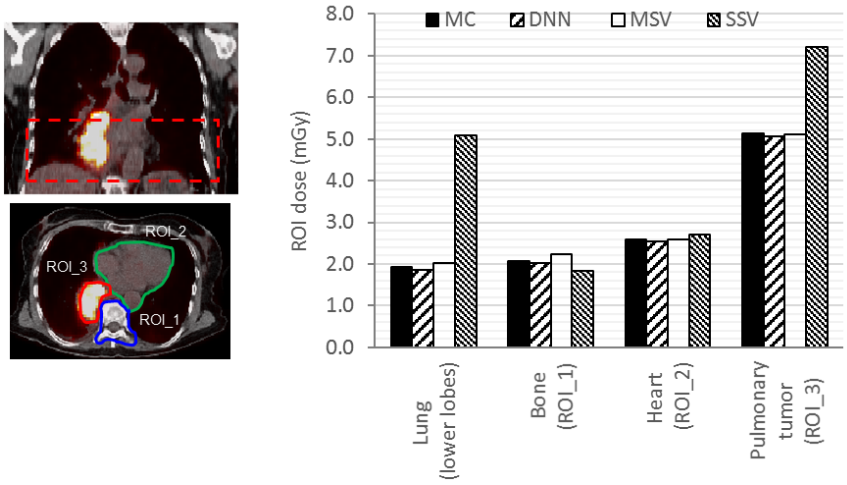


Figure 8. Anatomical region for dose evaluation (top left), axial view of delineated VOIs (bottom left). Average absorbed doses in defined VOIs obtained using DNN, MSV and SSV compared to MC calculations (right).

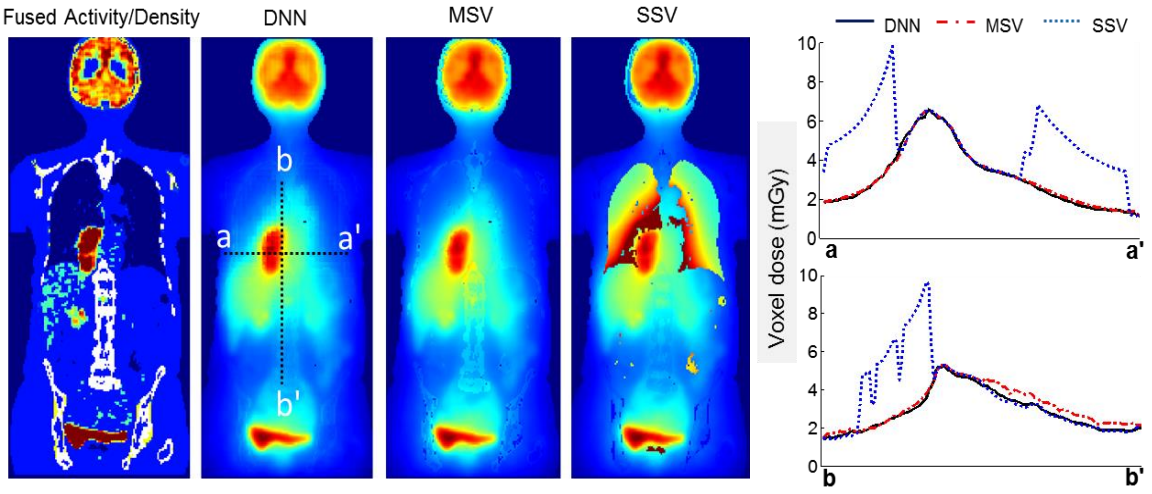


Figure 9. Voxelwise dose maps estimated using DNN, MSV, and SSV along with horizontal and vertical profiles drawn on the coronal view.

IV. Discussion

Despite the paramount importance of personalization in routine clinical setting, this paradigm is still in its infancy, and in the literature, only a few studies have addressed this issue. In this work, we propose a novel methodology to perform personalized radiation dose quantification, which is applicable in various nuclear medicine procedures including diagnostic, therapeutic, and theranostics. The current methodology has been employed in PET imaging dosimetry using 18F-FDG radiopharmaceutical, as the proof of concept. We developed a MC-based whole-body voxel-level dosimetry approach to enable studies that might provide answers to controversies on whether voxel-based dosimetry is superior to the mean absorbed dose approach [53]. Deep learning algorithms have been deployed to solve complex real-life problems by translating the fundamental physics behind the problem into the computer vision domain. In this work, we extended the core idea of the voxel-level MIRD dosimetry formalism by using DNN algorithms to predict medium-specific S-value kernels instead of using a single kernel obtained from deposited energy in a homogenous soft-tissue medium. The size of the kernel was 19.2 cm where the distance from the central voxel to the border is more than the mean free path of annihilation photons (511 KeV). In a kernel of 19.2 cm in three-dimensions representing voxelwise deposited energy in an infinite soft-tissue medium obtained from MC simulations, the ratio of the deposited energy at the border of the kernel to the central source voxel is about 10^{-4} order of magnitude confirming adequate size of the kernel. Although, increasing the size of the kernel up to three mean free paths of annihilation photons from the center of the kernel can improve the accuracy of dose estimation, it would induce considerably longer simulation time. The resolution of the kernel was defined based on regular axial resolution of PET images. The statistical uncertainty of MC simulations was less than 4%. It is obvious from Figure 2 that even in the border of the S-value kernels, the noise level owing to statistical uncertainty is not significant. To benchmark our assumption for extending the reciprocal theory to heterogeneous medium, we simulated a simple geometry consisting of soft-tissue, bone, and lung materials and calculated the deposited energy when the source and target were locally interchanged. The deposited energies calculated using the reciprocal theory were within 5% of those calculated by simulations.

The predicted 3D kernels exhibited good agreement with MC simulations with a MRAE of 4.5%. The DNN predicted S-value kernel underestimates the ground truth as illustrated in the joint histogram analysis. The comparison of the summation of the predicted 3D kernel against the summation of MC S-value kernel, as an index of total energy deposition in the medium, showed an overall 4% underestimation. Since the deposited energy follows an inverse square law with respect to the distance from the source, S-value kernels bear a very broad dynamic range of intensities. Hence, we implemented a nonlinear intensity normalization using a sigmoid function before feeding the kernels into the network. Owing to the non-linear behavior of the sigmoid function, increased prediction errors were observed for certain intensity values after applying inverse sigmoid function. However, the model performed overall much better using non-linear normalization.

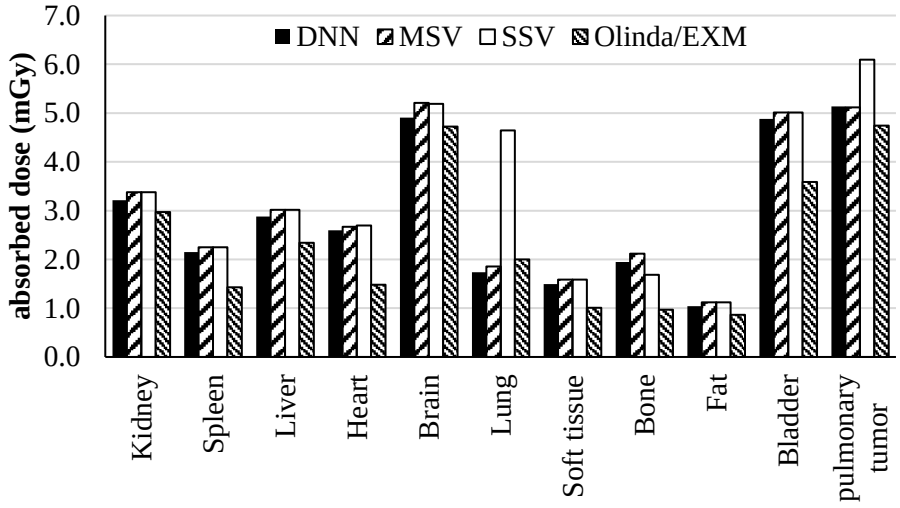


Figure 10. Whole-body organ-level absorbed doses estimated using DNN, MSV, SSV and Olinda/EXM software.

Voxelwise dose comparison between the proposed approach and conventional techniques revealed the limitations of SSV and MSV for internal dosimetry calculations. Since media with higher densities inherently contain more photon interactions, this causes higher energy deposition in the voxels, the summation on S-value kernels of higher density media has higher values compared to those with lower densities. The profiles of absorbed dose showed that SSV overestimated the deposited energy within media with a density lower than soft-tissue (e.g. lung) (Figure 5). This concept applies to densities higher than soft-tissue, such as bone where SSV underestimated the deposited energy. None of the above-mentioned seminal works compared their results with the MSV approach which showed a good agreement with direct MC simulations. However, since this approach relies on the assumption that most absorbed doses are contributed by self-absorption, dose estimation errors are commonly observed at the boundaries of heterogeneous media (see Figure 7 where few data points significantly deviated from the reference), which is not clinically important. Furthermore, MSV underestimated or overestimated the absorbed dose in VOIs with small size depending on the medium density. Figures 5, 8 and 9 confirm that MSV overestimated the results with respect to the ground truth in bones. These errors were predictable since the total deposited energy in the soft-tissue kernel is 54% higher than that in lung kernel with the same size, while this difference is about -34% between soft-tissue and cortical bone kernels. This limitation causes significant errors in absorbed dose estimation for small size lesions in media with different mass densities, e.g. pulmonary nodules, which is a critical issue in targeted radionuclide therapy. In addition, application of the MSV method is restricted to radiotracers with higher positron energy, since taking only self-absorption into account does not fulfill the requirements of accurate internal dosimetry. The Bland-Altman analysis demonstrated the lower bias and variance of DNN against MSV and SSV. The data points of the DNN method beyond the CI correspond to voxels at the boundary of body contour having no impact on dose calculation results. In addition, the data points of MSV method beyond the CI belong to voxels with heterogeneous boundaries, while for SSV three separate regions were formed corresponding to three different media. In nuclear medicine practice, knowledge of organ-scale absorbed dose according to the different radio-sensitivity of organs is required. Olinda/EXM is a commercial software package enabling estimation of organ-level absorbed doses according to the MIRD formalism. For the studied patient, it was observed that organ-level dosimetry leads to underestimation of absorbed dose compared to voxel-level approaches, except the lungs, as a result of ignoring the non-uniformity of organ activity distribution and inter-subject variability of anatomical characteristics (Figure 10). Another limitation of this software is the use of isolated sphere model for tumor dosimetry. This latter assumes that tumors are spheres with unit density and uniform activity distribution and there is no information about the cross-dose from a tumor to other organs or from other organs to a tumor. Because of this limitation, in the case study with a pulmonary tumor, we determined the total number of disintegrations within the lung and tumor as input kinetic parameters of the lung in Olinda/EXM, which led to an overestimation of lung self-absorbed dose by Olinda/EXM. Conversely, the underestimation of tumor dose lies in the fact that only self-absorbed dose is considered in Olinda whereas cross-irradiation is ignored [54, 55]. Absorbed doses in most organs

considered as soft-tissue were almost similar when using MSV and SSV techniques. MSV was able to correct the SSV errors in regions with a density different from soft-tissue.

The importance of accurate patient-specific voxel-scale internal dosimetry is rapidly growing thanks to recent advances in targeted radionuclide therapy and theranostics. Considering the advantages of voxel-level dosimetry in molecular radiotherapy in terms of providing dose indices, such as dose volume histograms, we developed a methodology for voxelwise dosimetry. The execution time for building a whole-body voxel dose map is less than 0.1% of the time required for direct MC simulations. However, the computational time is longer than that of MSV because it has one additional component for inferring the specific S-value kernels. The total computation time for the first step is about 0.7 h using NVIDIA GEFORCE RTX 2080 Ti platform, whereas the required time for the convolution process is about 0.1 h on a 10-core CPU and 32 GB RAM. The results presented in this work demonstrated that MSV provides reasonable accuracy for dose estimation in diagnostic nuclear medicine procedures. However, due to its limitations, it introduces significant uncertainties which might limit its adoption in therapeutic applications. The proposed method is robust and accurate and suitable for direct transfer to other molecular imaging modalities. Its advantages compared to other deep learning-based dosimetry techniques reported in the literature [35, 36] is that it does not require whole-body dose maps for the training step. In addition, a single trained model for a given radionuclide could be employed for all compounds labelled with this radionuclide. Furthermore, the fundamental principles and/or underlying physics of energy deposition have been considered in our model. The latter depends directly on the energy absorption coefficient of the target voxel and the probability of Compton scattering, which depends on the density of the medium. Previous works did not explicitly incorporate in their model Compton scattering and its contribution to the overall absorbed dose. In this regard, deep learning algorithms were employed to predict the absorbed dose map from the density/activity maps using an end-to-end scheme without explicitly modelling the underlying physical principles (Compton scattering and cross-irradiation). More importantly, we developed a simple network with a single input/output channel featuring detailed modelling of the underlying physical interactions, which enables efficient and versatile training of the algorithm with minimal risk of overfitting. Owing to the simple but efficient deep learning-based core of the proposed framework (smaller number of trainable parameters compared to end-to-end image translation), the model provides an accurate and robust solution using a small training dataset.

This work bears inherently some limitations that should be acknowledged, among them the long time required for simulation-based generation of ground truth dose maps. First, the size of S-value kernel is about one mean free path of annihilation photons. Second, extending the reciprocal theory to heterogeneous media is not straightforward. However, we proved the efficacy of the concept using a simple simulation study. Third, the effect of the limited size of the training and validation dataset warrants further investigation. However, a single patient study was presented as a proof of concept. Unlike organ-level dosimetry that is inherently subject-sensitive, estimation of voxel-wise dose distribution based on the voxel-based MIRD formalism is not subject-sensitive since it depends only on physical parameters (S-value kernel, density map and activity distribution). In this context, the accuracy of the results depends only on how the S-value kernels are determined. Let’s consider that SSV performs well in homogenous media, the accuracy of this method is not related to the type of medium or the activity distribution. The accuracy of this method directly depends on the S-value kernels applied for voxel-wise dosimetry. Likewise, the accuracy of the proposed methodology is linked to the accuracy of the specific S-value kernels while it is not dependent on patient-specific anatomy and activity distribution. Hence, in the first step, we evaluated our S-value prediction voxel-by-voxel to assess the accuracy of our approach (Figure 4). Lastly, we only provided a model for 18F, yet our method is extendable to all types of radionuclides/radiotracers where transfer learning can be exploited to obviate the need for regeneration of large ground truth dataset for training the network. In particular, for positron-emitting radiotracers with different positron energies, the generation of the ground truth should be repeated for a kernel size equal to the range of positrons. Since the deposited energy outside the positron range is contributed by the interactions of annihilation photons, for any pure positron-emitting radiotracer, the central part of S-value kernels should be replaced with the center of simulated S-value kernels for 18F generated in the current study.

V. Conclusion

We proposed a unified methodology for patient-specific voxelwise whole-body internal dosimetry using deep learning algorithms. The comparison of the proposed approach with standard of reference MC simulations revealed very good accuracy with a MRAE of 2.6%. Our technique also outperformed conventional voxel-level and organ-level MIRD-based formalisms. Future work will focus on exploiting the current methodology to generate whole-body voxelwise dose maps in few minutes to serve as Monte Carlo-based ground truth datasets. A network with two-channel inputs consisting of density/activity map pairs and one output channel corresponding to voxelwise dose maps obtained from the previous step is then trained to develop a model for straightforward prediction of whole-body dose maps from hybrid images.

Acknowledgments

This work was supported by the Swiss National Science Foundation under grant SNRF 320030_176052; the Swiss Cancer Research Foundation under Grant KFS-3855-02-2016 and Iran's Ministry of Science.

References

[1] M. Baumann *et al.*, "Radiation oncology in the era of precision medicine,"*Nat. Rev. Cancer,* vol. 16, no. 4, p. 234, 2016.

[2] M. G. Stabin, M. T. Madsen, and H. Zaidi, "Personalized dosimetry is a must for appropriate molecular radiotherapy.,"*Med Phys,* vol. 46, no. 11, pp. 4713-4716, 2019.

[3] H. Zaidi and X. G. Xu, "Computational anthropomorphic models of the human anatomy: the path to realistic Monte Carlo modeling in radiological sciences,"*Annu. Rev. Biomed. Eng.,* vol. 9, pp. 471-500, 2007.

[4] W. E. Bolch *et al.*, "MIRD pamphlet No. 17: the dosimetry of nonuniform activity distributions--radionuclide S values at the voxel level. Medical Internal Radiation Dose Committee,"*J Nucl Med,* vol. 40, no. 1, pp. 11s-36s, 1999.

[5] P. B. Johnson, S. R. Whalen, M. Wayson, B. Juneja, C. Lee, and W. E. Bolch, "Hybrid patient-dependent phantoms covering statistical distributions of body morphometry in the US adult and pediatric population,"*Proc. IEEE,* vol. 97, no. 12, pp. 2060-2075, 2009.

[6] A. Akhavanallaf, T. Xie, and H. Zaidi, "Development of a library of adult computational phantoms based on anthropometric indexes,"*IEEE Trans. Radiat. Plasma Med. Sci.,* vol. 3, no. 1, pp. 65-75, 2019.

[7] Y. H. Na, B. Zhang, J. Zhang, P. F. Caracappa, and X. G. Xu, "Deformable adult human phantoms for radiation protection dosimetry: anthropometric data representing size distributions of adult worker populations and software algorithms,"*Phys. Med. Biol.,* vol. 55, no. 13, p. 3789, 2010.

[8] T. Xie, A. Akhavanallaf, and H. Zaidi, "Construction of patient-specific computational models for organ dose estimation in radiological imaging,"*Med Phys,* vol. 46, no. 5, pp. 2403-2411, 2019.

[9] T. Xie and H. Zaidi, "Estimation of the radiation dose in pregnancy: an automated patient-specific model using convolutional neural networks,"*Eur Radiol,* vol. 29, no. 12, pp. 6805-6815, 2019.

[10] M. J. Berger, "Distribution of absorbed dose around point sources of electrons and beta particles in water and other media," National Bureau of Standards, Washington, DC, 1971.

[11] K. S. Kolbert *et al.*, "Implementation and evaluation of patient-specific three-dimensional internal dosimetry,"*J Nucl Med,* vol. 38, no. 2, pp. 301-8, 1997.

[12] H. B. Giap, D. J. Macey, J. E. Bayouth, and A. L. Boyer, "Validation of a dose-point kernel convolution technique for internal dosimetry,"*Phys Med Biol,* vol. 40, no. 3, pp. 365-81, 1995.

[13] H. Zaidi, "Relevance of accurate Monte Carlo modeling in nuclear medical imaging,"*Med Phys,* vol. 26, no. 4, pp. 574-608, 1999.

[14] I. Gardin *et al.*, "Voxeldoes: a computer program for 3-D dose calculation in therapeutic nuclear medicine,"*Cancer Biother. Radiopharm.,* vol. 18, no. 1, pp. 109-15, 2003.

[15] P. Papadimitroulas, G. Loudos, G. C. Nikiforidis, and G. C. Kagadis, "A dose point kernel database using GATE Monte Carlo simulation toolkit for nuclear medicine applications: comparison with other Monte Carlo codes,"*Medical physics,* vol. 39, no. 8, pp. 5238-5247, 2012.

[16] A. E. Besemer, Y. M. Yang, J. J. Grudzinski, L. T. Hall, and B. P. Bednarz, "Development and validation of RAPID: A patient-specific Monte Carlo three-dimensional internal dosimetry platform,"*Cancer Biother. Radiopharm.,* vol. 33, no. 4, pp. 155-165, 2018.

[17] M. Ljungberg and K. S. Gleisner, "3-D Image-Based Dosimetry in Radionuclide Therapy,"*IEEE Trans. Radiat. Plasma Med. Sci.,* vol. 2, no. 6, pp. 527-540, 2018.

[18] A. Dieudonne *et al.*, "Study of the impact of tissue density heterogeneities on 3-dimensional abdominal dosimetry: comparison between dose kernel convolution and direct Monte Carlo methods,"*J Nucl Med,* vol. 54, no. 2, pp. 236-43, 2013.

[19] G. Loudos *et al.*, "A radionuclide dosimetry toolkit based on material-specific Monte Carlo dose kernels,"*Nucl Med Commun,* vol. 30, no. 7, pp. 504-12, 2009.

[20] M. Khazaee Moghadam, A. Kamali Asl, P. Geramifar, and H. Zaidi, "Evaluating the Application of Tissue-Specific Dose Kernels Instead of Water Dose Kernels in Internal Dosimetry: A Monte Carlo Study,"*Cancer Biother. Radiopharm.,* vol. 31, no. 10, pp. 367-379, 2016.

[21] M. S. Lee *et al.*, "Whole-Body Voxel-Based Personalized Dosimetry: The Multiple Voxel S-Value Approach for Heterogeneous Media with Nonuniform Activity Distributions,"*J Nucl Med,* vol. 59, no. 7, pp. 1133-1139, 2018.

[22] I. Shiri *et al.*, "Deep-JASC: joint attenuation and scatter correction in whole-body (18)F-FDG PET using a deep residual network.,"*Eur J Nucl Med Mol Imaging,* p. *in press,* 2020.

[23] H. Xiang, H. Lim, J. A. Fessler, and Y. K. Dewaraja, "A deep neural network for fast and accurate scatter estimation in quantitative SPECT/CT under challenging scatter conditions.,"*Eur J Nucl Med Mol Imaging,* p. *in press,* 2020.

[24] X. Dong *et al.*, "Deep learning-based attenuation correction in the absence of structural information for whole-body PET imaging.,"*Phys Med Biol,* vol. 65, no. 5, p. 055011, 2020.

[25] A. Sanaat, H. Arabi, I. Mainta, V. Garibotto, and H. Zaidi, "Projection-space implementation of deep learning-guided low-dose brain PET imaging improves performance over implementation in image-space.,"*J Nucl Med*, p. *in press*, 2020.
[26] G. Zaharchuk, "Next generation research applications for hybrid PET/MR and PET/CT imaging using deep learning.,"*Eur J Nucl Med Mol Imaging*, vol. 46, no. 13, pp. 2700-2707, 2019.
[27] I. Shiri *et al.*, "Standard SPECT myocardial perfusion estimation from half-time acquisitions using deep convolutional residual neural networks.,"*J Nucl Cardiol*, p. *in press*, 2020.
[28] H. Seo *et al.*, "Machine learning techniques for biomedical image segmentation: An overview of technical aspects and introduction to state-of-art applications.,"*Med Phys*, vol. 47, no. 5, pp. e148-e167, 2020.
[29] M. Mardani, P. Dong, and L. Xing, "Deep-learning based prediction of achievable dose for personalizing inverse treatment planning,"*Int. J. Radiat. Oncol. Biol. Phys*, vol. 96, no. 2, pp. E419-E420, 2016.
[30] D. Nguyen *et al.*, "A feasibility study for predicting optimal radiation therapy dose distributions of prostate cancer patients from patient anatomy using deep learning,"*Sci. Rep.*, vol. 9, no. 1, p. 1076, 2019.
[31] M. Ma, M. K. Buyyounouski, V. Vasudevan, L. Xing, and Y. Yang, "Dose distribution prediction in isodose feature-preserving voxelization domain using deep convolutional neural network,"*Med. Phys.*, vol. 46, no. 7, pp. 2978-2987, 2019.
[32] V. Kearney, J. W. Chan, S. Haaf, M. Descovich, and T. D. Solberg, "DoseNet: a volumetric dose prediction algorithm using 3D fully-convolutional neural networks,"*Phys. Med. Biol.*, vol. 63, no. 23, p. 235022, 2018.
[33] D. Jarrett, E. Stride, K. Vallis, and M. J. Gooding, "Applications and limitations of machine learning in radiation oncology,"*Br J Radiol*, vol. 92, no. 1100, p. 20190001, 2019.
[34] P. Andreo, "Monte Carlo simulations in radiotherapy dosimetry,"*Radiat Oncol*, vol. 13, no. 1, p. 121, 2018.
[35] M. S. Lee, D. Hwang, J. H. Kim, and J. S. Lee, "Deep-dose: a voxel dose estimation method using deep convolutional neural network for personalized internal dosimetry,"*Sci Rep*, vol. 9, no. 1, p. 10308, 2019.
[36] T. I. Götz, C. Schmidkonz, S. Chen, S. Al-Baddai, T. Kuwert, and E. Lang, "A deep learning approach to radiation dose estimation,"*Phys. Med. Biol.*, vol. 65, no. 3, p. 035007, 2020.
[37] Z. Peng *et al.*, "A method of rapid quantification of patient-specific organ doses for CT using deep-learning-based multi-organ segmentation and GPU-accelerated Monte Carlo dose computing,"*Med. Phys.*, p. *in press*, 2020.
[38] M. Karbalaee, D. Shahbazi-Gahrouei, and M. B. Tavakoli, "An Approach in Radiation Therapy Treatment Planning: A Fast, GPU-Based Monte Carlo Method,"*J Med Signals Sens*, vol. 7, no. 2, pp. 108-113, 2017.
[39] X. Jia, P. Ziegenhein, and S. B. Jiang, "GPU-based high-performance computing for radiation therapy.,"*Phys Med Biol*, vol. 59, no. 4, pp. R151-182, 2014.
[40] R. Loevinger, P. Pfalzner, H. Eisenlohr, S. Malo, A. Sanielevici, and J. Nagl, "The Iaea Program in Medical Radiation Dosimetry,"*Ann. N.Y. Acad. Sci.*, vol. 161, no. 1, pp. 158-167, 1969.
[41] M. Cristy, "Applying the reciprocal dose principle to heterogeneous phantoms: practical experience from Monte Carlo studies,"*Phys. Med. Biol.*, vol. 28, no. 11, p. 1289, 1983.
[42] C. Lee, C. Lee, A. P. Shah, and W. E. Bolch, "An assessment of bone marrow and bone endosteum dosimetry methods for photon sources,"*Phys. Med. Biol.*, vol. 51, no. 21, p. 5391, 2006.
[43] J. Seuntjens, W. Strydom, and K. Shortt, "Dosimetric principles, quantities and units,"*Radiation oncology physics: a handbook for teachers and students. Podgorsak EB, editor. Vienna, Austria: IAEA*, 2005.
[44] D. L. Bailey, J. S. Karp, and S. Surti, "Physics and instrumentation in PET," in *Positron emission tomography*: Springer, 2005, pp. 13-39.
[45] W. Li, G. Wang, L. Fidon, S. Ourselin, M. J. Cardoso, and T. Vercauteren, "On the compactness, efficiency, and representation of 3D convolutional networks: brain parcellation as a pretext task,"*International conference on information processing in medical imaging*, pp. 348-360, 2017.
[46] W. Schneider, T. Bortfeld, and W. Schlegel, "Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions,"*Phys. Med. Biol.*, vol. 45, no. 2, p. 459, 2000.
[47] L. S. Waters, "MCNPX user's manual,"*Los Alamos Nat. Lab.*, 2002.
[48] S. Jan, C. Comtat, D. Strul, G. Santin, and R. Trebossen, "Monte Carlo simulation for the ECAT EXACT HR+ system using GATE,"*IEEE Trans. Nucl. Sci.*, vol. 52, no. 3, pp. 627-633, 2005.
[49] N. Zaker, F. Kotasidis, V. Garibotto, and H. Zaidi, "Assessment of Lesion Detectability in Dynamic Whole-Body PET Imaging Using Compartmental and Patlak Parametric Mapping,"*Clin. Nucl. Med.*, vol. 45, no. 5, pp. e221-e231, 2020.
[50] G. Fahrni, N. Karakatsanis, G. Di Domenicantonio, V. Garibotto, and H. Zaidi, "Does whole-body Patlak 18F-FDG PET imaging improve lesion detectability in clinical oncology?,"*Eur Radiol*, vol. 29, no. 9, pp. 4812-4821, 2019.

[51] J. H. Hubbell and S. M. Seltzer, "Tables of X-ray mass attenuation coefficients and mass energy-absorption coefficients 1 keV to 20 MeV for elements Z= 1 to 92 and 48 additional substances of dosimetric interest," National Inst. of Standards and Technology-PL. Gaithersburg, 1995.

[52] M. G. Stabin, R. B. Sparks, and E. Crowe, "OLINDA/EXM: The second-generation personal computer software for internal dose assessment in nuclear medicine.,"*J Nucl Med,* vol. 46, no. 6, pp. 1023-1027, 2005.

[53] C. Chiesa, M. Bardiès, and H. Zaidi, "Voxel-based dosimetry is superior to mean absorbed dose approach for establishing dose-effect relationship in targeted radionuclide therapy,"*Med. Phys.,* vol. 46, no. 12, pp. 5403-5406, 2019.

[54] D. M. Howard, K. J. Kearfott, S. J. Wilderman, and Y. K. Dewaraja, "Comparison of I-131 radioimmunotherapy tumor dosimetry: unit density sphere model versus patient-specific Monte Carlo calculations,"*Cancer Biotherapy and radiopharmaceuticals,* vol. 26, no. 5, pp. 615-621, 2011.

[55] A. Divoli, S. Chiavassa, L. Ferrer, J. Barbet, G. D. Flux, and M. Bardies, "Effect of patient morphology on dosimetric calculations for internal irradiation as assessed by comparisons of Monte Carlo versus conventional methodologies,"*J. Nucl. Med.,* vol. 50, no. 2, pp. 316-323, 2009.

Chapter 10

Personalized brachytherapy dose reconstruction using deep learning

Azadeh Akhavanallaf, Reza Mohammadi, Isaac Shiri, Yazdan Salimi, Hossein Arabi, and Habib Zaidi

Contribution: Study conception and design, methodological development, computer programming, analysis and interpretation of results, manuscript preparation.

Comput Biol Med, Vol. 136, pp 104755 (2021)

Abstract

Background and purpose: Accurate calculation of the absorbed dose delivered to the tumor and normal tissues improves treatment gain factor, which is the major advantage of brachytherapy over external radiation therapy. To address the simplifications of TG-43 assumptions that ignore the dosimetric impact of medium heterogeneities, we proposed a deep learning (DL)-based approach, which improves the accuracy while requiring a reasonable computation time.

Materials and methods: We developed a Monte Carlo (MC)-based personalized brachytherapy dosimetry simulator (*PBrDoseSim*), deployed to generate patient-specific dose distributions. A deep neural network (DNN) was trained to predict personalized dose distributions derived from MC simulations, serving as ground truth. The paired channel input used for the training is composed of dose distribution kernel in water medium along with the full-volumetric density maps obtained from CT images reflecting medium heterogeneity.

Results: The predicted single-dwell dose kernels were in good agreement with MC-based kernels serving as reference, achieving a mean relative absolute error (MRAE) and mean absolute error (MAE) of 1.16±0.42% and $4.2\pm2.7\times10^{-4}$ ($Gy.sec^{-1}$/voxel), respectively. The MRAE of the dose volume histograms (DVHs) between the DNN and MC calculations in the clinical target volume were 1.8±0.86 %, 0.56±0.56 %, and 1.48±0.72 % for D90, V150, and V100, respectively. For bladder, sigmoid, and rectum, the MRAE of D5cc between the DNN and MC calculations were 2.7±1.7 %, 1.9±1.3 %, and 2.1±1.7 %, respectively.

Conclusion: The proposed DNN-based personalized brachytherapy dosimetry approach exhibited comparable performance to the MC method while overcoming the computational burden of MC calculations and oversimplifications of TG-43.

I. Introduction

Brachytherapy is a radiation therapy technique where radiation sources are located at small distances from the tumors, temporarily as in high-dose rate brachytherapy (HDR-BT) or permanently as in low-dose rate brachytherapy. In routine clinical practice, the dose distributions are commonly calculated using a simplified formalism proposed by the American Association of Physicists in Medicine (AAPM) Task Group No. 43 (TG-43) [1] or its updated version (TG-43U1) [2]. In these formalisms, the patient's anatomy is considered as a water-filled homogenous medium. A number of factors that influence the dose distributions, such as tissue and applicator heterogeneities, the finite geometry of the patient, the source-source and source-cable attenuation, and electron contributions to absorbed dose are ignored. Commercial treatment planning systems (TPSs), such as SagiPlan (Eckert & Ziegler BEBIG Co., Germany), attempted to address the limitations of TG-43 in attenuating media like a shield and metallic applicator through multiplication of the dose distribution obtained from the TG-43 model by an analytical attenuation factor. However, the effects of transmission angles and attenuation dependency on distance from the applicator in the region behind the shield (caused by multiple scattering) are ignored. A number of studies have addressed the limitations of TG-43 considered to cause an over/underestimation of the estimated planned dose and consequently treatment evaluation parameters for the clinical target volume (CTV) and organs at risk (OARs) in different disease sites [3-12]. The AAPM published the recommendations of Task Group 186 on model-based dose calculation algorithms (MBDCAs) in brachytherapy beyond the TG-43 formalism [13]. They recommended the collapsed-cone [14], superposition/convolution [15], deterministic solutions using the linear Boltzmann transport equation [16] and Monte Carlo (MC) methods to improve the accuracy of dosimetric calculations in TPS. In MBDCAs, the exact definition of source and applicator geometry inserted within the patient-specific computational model and heterogeneity corrections are implemented into the model [13]. For photon-emitting sources at energies lower than 150 KeV, the predominance of photoelectric interactions makes a large difference in energy absorption coefficients between different tissue types, which necessitates the introduction of heterogeneity corrections in dosimetry calculations. However, the dosimetric impact of tissue heterogeneities and finite patient dimensions for high-energy photon sources of Co-60 and Ir-192 in different treatment sites has been reported to be about 2% difference of dose-volume histogram (DVH) parameters between TG-43 against MC ground truth for the CTV, while these differences exceeded 5% for OARs [17]. Desbiens *et al.* [7] studied the dosimetric impact of medium heterogeneities for Ir-192 in gynecologic HDR-BT using MC simulations. They reported about 1% error on DVH-driven indices by taking into account tissue heterogeneities, whereas they reported that excluding the air pocket and applicator material from DVH calculation produces about 8.7% difference in CTV D90 with respect to TG-43.

Personalized dosimetry is required to improve clinical outcomes while lowering the risk of radiation-induced toxicity by growing recognition of precision medicine as a new paradigm aiming at increasing treatment efficacy. In this context, direct MC simulation is considered the gold standard for dosimetry calculations. However, its heavy computational burden and long execution time made it prohibitive for routine clinical application. More recently, the clinical adoption of deep learning (DL) has been extended into radiation oncology through treatment planning optimization [18-25]. A number of studies assessed knowledge-based automatic treatment planning using deep learning algorithms for external beam radiation therapy to overcome the computational burden of MC-based dose distribution [26-29] for head and neck cancer patients. In our previous work, we developed a framework for patient-specific internal dosimetry, where the core idea of the Medical Internal Radiation Dose Committee (MIRD) formalism was employed by training a physics-informed neural network to predict specific deposited energy kernels in a heterogeneous medium [30]. It was shown that the deep learning-based model outperformed conventional MIRD approaches compared to reference MC simulation. We further extended our work to patient-specific dose distributionning in brachytherapy. An independent work was simultaneously carried out by Mao *et al.* [31] on Ir-192-based HDR-BT dose prediction using deep learning. They designed a modified U-Net to predict the dose distribution considering contoured structures of patients, where 3D dose map obtained from MC simulations serving as ground truth. Considering the current literature, the main contribution of this work is (i) developing a physics-informed DL-based framework through feeding full voxel density map into the network (considering the presence of dense objects, such as metallic applicators, ovoid caps and air pockets); (ii) introducing the volumetric dose map as a whole (taking into account attenuation/ Compton scattering and their contribution to the overall absorbed dose).

In this work, we investigated the potential of predicting brachytherapy dose distributions using a 3D Deep Neural Network (DNN), wherein a two-channel input consists of a density map obtained from CT images along with the deposited energy kernel in water was fed into the network to predict patient-specific planned dose distributions.

II. Materials and methods

Method description

In the first step, we developed a MC-based Personalized Brachytherapy Dosimetry Simulator (PBrDoseSim) wherein CT and DICOM-RT plan dataset are imported into the system to estimate patient-specific dose distribution. In the second step, we employed a DL model to generate personalized dose distributions from patient's CT images and treatment plans (Figure 1). The DL network was designed as an image-to-image regression model to predict the specific energy deposition kernel when the radioactive source is located in the dwell position obtained from the treatment plan. The input data for the training is a two-paired channel composed of dose distribution kernel in water medium along with volumetric density maps. The corresponding output is a 3D dose distribution inspired from MC simulations. Given the input matrices, PBrDoseSim was employed to produce the dose distribution kernel for single-dwell sources considering the radioactive seed with unit activity located in the dwell position. The final dose distribution was reconstructed by superposition of single-dwell-position dose maps through dwell-time-weighted linear combination.

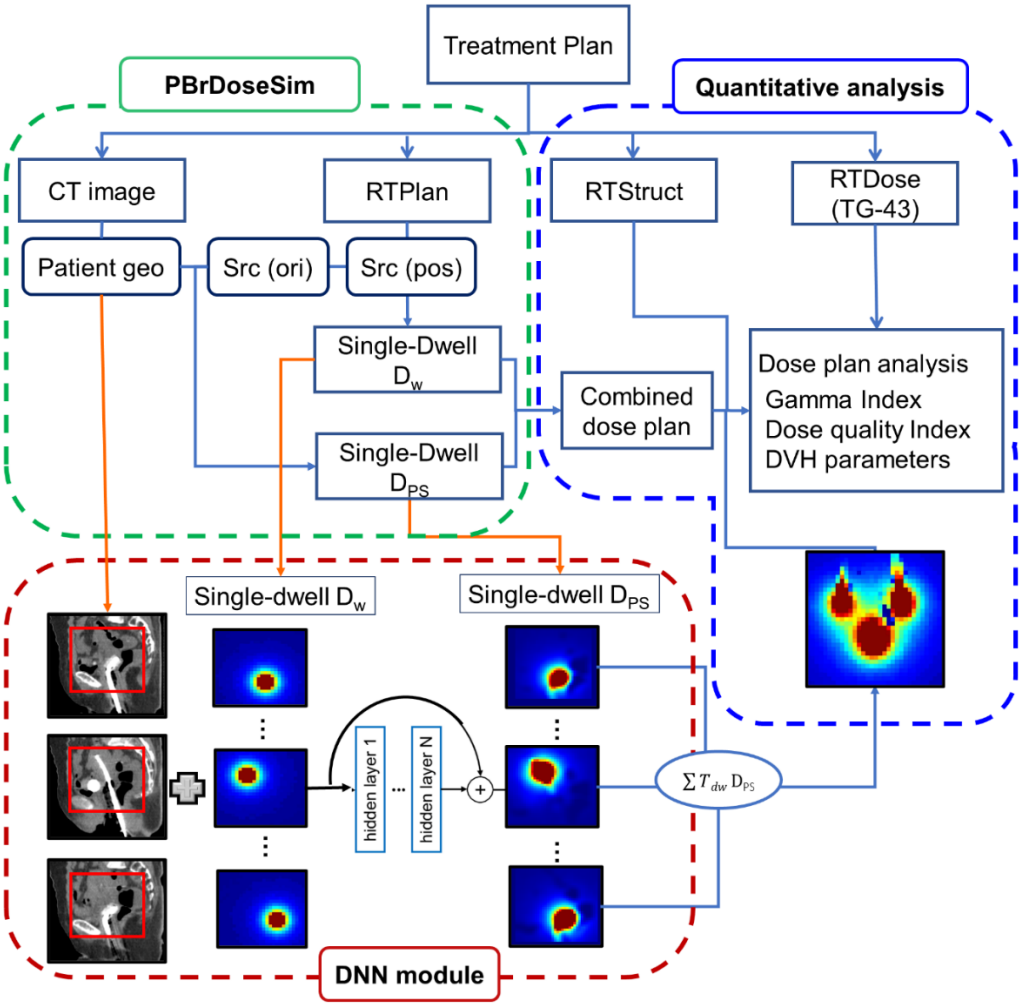


Figure 1. graphical abstract of the dose reconstruction procedure. The green panel shows the MC simulator structure, the red panel represents deep learning-based single-dwell dose kernel prediction, and the blue panel represents post-processing and analysis steps. D_w: single-dwell dose distribution in water, D_{PS}: single-dwell dose distribution in patient-specific geometry. geo: geometry, Src: source, pos: position, ori: orientation.

Data preparation

Data from 78 patients with locally-advanced cervical cancer treated with 3D conformal radiotherapy (45 Gy in 25 fractions) and HDR-BT (Co-60 source, 3/4 fractions and 8.6/7 Gy per fraction) as boost using a fletcher applicator were collected. A total of 2355 single-dwell positions (on average 30 dwell positions per patient) were obtained. Detailed patient demographics are given in Table 1. We randomly divided our patient data into training (70) and external validation (8) datasets. Division by patient class was adopted to avoid sharing the patient geometry information between the training dataset and unseen external validation set.

Table 1. Patients' demographic information.

Patients split	CTV_volume (cm^3)	Pescribed dose-D_{90} (Gy)	Dwell positions (#)			Dwell-pos Sparsity
			Ch. 1	Ch. 2	Ch. 3	
Training (No. 70)	21.94±16.11	8.45±0.85	7.4±3.0	15.7±5.0	6.95±2.0	20.73±3.84
Validation (No. 8)	20.22±8.83	8.62±0.86	7.3±1.1	16.8±3.6	7.7±1.0	20.16±3.32

Ch: applicator channel (Ch. 1: left oviod, Ch. 2: tandem and Ch. 3 right ovoid)
Dwell-pos Sparsity: $\frac{1}{N}\sum_N r_{dwell-position} - \bar{r}$, N: total number of dwell positions

Monte Carlo simulation

An in-house developed software for brachytherapy dose planning based on MCNP transport code was developed [32]. This program was implemented in Matlab and comprises four modules: i) geometry, ii) source definition, iii) simulation parameters, iv) output definition. The geometry module is fed by CT images. This module reads patient's CT images to provide a density map through the conversion of CT Hounsfield units (HUs) into different density classes [33]. In this module, the resolution size is a variable that can be modified by the user. Besides, the densities and corresponding material compositions assigned to each voxel are available to users for modification as an external library. The source module provides the definition of the source in MCNP code consisting of the seed model, particles and energy definition, as well as position and orientation of the source. The source geometry is defined using MCNP microbodies according to the source design. The mode and source energy spectrum are defined based on the radioactive source material. The position and orientation of the seed is extracted from DICOM-RT data (seed orientation is defined based on the two end-to-end dwell positions). The simulation parameters include the number of histories tracked in the model and variance reduction techniques. The output module defines the parameters of the scored MCNP tally and the geometry of the output that can be fine-tuned by the user (Supplemental Table 1).

Dw representing the dose distribution in water was generated through the simulation of specified source seed in an infinite homogenous water medium. Accodrding to TG-43 formalism, heterogeneity correction of metallic applicator was applied on the planned dose by multiplication of an attenuation factor depending on the applicator material and wall thickness.

In this work, density maps consisting of 13 density classes (air, lung, fat, soft tissue, and bone) where values above 100 HU were divided into eight discrete density groups. Afterward, the generated density maps were resampled to 3 mm3 voxel size and were directly imported to the MCNP code, representing a heterogeneous medium of patient's anatomical structures. The material compositions were defined based on Schneider et al. [33]. BEBIG Co-60 HDR source (Model Co60.A86) used in this study was designed according to the company model (Supplemental Figure 1). The mode and energy spectrum of the emitted source particles was defined based on Co-60 (photons with two equal emission probability energy bins of 1.33 MeV and 1.17 MeV). The position and the source orientation were extracted from the DICOM-RT file and used as input to MCNP code considering the original coordinate in the TPS. Metal applicator (Fletcher tandem and oviods, Eckert & Ziegler BEBIG Co., Germany) segmented on CT images was modelled as foreign objects (density of 4.51 gr/cm^3) within the patients. 5 million (5×10^6) particles were tracked in these simulation sets and truncation methods, i.e., energy cut, were used as MCNP variance reduction techniques. Energy deposition mesh tally (type 3) was used in this simulation. 3D dose grids with a size of 34×34×34 voxels and a resolution of $3mm^3$ were designed to score voxelwise energy deposition. To benchmark our simulator, we designed a single dwell position treatment plan in a water sphere of

5 cm radius (Supplemental Table 2) and quantitatively analyzed the planned dose obtained from the TPS against our simulation.

Deep neural network architecture

In this work, a modified ResNet [45] architecture implemented on the TensorFlow platform was employed. The ResNet architecture consists of 20 cascaded convolutional layers with three levels of dilated residual blocks (Supplemental Figure 2). The first convolution layer is cascaded by three residual blocks with a 3×3×3 voxel convolution used to extract low-level features from the data. The next three residual blocks were designed to extract medium-level features using a dilation convolution operation by a factor of 2, whereas the last three residual blocks capture high-level features by a dilation factor of 4. The network was trained to learn heterogeneity correction on D_w according to the information derived from density maps. The optimization of the network was defined based on L2 norm as objective function (OF) in addition to a regularization term of L1 in the following form:

$$OF\ (regularized) = \frac{1}{2}\sum(\hat{y} - y)^2 + \frac{\lambda}{2}\sum w \qquad (1)$$

where y, $\hat{y}$ and λ are ground truth, prediction, and decay factor, respectively. w represents trainable parameters. The following setting was used for the training: optimizer = Adam, learning rate = 0.0001, batch size = 20, decay = 0.00001. Pairs of volumetric density maps representing medium heterogeneity and D_w reflecting the dose distribution around dwell position in homogenous medium as input, and corresponding patient-specific dose map (D_{ps}) as output were fed into the DNN. D_{ps} obtained from PBrDoseSim has a large dynamic range owing to the steep dose gradients with the distance to the source. Hence, it was non-linearly normalized for the sake of effective training of the network. 5% of the training dataset was isolated for validation within the training of the model.

Evaluation strategy

Quantitative analysis of model performance was conducted by evaluating the mean relative absolute error (MRAE), voxelwise mean absolute error (MAE), and kernelwise absolute mean error (AME) between DNN and MC-based single-dwell dose maps. Furthermore, dose distribution quality metrics, i.e. 3D Gamma analysis for multiple criteria (1%/ 3% dose deviation (DD), 3 mm distance-to-agreement (DTA), total plan volume/ 100% isodose volumes), conformity index (COIN), dose homogeneity index (DHI), dose non-uniformity ratio (DNR), and dose-volume histogram (DVH) parameters were investigated as clinically relevant indices. Two groups of indicators consisting of CTV-based indices and OAR-based indices were defined. D_{xx} is the absorbed dose received by xx % of the target volume, whereas V_{xx} is the percentage of the target volume receiving at least xx % of the prescribed dose. D_{xcc} represents the minimum dose received by x cm^3 of an OAR. The distribution of the results was analyzed using Kolmogorov–Smirnov test and pairwise comparison between different methods against reference was analysed using post-hoc-test ($p < 0.05$).

III. Results

The details of PBrDoseSim evaluation in water phantom are summarized in Supplemental C. A single-dwell dose kernel and combined dose distribution along with axial dose profiles are illustrated in Figure 2, where the relative difference between the two profiles are about 0.8% and 1.59%, respectively. The MRAE in a 10×10×10 window around the dwell position ($MRAE_c$) was calculated owing to the large contribution of absorbed doses in proximity to the source position, yielding 1.16±0.42 % difference between DNN and MC results.

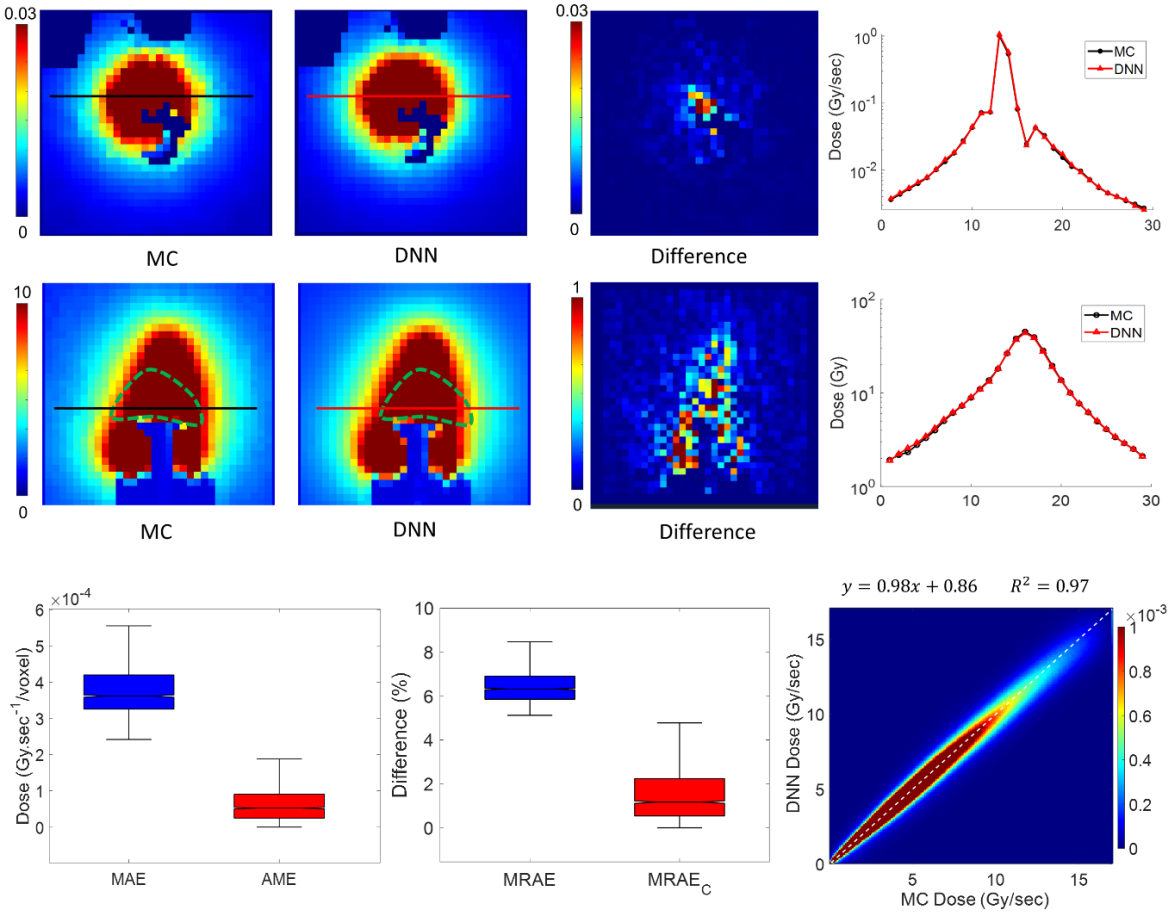


Figure 2. Single-dwell dose kernel (axial view, top panel) and combined dose map (pear-shaped dose distribution, coronal view, pear shape, middle panel) obtained from MC simulations and DNN model along with bias map in absolute unit of Gy and line profiles across the kernels. Comparison of MAE and AME (bottom left), MRAE (%) and MRAEC (%) (bottom middle) obtained within planned dose when using DNN against MC-based approaches. Joint histogram analysis displaying the correlation between the local voxels between predicted dose maps against their corresponding MC-based ground truth (bottom right).

Clinical studies

Dose quality indices obtained from MC simulation through the whole dataset are illustrated in Figure 3. The average value (95% Confidence Interval (CI95%)) of COIN, DNR and DHI indices were calculated as 0.24 (0.15), 0.65 (0.15), and 0.34 (0.15), respectively. The DNN outperformed other approaches by achieving the lowest bias (0.05%) and the smallest variance (0.48%) against MC calculations. Voxelwise gamma analysis in the form of cumulated volume histogram of predicted DNN-based dose distribution compared to MC-based dose map for multiple criteria of all studied cases are presented in Figure 3-bottom. This graph confirms the gamma passing rate (Gamma-value<1) of about 99.9% for all analyzed criteria.

The predicted DNN-based DVH shows a consistent shape with those obtained from the MC method serving as reference (Figure 4). The mean relative absolute error of DVH-driven dose metrics between DNN and D_w approach with respect to MC-based results were 1.4±0.9 % and 2.4±2.1, respectively. While for volumetric metrics, these were 1.05±1.18 % and 2.1±3.2 %, respectively. In the current cervical patient dataset, three organs were delineated as OARs (sigmoid, bladder, and rectum) where D_{5cc} and D_{2cc} metrics are shown in Figure 5. The mean relative absolute error of DVH parameters in OAR regions between DNN and D_w compared to MC simulations resulted in 1.6±1.6 % and 8.7±14.94 %, respectively. According to Kolmogorov–Smirnov test on the current dataset, non-parametric statistical analysis was utilized. The post-hoc comparison confirms that the differences between DVH-driven metrics obtained from DNN (*p-value*= 0.99) and D_w (*p-value*= 0.68) with respect to those from MC-based approach are not statistically significant.

3D gamma analysis for the unseen external validation set is illustrated in Figure 6, which compares the predicted DNN-based dose distribution against MC-based dose map for multiple criteria, namely 1% / 3% DD in 3mm DTA and local normalization, for both total dose distribution volume and 100% isodose volume. The cumulated volume histogram depicts the gamma passing rate (Gamma-value<1) of 99.9% for all analyzing criteria. Furthermore, the maximum intensity projection of 3D gamma maps of validation dataset is shown in Supplemental Figure 5. Quality indices obtained from DNN and D_w dose distributions are depicted along with MC-based as ground truth (Figure 6). The MRAE of COIN and DNR indices between DNN models compared to MC approach were 0.8±0.5 % and 1.8±2.3 % while these differences exceeded 0.8±0.4 % 3.4±4.6 % for D_w model compared to the reference.

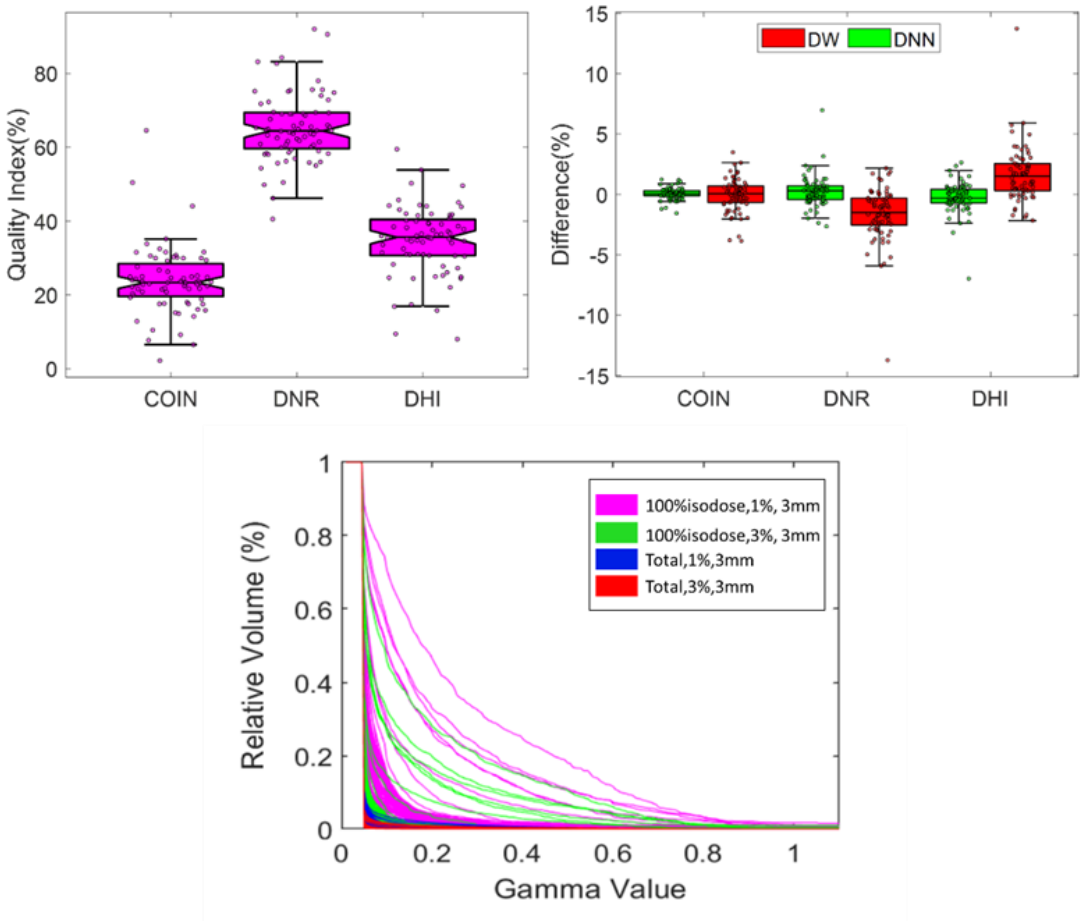


Figure 3. Dose quality metrics of the whole dataset according to the dose distribution obtained from MC-based calculations (top left) along with the differences between quality indices calculated by DNN and Dw models against MC ground truth (top right). Cumulated volume histogram of gamma analysis between dose maps obtained from DNN with respect to MC-based dose distributions for 100%isodose volumes, (1% and 3%) DD, 3mm DTA, [coded as 100%isodose, (1% and 3%), 3mm] and total dose volumes, (1% and 3%) DD, 3mm DTA, [coded as total, (1% and 3%), 3mm] (bottom).

The performance of the proposed DNN model on the external validation in terms of DVH-driven parameters are summarized in Tables 2-3. The MRAE of DVH metrics between DNN and MC was 1.5±0.88 %, 1.8±0.86 %, 1.3±1 %, 0.85±0.43 %, 0.56±0.56 %, 1.48±0.72 %, 0.26±0.38 % for D95, D90, D50, V200, V150, V100 and V50 in the CTV region, respectively. Conversely, the D_w approach compared to MC yielded 2.45±2 %, 2.56 ±1.4 %, 3.9±2.1 %, 2.6±2.5 %, 2.9±1.9 %, 2.5±1.6 %, 0.37±0.55 % for the same metrics, respectively. For D2cc of the bladder, sigmoid, and rectum, the MRAE between DNN and MC method was 3.2±1.9 %, 2.4±1.6 %, 2.5±2 %, respectively. The post-hoc-test revealed no statistically significant differences between the metrics obtained from DNN with respect to the MC-based approach (p-value = 0.82). An intraclass correlation coefficient of 99.8% confirms that the results obtained from the proposed DNN method are in excellent agreement with MC serving as reference.

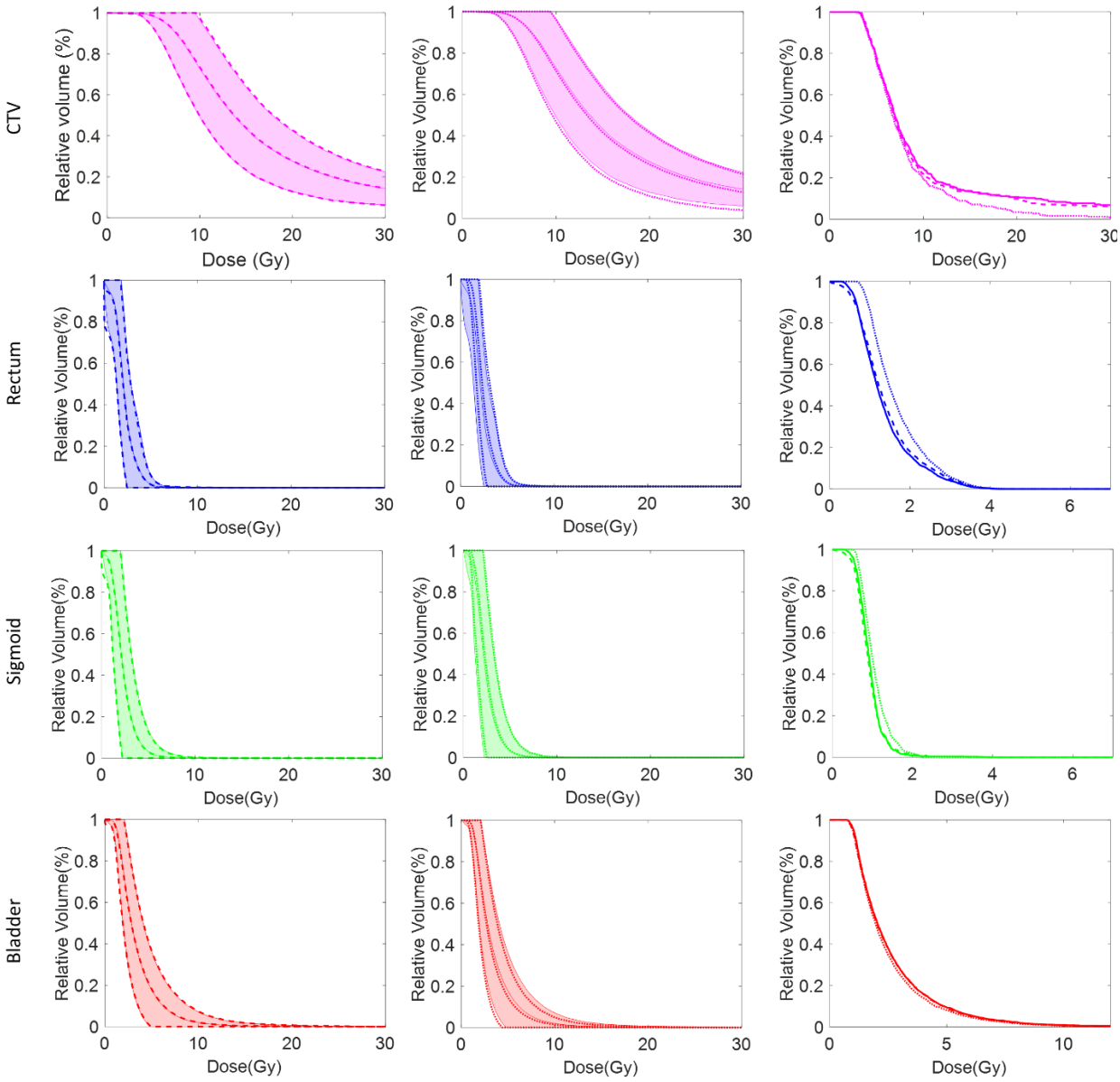


Figure 4. Mean DVH plots along with CI95% (light shaded area) for the study population comparing the proposed DNN (left, dashed line), and Dw (middle, dotted line) plans against MC results (continuous line) are given for CTV (pink), rectum (blue), sigmoid (green) and bladder (red). DVH comparisons between DNN (dashed line), Dw (dotted line) approach with respect to MC-based DVH (continuous line) are plotted for a randomly selected case study (right).

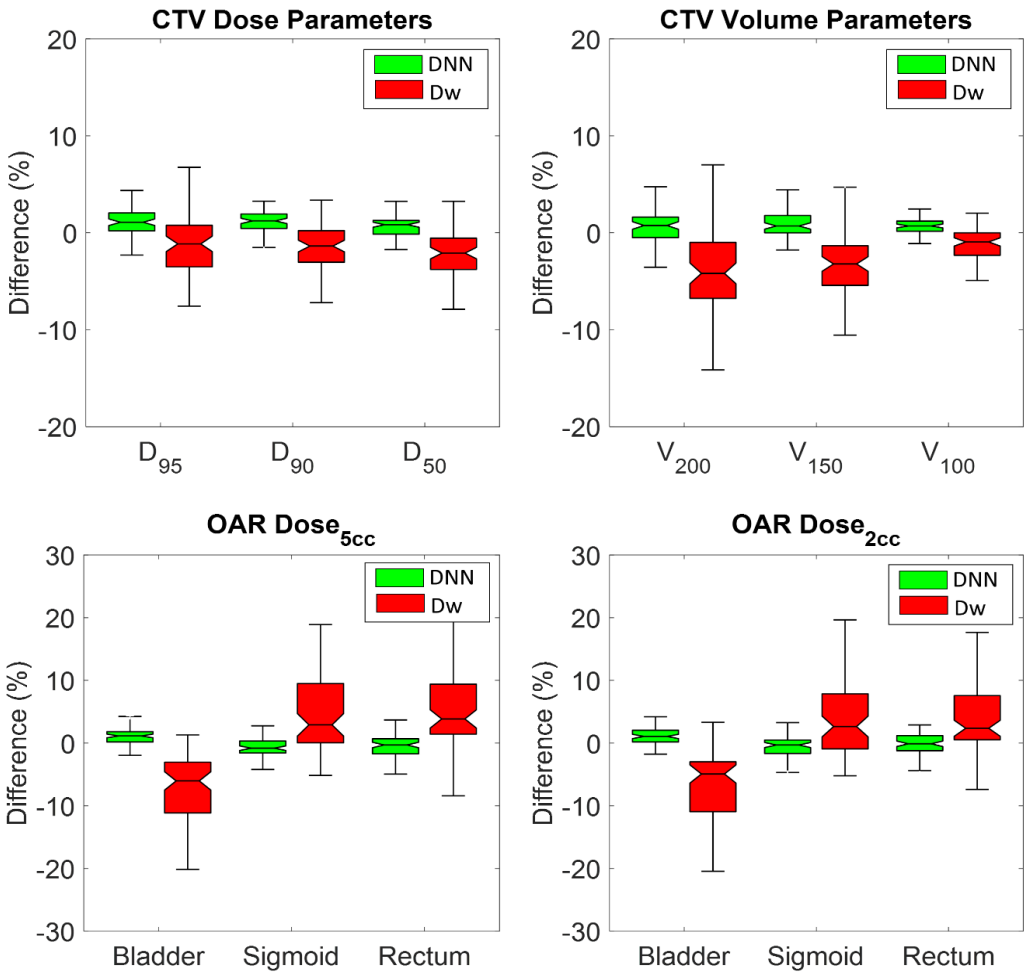


Figure 5. Relative difference of DVH driven dose-volume indices between DNN and Dw method against MC-based approach for CTV (top) and OAR (bottom) for the studied population.

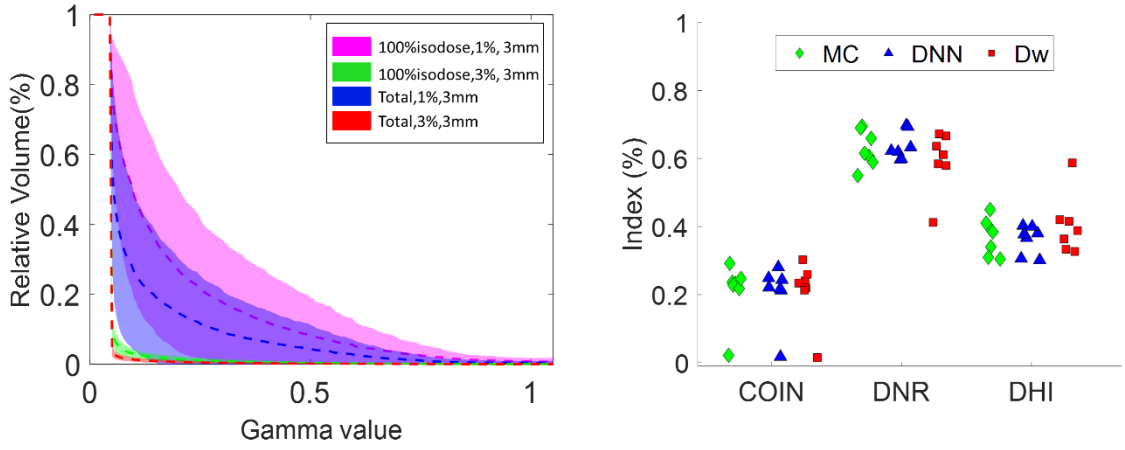


Figure 6. Left panel: Mean gamma volume histogram (dashed) plots along with CI95% (light shade) for the external validation set comparing the proposed DNN model against MC-based dose distribution are given for 100% isodose volumes, (1% and 3%) DD, 3mm DTA [coded as 100% isodose, (1% and 3%), 3mm] and total dose volumes, (1% and 3%) DD, 3mm DTA [coded as total, (1% and 3%), 3mm]. Right panel: Quality indices calculated by MC, DNN and D_w models.

Table 2. Comparison between the proposed DNN and Dw models with respect to MC. The mean value of the DVH-driven dose metrics obtained from MC calculations are provided in absolute unit of Gy. The interquartile difference of dose indices in absolute unit of Gy and mean relative error in percent difference between DNN and Dw model against MC are presented for CTV and OARs.

ROI	DVH metrics	MC (Gy)	DNN vs. MC Diff (Gy)			D_w vs. MC Diff (Gy)			DNN_MC Diff±std (%)	D_w_MC Diff±std (%)
		(mean±CI)	25%	50%	75%	25%	50%	75%		
CTV	D95	5.90± 1.6	-0.08	-0.04	0.07	-0.15	-0.06	0.02	-0.05±1.86	-0.7± 3.3
	D90	6.69± 1.7	-0.14	-0.11	0.10	-0.19	-0.10	-0.06	-0.53± 2.11	-1.7±2.5
	D50	12.12± 2.8	-0.20	-0.10	-0.02	-0.61	-0.46	-0.31	-0.90± 1.45	-3.9±2.1
Bladder	D5cc	5.9±1.4	-0.08	0.05	0.25	-0.63	-0.37	-0.12	1.61±3	-5.93±3.5
	D2cc	7.56±1.66	-0.11	0.20	0.37	-0.54	-0.44	-0.20	1.92±3.46	-5.91±4.3
Sigmoid	D5cc	2.34±1.36	-0.11	-0.04	0.01	-0.01	0.06	0.15	-1.43±2	3.29±6.5
	D2cc	2.94±1.67	-0.11	-0.04	0.01	-0.01	0.14	0.21	-0.80±3	4.23±6
Rectum	D5cc	3.21±0.53	-0.09	-0.04	0.00	0.04	0.09	0.15	-0.65±2.87	2.87±3.9
	D2cc	3.91±0.61	-0.14	-0.03	0.03	-0.04	0.02	0.11	-1.32±3.1	0.84±3.3

Table 3. Mean value of the DVH-driven volume metrics obtained from MC calculations provided in absolute unit of cm^3 and relative percent of CTV volume. The interquartile difference of volume indices in (%) between the DNN and D_w models against MC is presented for the CTV.

ROI	DVH volume metrics	MC Reference (cm^3)	MC Reference (%)	DNN vs. MC (%)			D_w vs. MC (%)		
		(mean±CI)	(mean±CI)	25%	50%	75%	25%	50%	75%
CTV	V200	6.36± 3	30.8± 10.8	-0.83	-0.65	0.35	-3.71	-2.04	-0.81
	V150	9.73± 4.7	46.7±16.3	-0.83	-0.27	0.00	-4.55	-2.78	-1.11
	V100	15.29±7.5	72.7±22.8	-1.63	-1.02	0.40	-3.72	-2.44	-0.77
	V50	19.81±9.2	96.3±8.5	-0.10	0.00	0.09	-0.24	0.00	0.19

IV. Discussion

Accurate calculation of the absorbed dose delivered to the tumor and specific surrounding OARs enables to maximize the treatment gain factor, which is the main advantage of brachytherapy over external beam radiotherapy. In this study, we developed a unified MC-based dosimetry platform enabling the transition from TG-43 to MBDCAs approach while covering the limitations of analytical models in MBDCA [34, 35]. We designed a novel DNN architecture to solve a complex problem, i.e. Boltzmann transport equation, by translating the underlying fundamental physics of particle interactions into the computer vision domain.

We prepared a dataset of 78 cervical cancer patients from which only one treatment session was chosen to increase anatomical variabilities. The simulator was fed by with patient CT images and treatment plans providing patient-specific geometry and radioactive source position and orientation within the patient. 10 cm dose grid's size was selected based on the steep dose gradient with radial distance (inverse square law) [36] that covers the desired volume of interest for clinical dose evaluation. To confirm this assumption, isodose contours were calculated at the boundaries of selected dose grids for the whole dataset corresponding to 15 % (±4%) of the prescribed dose. In addition, none of the evaluated metrics, i.e. DVH-driven indices were affected by this kernel size.

The predicted 3D dose kernels from DNN exhibited good agreement with MC serving as reference confirmed by joint histogram analysis (Figure 2). However, MRAE of 6.3% in the full planning volume can be attributed to the larger statistical uncertainty associated with MC simulations in far distances from the source. Voxel dose difference obtained from D_w model compared to MC was calculated with MRAE of about 13±3 % while $MRAE_c$ exceeds 31±6 % in the vicinity of dwell position (Supplemental Figure 4). Gamma analysis shows that at least 99.99% of points passed all criteria through the whole dataset (3 mm DTA was restricted by voxel size of the dose grids). In terms of DVH indices, DNN shows a comparable performance against MC calculations with an average

relative bias of 0.7 (±1.4) % in the CTV volume. While, Dw showed an overall negative bias (-2.2 %) and higher variance (3.34 %) against MC results (Figure 5). For OARs, the inserted balloon within the bladder caused a considerable underestimation of absorbed dose obtained from Dw, while DNN was trained to correct for its effect. For the sigmoid and bladder, there is low bias since OARs are almost composed of soft tissues and not located in the high dose regions, while the high variance is mainly caused by the presence of air pockets ignored in Dw models (low dose area in Figure 4). External validation revealed good generalizability of the model with an average bias of 0.49 (±1.8)% in CTV dose indices with respect to the MC ground truth (details on case study interpretations are presented in Supplemental F). Mao *et al.* [31] reported the same magnitude of relative error (prostate: CTV D90=0.73%, OAR D2cc~1.1%; cervix: CTV D90=1.7%, OAR D2cc~ 2%) between their DL model against MC as ground truth using Ir-192. Although, they reported slightely less error for quantitative indices in their original model evaluation (prostate), they did not provide any evaluation on their full dose distribution compared to ground truth. Furthermore, they did not provide any baseline comparison, i.e. against Dw, thus, direct comparison of different models is not fair/insightful (owing to different dataset). One of the major limitation of their model is that it relies on contoured structures rather than original density map obtained from CT images. Therefore, it is unable to account for dense materials such as metallic applicator, ovoid caps, CT contrast agents, air pockets, etc, while our framework was based on a realistic physical model considering original voxel density maps as input. Furthermore, they used modified U-Net architecture composed of encoder-decoder (down and up sampling/polling layers) that converts original images to feature space by losing image resolution that would be issue in the presence of small heterogenieties. while in our proposed modified ResNet algorithm, image size is steady through all layers without losing image resolution. We modeled the fundamental principles of energy deposition [37] through fully volumetric dose map rather than patchwise learning. However, previous studies on Dl-based dose prediction did not explicitly incorporate compton scattering in their network and ignored its contribution to the overall dose distribution [31, 38, 39].

Overall, the DNN model outperformed TG-43-based approach in terms of heterogeneity correction in clinically relevant parameters of the HDR-BT planned dose. The required time for prediction of a combined dose (~30 dwell positions) was about 0.6 sec (2080TI GPU, Xeon 2.30) compared to 540 min (10 core CPU, 64-GB RAM). This work bears some limitations that should be acknowledged. First, the choice of dose grid size/ resolution and particle histories were restricted by the long simulation time and limited GPU memory for DL training process. Source position/ orientation within the voxel is affected by course grid resolution, however, it has a local effect (first vicinity voxel) on deposited energy distribution and do not impact DVH-derived parameters. Second, the effect of the limited size of the training and validation dataset warrants further investigation. Third, we only provided a model for cervical HDR-BT using Co-60. Yet, this methodology is extendable to all types of brachytherapy treatments and different disease sites, where transfer learning can be exploited to obviate the need for a large ground truth dataset for model training.

It is worth highlighting the potential opportunities and challenges in the utilization of deep learning into brachytherapy personalized dose distributionning. In this context, DL algorithms can provide a solution for fast personalized dosimetry without compromising the accuracy. One of the challenges that DL can address is the construction of patient-specific computational models using structural images. It highly impacts the accuracy of MC-based dose calculation. Furthermore, it can be directly deployed for construction of planned dose for verification of clinical TG-43 dose distributions, inverse planning and treatment outcome prediction.

V. Conclusion

We developed a unified pipeline for MC-based dosimetry in HDR-BT that has been used to provide an accurate set of MC simulations on a large retrospective cohort. We further developed a DNN model to provide an alternative solution for accurate personalized dose distributionning in brachytherapy to overcome the computational burden of MC simulations. The proposed algorithm achieved good agreement with MC calculations while outperforming the conventional TG-43-based formalism. Future work will focus on extending the core idea to different radioactive seeds for various disease sites.

Acknowledgements

This work was supported by the Swiss National Science Foundation under grant SNRF 320030_176052, the Private Foundation of Geneva University Hospitals under grant RC-06-01 and Iran's Ministry of Science.

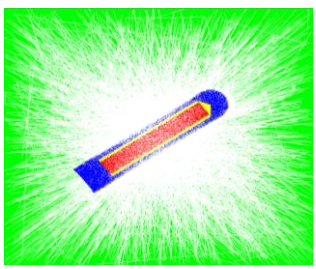
Supplemental material

Monte Carlo-based dose simulation

A unified MC-based dose calculation engine, PBrDoseSim, was developed for personalized brachytherapy based on MCNP radiation transport code. The simulator has the capability of modifying all modules, e.g. source model, applicator model, tally score definition, and simulation setting depending on the application, supporting different types of input images, e.g. DICOM, NIFTY, NRRD, etc, and modifying material composition and density of the patient-specific computational model.

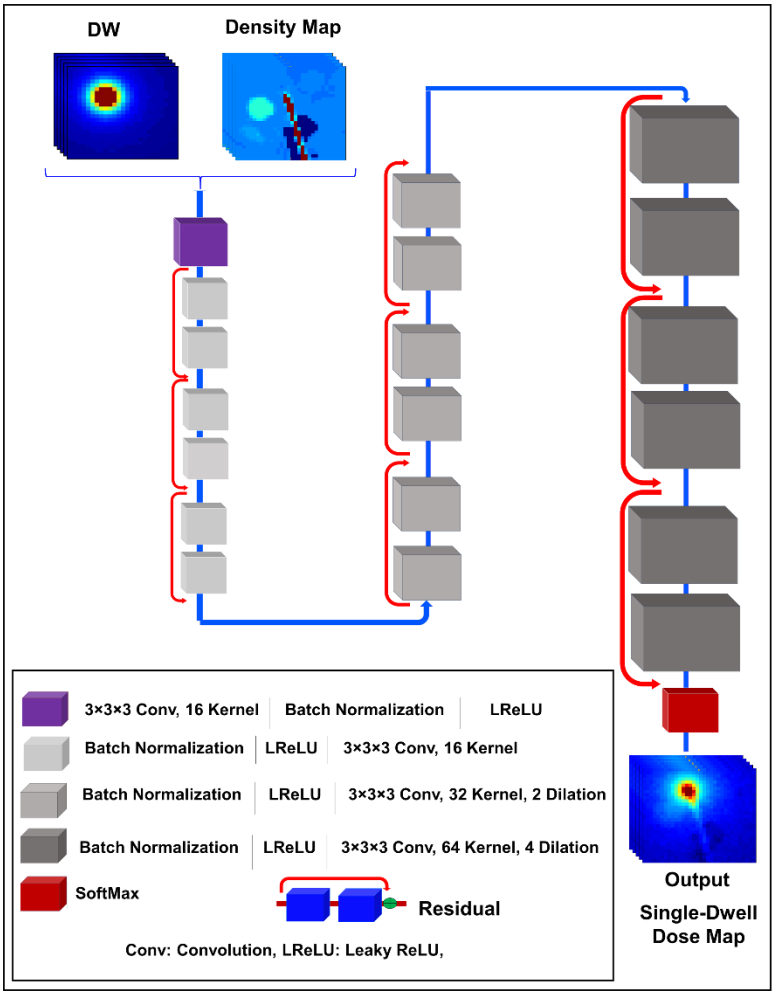
Supplemental Table 4. Monte Carlo simulations description as recommended by TG-268 task group.

Item name	Description
Code, version	MCNPX (version 2.6)
Validation	Validated against phantom study
Timing	18 mins on a 10 Intel core node from 3.7 GHz processor.
Geometry	Patient-specific voxelwise computational model obtained from CT images segmented by multi-level thresholding (Voxel resolution: $(3mm)^3$). Phantom material: mass density obtained from HU values (CT images) and assigning material composition from [40]
Source description	BEBIG 60Co HDR source (model Co60.A86). Energy spectrum of 60Co was defined as photon-emitting source with two equal emission probability energy bines of 1.33 MeV and 1.17 MeV. Source positions was extracted from RTplan and source orientation was derived from the direction of two consecutive dwell positions (Supplemental Figure 1).
Cross sections	ENDF/B-VI Release 8 Photoatomic Data.
Transport parameters	Energy Deposition Mesh Tally (type 3), Tally scores energy deposition data in which the energy deposited per unit volume from all particles is included (Electron/ photon transport).
Variance reduction	Energy cut-off in 10 KeV electron and photon particles
H**istories** (**statistical uncertainty**)	5×10^6 photon histories. Statistical uncertainty per particle history (type A): <0.05 for single-dwell simulation in 1 Gy isodoses (10% of maximum prescribed isodose).
Postprocessing	Convert to voxel dose in absolute unit of Gy by multiplication in a constant factor (UC×Dwell_time×gamma yield×ReferenceAirKermaRate/306). UC: unit conversion in Gy.



Supplemental Figure 7. BEBIG Co-60 HDR source (model Co60.A86) collision score simulation within a homogeneous water medium.

Network architecture



Supplemental Figure 8. Schematic diagram of the ResNET architecture.

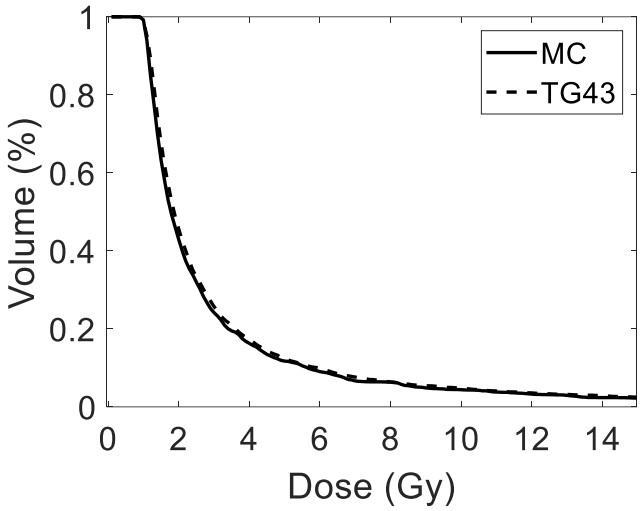
Monte Carlo simulator evaluation using water phantom

The details of the developed MC simulator are summarized in Supplemental Table 1 according to the recommendations of TG-268 [41]. The benchmark study performed on a 5 cm radius water sphere with a single central source yielded a gamma index (3%/3 mm) of 0.9999 between the simulated dose map and TG-43-based planned dose. DVH comparison between the two methods is shown in Supplemental Figure 3. The absolute percent difference between DVH parameters, i.e., D95, D90, V200, V150, V100, V50, was zero except for D90 where it was about 3% (likely caused by the statistical uncertainty of MC results, < 5%).

RTplan parameters for irradiation of a spherical water phantom while the source was located in the center of the phantom with a single dwell position is illustrated in Supplemental Table 2. The DVH plots are shown in Supplemental Figure 3.

Supplemental Table 5. RTPlan parameters of water phantom dose planning.

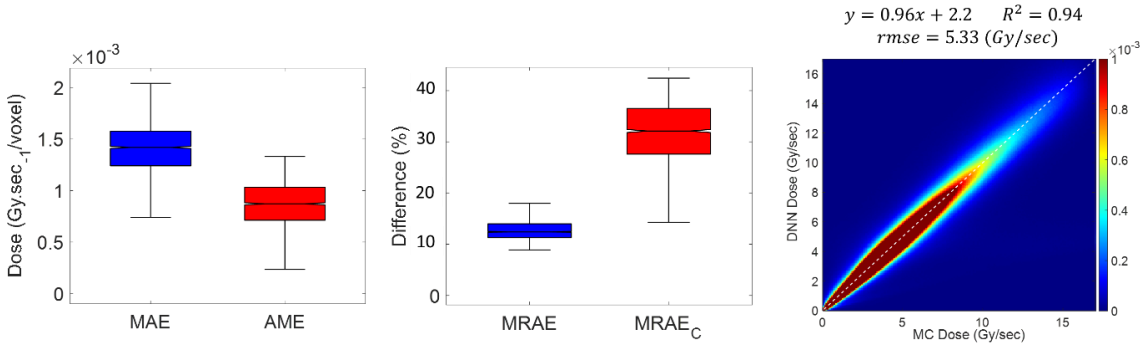
RT plan parameters	
Source model	BEBIG 60Co (Co0.A86)
Size of CTV	523 cm^3
Dwell positions	1 (central)
Dwell time	820 sec
Prescribed dose	1.2 Gy



Supplemental Figure 9. DVH plot for dose plan estimated by MC versus dose plan provided by Sagiplan software based on TG43 formalism.

D_w model performance

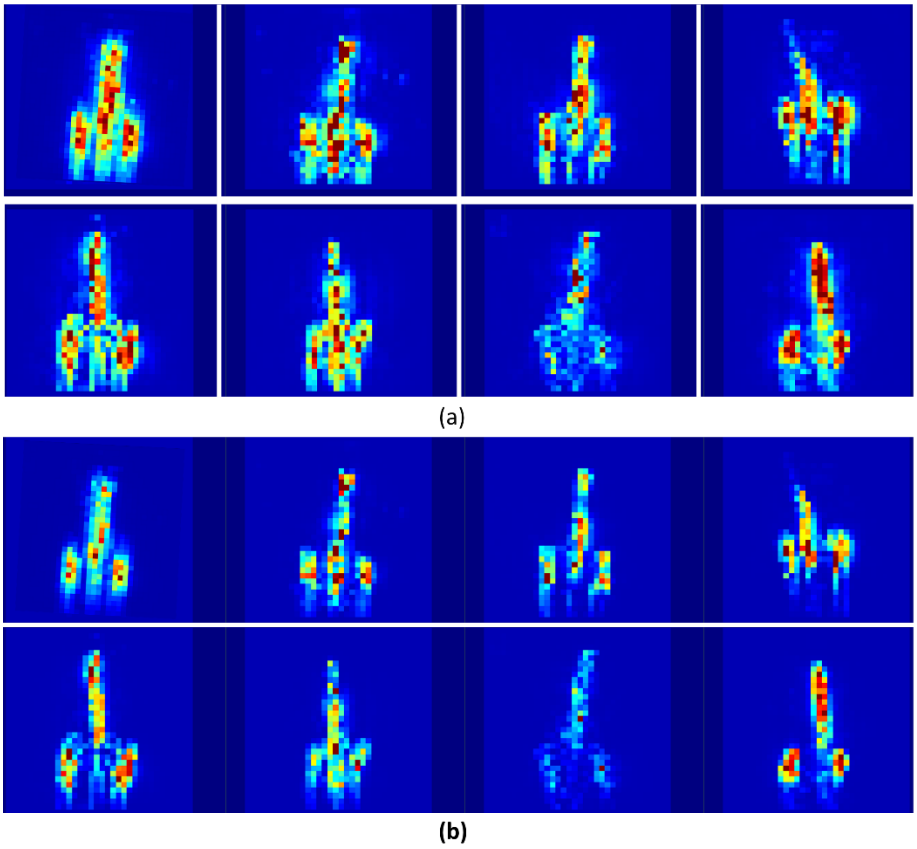
The analysis revealed that D_w resulted in higher bias but lower variance in dose calculation as a result of ignoring tissue heterogeneity. Unlike the DNN model, D_w underestimated the deposited energy in the vicinity of the dwell position owing to ignoring the applicator's material. We further analyzed tissue-specific voxel dose differences. According to the dosimetric principles, the deposited energy in a voxel depends on the energy fluence multiplied by the mass energy-absorption coefficient of the medium [37], wherein both factors are affected by medium density. Overall, it was observed that D_w overestimated the deposited energy in fat and soft-tissue whereas it underestimated the deposited energy in tissues with density higher than soft-tissue. The DNN model was trained to learn heterogeneity corrections to the medium density.



Supplemental Figure 10. Comparison of (left) MAE and AME, (middle) MRAE (%) and $MRAE_C$ (%) obtained in dose plan regions and a window surrounding dwell source position when using D_w against MC-based dosimetry approaches. (right) Voxelwise joint histogram plot depicting the correlation of predicted single-dwell dose maps with respect to their corresponding MC-based ground truth dose kernels.

Maximum Intensity Projection (MIP) of 3D gamma analysis for validation set

Gamma indices were analyzed for multiple volume of interests (total dose volume and 100% isodose volume) and DD factors (3% and 1%) while the setting of local normalization and 3mm DTA was constant. As expected, gamma passing rate in 3% DD is higher than 1% DD because of a larger dose gradient criterion.



Supplemental Figure 11. MIP of 3d gamma analysis of dose grids comparing DNN against MC ground truth on the validation set for (a) 1% DD and (b) 3% DD.

Case study interpretation

We compared the performance of our method against TG-43. However, in the case of high-energy photon emitting brachytherapy sources, such as Co-60, improving water-based dose kernels is challenging since the dosimetric impact of MBDCAs in this energy range is small (on average less than 5%) mainly owing to the predominance of Compton scatter conditions, which is less influenced by the heterogeneity of media. Furthermore, in this work, we focused on cervix site assumed to consist of soft-tissue media where no substantial density differences are expected for normal patients (without metallic implants). Therefore, the trained model mostly corrects the effect of metallic applicator, presence of fat, inserted balloon within bladder and air packets. Our dataset was composed of 78 patients, where in about 12 cases, at least one of the DVH-driven parameters obtained from Dw yielded a relative error higher than 5% against MC-based metrics. We performed a meta-analysis of these patients' information to address the limitations of Dw. It was concluded that when the size of CTV is small and the applicator is not excluded from the CTV volume, the effect of applicator heterogeneity distorts the DVH plot and underestimates the absorbed dose. The D50, V200 and V150 metrics are mostly affected since these parameters represent the DVH behavior in high dose areas. In a few cases where the CTV was large and partially located in the vicinity of ovoids, a distortion of V200 and V150 was observed owing to the ovoid heterogeneity effects. TG-43-based planned dose for a patient with a large metal implant in the femur underestimated all DVH parameters by a factor of 7% on average. Air packets around the applicator that were not excluded from the CTV contour also affected the DVH plot. In the external validation set, DNN outperform TG-43 based algorithm in almost all quantitative metrics.

A transition phase from TG-43 to MBDCA using radiological images seems feasible in clinical setting using some simplified corrections, such as heterogeneity correction according to the voxelwise energy absorption coefficients and excluding air pockets and applicator from DHV plots [7]. Although, accurate patient-specific dose planning requires detailed knowledge of material composition of the irradiated tissues, applicator characteristics, shield heterogeneity, etc. MC methods are deemed the gold standard. In addition, the accuracy of MC-based dose planning is highly dependent on the accuracy of patient-specific computational model obtained from radiological images that is fed into the MC simulator. In clinical radiation oncology, personalized computational phantoms are mostly constructed through conversion of HUs to density maps and assigning an elemental tissue according to the density [33]. This process is prone to error owing to photon starvation and the impact of reconstruction algorithms. In this work, we observed metallic artifacts in tissues around the applicator, which led to a higher density in the area surrounding the applicator and consequently to a higher deposited energy. We compared the mean absorbed doses in the CTV which showed a considerable underestimation of Dw against MC-based mean absorbed dose (on average 7%) mainly caused by the effect of the applicator and the metal artifact in the applicator surrounding area. This underestimation was more likely observed in cases with small size of CTV where the applicator was not excluded.

References

[1] R. Nath, L. L. Anderson, G. Luxton, K. A. Weaver, J. F. Williamson, and A. S. Meigooni, "Dosimetry of interstitial brachytherapy sources: recommendations of the AAPM Radiation Therapy Committee Task Group No. 43. American Association of Physicists in Medicine,"*Med Phys,* vol. 22, no. 2, pp. 209-34, 1995.

[2] M. J. Rivard *et al.*, "Update of AAPM Task Group No. 43 Report: A revised AAPM protocol for brachytherapy dose calculations,"*Med Phys,* vol. 31, no. 3, pp. 633-74, 2004.

[3] J. Mason, B. Al-Qaisieh, P. Bownes, A. Henry, and D. Thwaites, "Investigation of interseed attenuation and tissue composition effects in (125)I seed implant prostate brachytherapy,"*Brachytherapy,* vol. 13, no. 6, pp. 603-10, 2014.

[4] H. Afsharpour *et al.*, "Influence of breast composition and interseed attenuation in dose calculations for post-implant assessment of permanent breast 103Pd seed implant,"*Phys Med Biol,* vol. 55, no. 16, pp. 4547-61, 2010.

[5] G. Landry *et al.*, "Sensitivity of low energy brachytherapy Monte Carlo dose calculations to uncertainties in human tissue composition,"*Med Phys,* vol. 37, no. 10, pp. 5188-98, 2010.

[6] T. Brun and E. Torfeh, "EP-1796: Dosimetric comparison between TG43/TG186 algorithms and manual/inverse optimization in brachytherapy,"*Radiotherapy and Oncology,* vol. 123, p. S986, 2017.

[7] M. Desbiens *et al.*, "Monte Carlo dosimetry of high dose rate gynecologic interstitial brachytherapy,"*Radiother Oncol,* vol. 109, no. 3, pp. 425-9, 2013.

[8] K. A. Gifford, G. P. Fonseca, S. Thrower, and F. Verhaegen, "Comparison of Two Commercially Available Model Based Dose Calculation Algorithms and TG-43 for APBI Brachytherapy,"*Brachytherapy,* vol. 17, no. 4, p. S32, 2018.

[9] D. Jacob, M. Lamberto, L. DeSouza Lawrence, and F. Mourtada, "Clinical transition to model-based dose calculation algorithm: A retrospective analysis of high-dose-rate tandem and ring brachytherapy of the cervix,"*Brachytherapy,* vol. 16, no. 3, pp. 624-629, 2017.

[10] K. Zourari, T. Major, A. Herein, V. Peppa, C. Polgár, and P. Papagiannis, "A retrospective dosimetric comparison of TG43 and a commercially available MBDCA for an APBI brachytherapy patient cohort,"*Phys Med,* vol. 31, no. 7, pp. 669-76, 2015.

[11] G. Anagnostopoulos, D. Baltas, E. Pantelis, P. Papagiannis, and L. Sakelliou, "The effect of patient inhomogeneities in oesophageal 192Ir HDR brachytherapy: a Monte Carlo and analytical dosimetry study,"*Phys Med Biol,* vol. 49, no. 12, pp. 2675-85, 2004.

[12] E. Poon, J. F. Williamson, T. Vuong, and F. Verhaegen, "Patient-specific Monte Carlo dose calculations for high-dose-rate endorectal brachytherapy with shielded intracavitary applicator,"*Int J Radiat Oncol Biol Phys,* vol. 72, no. 4, pp. 1259-66, 2008.

[13] L. Beaulieu *et al.*, "Report of the Task Group 186 on model-based dose calculation methods in brachytherapy beyond the TG-43 formalism: current status and recommendations for clinical implementation,"*Med Phys,* vol. 39, no. 10, pp. 6208-36, 2012.

[14] A. Ahnesjö, "Collapsed cone convolution of radiant energy for photon dose calculation in heterogeneous media,"*Med Phys,* vol. 16, no. 4, pp. 577-92, 1989.

[15] K. Zourari *et al.*, "Dosimetric accuracy of a deterministic radiation transport based 192Ir brachytherapy treatment planning system. Part I: single sources and bounded homogeneous geometries,"*Med Phys,* vol. 37, no. 2, pp. 649-61, 2010.

[16] H. Hedtjärn, G. A. Carlsson, and J. F. Williamson, "Accelerated Monte Carlo based dose calculations for brachytherapy planning using correlated sampling,"*Phys Med Biol,* vol. 47, no. 3, pp. 351-76, 2002.

[17] I. Fotina, K. Zourari, V. Lahanas, E. Pantelis, and P. Papagiannis, "A comparative assessment of inhomogeneity and finite patient dimension effects in (60)Co and (192)Ir high-dose-rate brachytherapy,"*J Contemp Brachytherapy,* vol. 10, no. 1, pp. 73-84, 2018.

[18] H. H. Zhang, W. D. D'Souza, L. Shi, and R. R. Meyer, "Modeling plan-related clinical complications using machine learning tools in a multiplan IMRT framework,"*Int J Radiat Oncol Biol Phys,* vol. 74, no. 5, pp. 1617-26, 2009.

[19] L. V. van Dijk *et al.*, "Improving automatic delineation for head and neck organs at risk by Deep Learning Contouring,"*Radiother Oncol,* vol. 142, pp. 115-123, 2020.

[20] S. Kazemifar *et al.*, "Segmentation of the prostate and organs at risk in male pelvic CT images using deep learning,"*Biomedical Physics & Engineering Express,* vol. 4, no. 5, p. 055003, 2018.

[21] Z. Zhu, E. Albadawy, A. Saha, J. Zhang, M. R. Harowicz, and M. A. Mazurowski, "Deep learning for identifying radiogenomic associations in breast cancer,"*Comput Biol Med,* vol. 109, pp. 85-90, 2019.

[22] V. S. Parekh and M. A. Jacobs, "Deep learning and radiomics in precision medicine,"*Expert Rev Precis Med Drug Dev,* vol. 4, no. 2, pp. 59-72, 2019.

[23] Y. Peng *et al.*, "Magnetic resonance-based synthetic computed tomography images generated using generative adversarial networks for nasopharyngeal carcinoma radiotherapy treatment planning,"*Radiother Oncol*, vol. 150, pp. 217-224, 2020.
[24] A. Yonekura, H. Kawanaka, V. B. S. Prasath, B. J. Aronow, and H. Takase, "Automatic disease stage classification of glioblastoma multiforme histopathological images using deep convolutional neural network,"*Biomed Eng Lett*, vol. 8, no. 3, pp. 321-327, 2018.
[25] J. Fan, J. Wang, Z. Chen, C. Hu, Z. Zhang, and W. Hu, "Automatic treatment planning based on three-dimensional dose distribution predicted from deep learning technique,"*Med Phys*, vol. 46, no. 1, pp. 370-381, 2019.
[26] B. Vanderstraeten, B. Goddeeris, K. Vandecasteele, M. van Eijkeren, C. De Wagter, and Y. Lievens, "Automated Instead of Manual Treatment Planning? A Plan Comparison Based on Dose-Volume Statistics and Clinical Preference,"*Int J Radiat Oncol Biol Phys*, vol. 102, no. 2, pp. 443-450, 2018.
[27] A. T. Y. Chang *et al.*, "Comparison of Planning Quality and Efficiency Between Conventional and Knowledge-based Algorithms in Nasopharyngeal Cancer Patients Using Intensity Modulated Radiation Therapy,"*Int J Radiat Oncol Biol Phys*, vol. 95, no. 3, pp. 981-990, 2016.
[28] J. Krayenbuehl, I. Norton, G. Studer, and M. Guckenberger, "Evaluation of an automated knowledge based treatment planning system for head and neck,"*Radiat Oncol*, vol. 10, p. 226, 2015.
[29] X. Chen, K. Men, Y. Li, J. Yi, and J. Dai, "A feasibility study on an automated method to generate patient-specific dose distributions for radiotherapy using deep learning,"*Med Phys*, vol. 46, no. 1, pp. 56-64, 2019.
[30] A. Akhavanallaf, I. Shiri, H. Arabi, and H. Zaidi, "Whole-body voxel-based internal dosimetry using deep learning,"*Eur J Nucl Med Mol Imaging*, 2021.
[31] X. Mao, J. Pineau, R. Keyes, and S. A. Enger, "RapidBrachyDL: Rapid Radiation Dose Calculations in Brachytherapy Via Deep Learning,"*Int J Radiat Oncol Biol Phys*, vol. 108, no. 3, pp. 802-812, 2020.
[32] L. S. Waters, "MCNPX user's manual,"*Los Alamos National Laboratory*, 2002.
[33] W. Schneider, T. Bortfeld, and W. Schlegel, "Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions,"*Phys Med Biol*, vol. 45, no. 2, pp. 459-78, 2000.
[34] B. Van Veelen, Y. Ma, and L. Beaulieu, "Whitepaper: ACE Advanced Collapsed cone Engine,"*Veenendal, the Netherlands: Elekta Corporation*, 2015.
[35] K. Zourari *et al.*, "Dosimetric accuracy of a deterministic radiation transport based (192)Ir brachytherapy treatment planning system. Part III. Comparison to Monte Carlo simulation in voxelized anatomical computational models,"*Med Phys*, vol. 40, no. 1, p. 011712, 2013.
[36] M. J. Berger, "Energy deposition in water by photons from point isotropic sources,"*J Nucl Med*, pp. Suppl 1:17-25, 1968.
[37] J. Seuntjens, W. Strydom, and K. Shortt, "Dosimetric principles, quantities and units,"*Radiation oncology physics: a handbook for teachers and students. Podgorsak EB, editor. Vienna, Austria: IAEA*, 2005.
[38] Y. Murakami, T. Magome, K. Matsumoto, T. Sato, Y. Yoshioka, and M. Oguchi, "Fully automated dose prediction using generative adversarial networks in prostate cancer patients,"*PLoS One*, vol. 15, no. 5, p. e0232697, 2020.
[39] M. S. Lee, D. Hwang, J. H. Kim, and J. S. Lee, "Deep-dose: a voxel dose estimation method using deep convolutional neural network for personalized internal dosimetry,"*Sci Rep*, vol. 9, no. 1, p. 10308, 2019.
[40] W. Schneider, T. Bortfeld, and W. Schlegel, "Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions,"*Physics in Medicine & Biology*, vol. 45, no. 2, p. 459, 2000.
[41] I. Sechopoulos *et al.*, "RECORDS: improved Reporting of montE CarlO RaDiation transport Studies: Report of the AAPM Research Committee Task Group 268,"*Med Phys*, vol. 45, no. 1, pp. e1-e5, 2018.

Chapter 11

The predictive value of ^{68}Ga-Dotatate PET and baseline biomarkers in ^{177}Lu-DOTATATE tumor dosimetry

Azadeh Akhavanallaf, Avery Peterson, Kellen fitzpatric, Moly Roseland, Kakit Wong, Issam El naqa, Habib Zaidi and Yuni Dewaraja.

Contribution: Study conception and design, methodological development, computer programming, analysis and interpretation of results, manuscript preparation.

Eur J Nucl Med Mol Imaging (2023) submitted

Abstract

Purpose: Metastatic neuroendocrine tumors (NETs) overexpressing type 2 somatostatin receptors are the target for peptide receptor radionuclide therapy (PRRT) through the theragnostic pair of 68Ga/177Lu-DOTATATE. The main purpose of this study was to develop machine learning models to predict therapeutic tumor dose using pre therapy 68Ga-PET and clinicopathological biomarkers.

Methods: We retrospectively analyzed 90 segmented metastatic NETs from 25 patients (M14/F11, age 63.7±9.5, range 38-76) treated by 177Lu-DOTATATE at our institute. Patients underwent both pretherapy 68Ga-DOTATATE PET/CT and four timepoint SPECT/CT at ~4, 24, 96 and 168 hours post-177Lu-DOTATATE infusion. Tumors were segmented by a radiologist on baseline CT or MRI and transferred to co-registered PET/CT and SPECT/CT and normal organs were segmented by deep learning-based method on CT of the PET and SPECT. The SUV metrics and tumor-to-normal tissue SUV ratios (SUV_TNRs) were calculated from 68Ga-PET at the contour-level. Posttherapy dosimetry was performed based on the co-registration of SPECT/CTs to generate time-integrated-activity, followed by an in-house Monte Carlo-based absorbed dose estimation. The correlation between delivered 177Lu Tumor absorbed dose and PET-derived metrics along with baseline clinicopathological biomarkers (such as Creatinine, Chromogranin A and prior therapies) were evaluated. Multiple interpretable machine-learning algorithms were developed to predict tumor dose using these pretherapy information. Model performance on a nested 10-fold cross-validation was evaluated in terms of coefficient of determination (R^2), mean-absolute-error (MAE) and mean-relative-absolute-error (MRAE).

Results: SUV_{mean} showed a significant correlation (q-value <0.05) with absorbed dose (Spearman ρ=0.64), followed by $TLSUV_{mean}$ (SUV_{mean} of total-lesion-burden) and SUV_{peak} (ρ=0.45 and 0.41, respectively). The predictive value of PET-SUV_{mean} in estimation of posttherapy absorbed dose was stronger compared to PET-SUV_{peak}, and SUV_TNRs in terms of univariate analysis (R^2 = 0.28 vs. $R^2 \leq 0.12$). An optimal trivariate random forest model composed of SUV_{mean}, $TLSUV_{mean}$ and total liver SUV_{mean} (normal and tumoral liver) provided the best performance in tumor dose prediction with R^2=0.64, MAE=0.73 Gy/GBq and MRAE=0.2.

Conclusion: Our preliminary results demonstrate the feasibility of using baseline PET images for prediction of absorbed dose prior to 177Lu-PRRT. Machine learning models combining multiple PET-based metrics performed better than using a single SUV value and using other investigated clinicopathological biomarkers. Developing such quantitative models forms the groundwork for the role of 68Ga-PET not only for the implementation of personalized treatment planning but also for patient stratification in the era of precision medicine.

I. Introduction

The theragnostic principle has been summed up as: "*We treat what we see, and We see what we treat*"[1] [1]. This concept of "see and treat" in nuclear medicine therapy has led to the development of theragnostic pairs, consisting of an imaging radiotracer for staging and molecular targeting and its therapeutic counterpart, usually a *beta*- or *alpha*-emitter for tumor ablation. Neuroendocrine tumors (NET) commonly express somatostatin receptors (SSTR), predominantly subtype 2, which is the basis for the use of SSTR PET imaging and peptide receptor radionuclide therapy (PRRT). For the management of NET, the theragnostic pair of 68Ga/177Lu-DOTATATE has been widely used since 2018 when 177Lu-DOTATATE (Lutathera) was approved by the U.S. Food and Drug Administration (FDA) on the basis of NETTER-1 trial results [2, 3].

In current 177Lu-PRRT clinical practice, pretherapy 68Ga-DOTATATE (68Ga-PET) is required for candidate eligibility to confirm sufficient tumor SSTR expression (peformed via qualitative assessment with Krenning score). The approved empiric protocol for 177Lu-PRRT is 4 cycles of 7.4 GBq infusions (~ 2 months intervals). Although 177Lu-PRRT has been showed to improve progression-free survival (65% at 20 months, compared to long-acting octreotide 11%), objective responses are uncommon (20%) and complete responses are rare (1-2%). Therefore, to optimize 177Lu-PRRT outcomes, either patient selection criteria must be improved, or a personalized treatment approach must be developed. Precision nuclear medicine for PRRT has been proposed, with pretreatment 68Ga-PET used for patient selection and additional posttherapy imaging valuable to provide individualized measurements relevant to treatment safety and efficacy [4].

Dosimetry-guided personalized radiopharmaceutical therapy (RPT) involves modulation of the number of treatment cycles or the administered dose per cycle based on posttherapy dosimetry, which has been shown positive impact on treatment response [5-8]. Predictions of therapeutic absorbed doses from 177Lu-PRRT have also been performed using pretherapy 68Ga-PET, which is particularly desirable for planned alterations in the first dose of 177Lu-PRRT, which has the potential to maximize tumor dose while limiting potential toxicity [9, 10]. Two previous studies [9, 11] reported on the ability to predict renal dose using pretherapy imaging, as the kidney toxicity is a limiting factor for 177Lu-labeled RPTs [11]. Knowledge of expected renal dose exposure per cycle is especially important if escalation of administered activity is considered in the first RPT cycle; while not currently performed routinely in clinical practice, prioritizing higher doses early on may be preferable, since there is an observed decrease in absorbed tumor dose per administered activity (Gy/GBq) in subsequent cycles [12].

According to the principles of RPT and cellular irradiation, the likelihood of tumor response is expected to be correlated with the tumor absorbed dose. Various studies have shown dose-response correlations in 177Lu-PRRT [3, 5, 7, 13]. Furthermore, some authors have reported on the correlation of 68Ga-PET uptake with treatment outcome [14, 15]. In this context, tumor absorbed dose estimation prior to the therapy could provide a quantitative metric for response with potential to improve patient-selection criteria. We therefore sought to develop models that predict the mean tumor absorbed doses delivered by 177Lu-DOTATATE using pretherapy 68Ga-DOTATATE PET plus a comprehensive set of clinicopathological biomarkers. The contribution of this work to the field of RPT is threefold: 1) using a previously validated Monte Carlo-based dosimetry workflow with a patient cohort that includes four-posttherapy SPECT/CT scans [16]; 2) including a complete set of clinical biomarkers in addition to 68Ga-PET in the dosimetry prediction models; and 3) implementation of interpretable machine learning algorithms for dose prediction.

II. Materials and methods

Patient population

This study comprised of 25 patients with histologically proven metastatic NETs, progressive on prior therapy, who received at least the first cycle of standard 177Lu-DOTATATE PRRT and underwent four time-point SPECT/CT dosimetry at the University of Michigan Hospital. As part of an ongoing research study approved by the Institutional Review Board, all patients provided written informed consent to participate in the study, which included serial SPECT/CT imaging following standard treatment. Patients' demographic information is presented in supplemental-Table. 1.

[1] Prof. Richard Baum

Tumor and organ delineation

Up to five index lesions larger than 2 mL were manually segmented by a radiologist (MER) on diagnostic-quality baseline CT or MRI and then transferred to the subsequent PET/CT and SPECT/CT scans using co-registration. The spleen was manually segmented by a technologist while kidneys and liver were segmented using a deep learning algorithm on the CT of the PET/CT and SPECT/CT [16]. The normal liver was sampled from uniform uptake regions using three sphere volumes-of-interest (8 cm^3 diameter). The organ segmentations were verified and adjusted by the radiologist as needed.

68Ga PET/CT imaging and PET-derived metrics

Patient preparation required PET scans to be acquired 4 weeks after any long-acting somatostatin analogue treatment. PET/CTs were acquired at ~60 min (range: 54-77 min) post-intravenous injection of ~160 MBq of 68Ga-DOTATATE (range: 144-196 MBq). Data were reconstructed using vendor-specific recommended parameters. Partial volume correction was performed using volume dependent recovery coefficients from a sphere-phantom measurement [2].

Image-derived features, both activity and SUV (standardized uptake value) metrics, were calculated for the transferred contours. Tumor SUV metrics including mean, peak, coefficient of variation (CoV: standard deviation divided by SUV_{mean}), skewness and kurtosis, and mean activity (Bq/mL) corrected to the injection time and normalized by injected activity were extracted. In addition, SUV_{mean} of the spleen, healthy liver, and kidneys along with blood pool (SUV_{mean} in aortic arch) were quantified. The relative tumor uptake was calculated as tumor-to-normal tissue ratios (TNR) using tumor SUV_{mean} relative to the SUV_{mean} of normal spleen (SUV_TNR_{spleen}), normal liver (SUV_TNR_{liver}) and blood pool (SUV_TNR_{blood}). In addition, SUV_{mean} of the total liver volume encompassing both healthy tissue and lesions is quantified as $TotLiverSUV_{mean}$.

To quantify total lesion burden-related metrics, whole-body PET-SUV images were segmented using an empiric SUV threshold (whole-body SUV-cutoff=5, liver SUV-cutoff=10). The generated mask from thresholding was adjusted to add lesions not included in initial segmentation and remove physiological uptake in organs and then verified by the nuclear medicine clinician (KW). Therefore, three independent metrics based on the segmented mask encompassing total 68Ga-DOTATATE-avid lesion volume were defined: Total Lesion Volume (TLV) in mL, average SUV of the Total Lesion Volume ($TLSUV_{mean}$), and Total Lesion Somatostatin Expression (TL-SSE) defined as $TLV \times TLSUV_{mean}$.

Clinicopathological biomarkers

A total of 25 clinical, pathologic, and laboratory variables were included in our study, all of which we believed had theoretical potential to influence patient overall health, tumor behavior, and treatment response. Clinical patient data and lab values were obtained through review of the electronic medical record.

The total variable set, including 16 quantitative and 3 qualitative 68Ga-PET features, 8 treatment history, and 11 blood-test biomarkers, is detailed in Table 1.

177Lu SPECT/CT imaging and dosimetry workflow

Our patient data regarding dosimetry in patients undergoing 177Lu-DOTATATE comes from an ongoing research study that includes serial post-therapy SPECT/CT imaging at ~ 4, 24, 96, and 168 hours after the first cycle. A 25 min single-bed SPECT/CT acquisition is performed on a Siemens Intevo using manufacturer-recommended protocol and reconstructed with Siemens xSPECT Quant using 48 iterations and 1 subset and no post filtering [17].

Table 1. Complete variable set, including PET and clinicopathological features, used in the development of a predictive model for tumor delivered dose from 177Lu-PRRT.

features	Name of feature	Description
Shape	Volume	Volume of index tumors (segmented by radiologist)
PET Uptake/ SUV	SUV_{mean}	Mean SUV value
	SUV_{peak}	Average SUV within a 1 mL sphere centered on the site of highest uptake in a tumor
	SUV_{kurt}	Measure of the shape of the peak of the SUV distribution (kurtosis)
	SUV_{skew}	Measure of the asymmetry of the SUV distribution (skewness)
	$SUV_{BloodPool}$	SUV_{mean} in the aortic arch
	SUV_{Spleen}	SUV_{mean} of the Spleen contour
	SUV_{Liver}	Average SUV_{mean} of three spheres (8 mL) sampled from the normal liver tissue
	$SUV_{Kidneys}$	Average of SUV_{mean} from right and left kidney contours
	SUV_TNR_{blood}	Ratio of Tumor SUV_{mean} to blood pool SUV_{mean}
	SUV_TNR_{spleen}	Ratio of Tumor SUV_{mean} to SUV_{mean} of the spleen
	SUV_TNR_{liver}	Ratio of Tumor SUV_{mean} to SUV_{mean} of the liver
	$TotLiverSUV_{mean}$	SUV_{mean} of the whole liver including both normal and tumoral tissues
	TLV	Total lesion volume
	$TL\text{-}SUV_{mean}$	Average SUV of the entire total lesion volume
	TL-SSE	Total lesion somatostatin expression (TLV× $TL\text{-}SUV_{mean}$)
Diagnostic	Liver metastasis	Disease present in liver (Based on Dotatate PET)
	Bone metastasis	Disease present in bone (Based on Dotatate PET)
	Node metastasis	Disease present in lymph nodes (Based on Dotatate PET)
	Tumor location	Anatomical location of the index tumor
Histological	Grade	Histologic grade (using Ki67 index) of primary tumor from biopsy/surgery
	Primary tumor site	Primary Tumor Site
Treatments	#Systemic therapy	Number of prior systemic treatments (Chemotherapy or other)
	#Directed therapy	Number of prior liver directed treatments (TACE, Y90, cryotherapy)
	Y90-SIRT	Prior treatment liver with Y90-SIRT
	Everolimus	Prior treatment with everolimus (Systemic MTOR inhibitor)
	Capecitabine/ temozolomide	Prior treatment with capecitabine and temozolomide (Chemo, systemic)
	Sunitinib	Prior treatment with Sunitinib (Multi-Kinase Inhibitor, systemic)
Blood tests	White Blood Cells	White Blood Cells (K/cmm)
	Lymphocytes	Lymphocytes
	Absolute Neutrophil	Absolute Neutrophil Counts (K/cmm)
	Hemoglobin	Hemoglobin (g/dL)
	Platelet	Platelet Count (K/cmm)
	eGFR	Estimated Glomerular Filtration Rate (Calculated)
	Creatinine	Creatinine (mg/dL)
	Bilirubin	Bilirubin (mg/dL)
	Albumin	Albumin (mg/dL)
	Alkaline Phosphatase	Alkaline Phosphatase (ALK, ALP, ALKP, or ALK PHOS) (IU/L)
	CgA	Chromogranin A (Tumor Marker) (ng/mL)

For dosimetry, we employed an integrated workflow implemented within MIM software that has been elaborated in a recent article by Dewaraja et al [16]. The workflow is composed of the following steps:

1. A contour-guided intensity-based registration was used to align four posttherapy SPECT images.
2. Time Integrated Activity (TIA) was calculated by integration of the time-activity curve, a mono/ bi-exponential function ($TIA = \int_{t_0}^{\infty} \mathrm{C}\,(e^{-\lambda_1 t} - e^{-\lambda_2 t})$). Here, C scales the curve up or down, $\lambda_1$is the clearance/elimination rate, and λ_2 is the uptake/absorption rate. The term effective half-life (T_{eff}) refers to the slower exponential component ($i.e. T_{\mathrm{eff}} = \frac{\ln(2)}{\lambda_1} \mid \lambda_1 \ll \lambda_2$).
3. TIA along with the corresponding density map (obtained from CT) were coupled with a fast Monte Carlo (MC) simulator, developed at the University of Michigan [18], to generate the voxel-level absorbed dose map.

Statistical analysis and predictive modeling

For the statistical analysis, the Spearman rank correlation between predictive features and tumor dose were analyzed, followed by Benjamini and Hochberg p-value correction, where q-value <0.05 considered significant.

To predict tumor absorbed dose using PET-derived features and biomarkers, a cross-combination of different regression models were analyzed. We compared linear and supervised random forest regression algorithms through univariate, bivariate, and multivariate analysis implemented in MATLAB 2022 (MathWorks Inc., Natick, MA, USA). We adopted nested cross-validation (CV), whereby the outer-loop CV was repeated 10 times to consolidate the results of 10-fold inner-loop CV [19]. During inner-loop CVs, 10% (8/80) of the whole dataset was considered as unseen validation-set and 90% used as training-set. Due to the intrinsic heterogeneity and limited size of our data, bootstrap aggregation strategy (500 bootstrap samples with replacement) was implemented to improve model stability and avoid overfitting (algorithm flowchart in supplemental-Figure 1).

We designed a hierarchical interpretable feature selection strategy using main-effect analysis to select the most important predictors. First, using a univariate linear regression model, the best variable with the highest coefficient of determination (R2) was determined. In the second step, a set of bivariate regression models was generated using two independent variables, i.e. the selected variable from the univariate analysis followed by a second variable from the predictor-set. In the third step, the five best bivariate models that most increased predictive likelihood (with the highest R2) were selected, forming the basis for a set of trivariate models. We extended the process up to four-variable models, but because we saw no further significant improvement in predictive likelihood, this process was stopped. In addition, we employed ElasticNet and Permutation-based Random Forest variable-Importance (PRFvI) feature selection algorithms. The feature-selection algorithms were implemented in a bootstrap ensemble framework as elaborated in supplemental-Figure 2. A maximum of 8 features were selected based on the recommended number of at least ten observations per predictor [20].

We employed the proposed hierarchical feature selection algorithm in both linear and random forest algorithms. In the linear model, we used a generalized linear regression model based on the least square loss function. In the case of random forest algorithms, we used a bootstrap aggregation between two models including random forest (ensemble tree) [21] and generalized additive model [22] (supplemental-Figure 2). To reduce overfitting and improve generalizability, we grew a shallow tree by forcing the number of observations per leaf to be at least 10 or the number of splits per predictor to be at most 5. The number of ensembled trees (=200) was obtained from hyperparameter optimization. We implemented the proposed hierarchical feature selection algorithms on both linear and decision tree regression models. Additionally, the selected features from ElasticNet were fed to a multivariate generalized linear model and those selected based on PRFvI algorithm were tested in the decision-tree model. The model performance was evaluated based on nested CV 10-fold R^2, mean-absolute-error (MAE) and mean-relative-absolute-error (MRAE) compared to ground truth.

We further tested sensitivity and specificity of the best model for predicting tumor absorbed dose >25 Gy/cycle for response. This threshold dose was chosen as it is a previously reported potential cutoff for tumor response following 177Lu-PRRT [5, 6].

III. Results

A total of 25 patients (M14: F11, age 63.7±9.5, range 38-76) with 90 neuroendocrine tumors larger than 2 mL (mean= 65.6±139.9 mL, range: 2.1-1039 mL) met the study criteria. An example of corresponding 68Ga-PET, post-treatment 177Lu SPECT/CT, and resulting time-activity curves of target lesions are given in Figure 1. PET-SUV_{mean} and SUV_{peak} measured from the 90 studied tumors were 16±6.4 (5.6-34.2) and 26.4±15.5 (6.1-104), respectively, while SUV_{mean} for normal liver, spleen and kidneys were 6.9±2.4 (2.2-11.4), 13.1±3.5 (7-19.2) and 5.4±2.7 (5.4-19.2), respectively. The mean tumor absorbed dose averaged 2.68±1.89 Gy/GBq (0.23-10.26 Gy/GBq), while the average value of T_{eff} was 91.6±26.6 h (27.9-159.5 h).

The statistical variability of the investigated predictors dichotomized based on ANOVA-test of the absorbed dose are illustrated in Table 2. The dichotomization cut-offs of the continuous predictors were calculated from an iterative process (1000 iterations), in which a random number within the range of predictor's quantiles (0.05-0.95) were generated to binarize the predictor values. Then, one-way ANOVA test was applied on dose vector according to the binarized predictor; thus, the cut-off was selected based on the minimum p-value obtained from ANOVA-test.

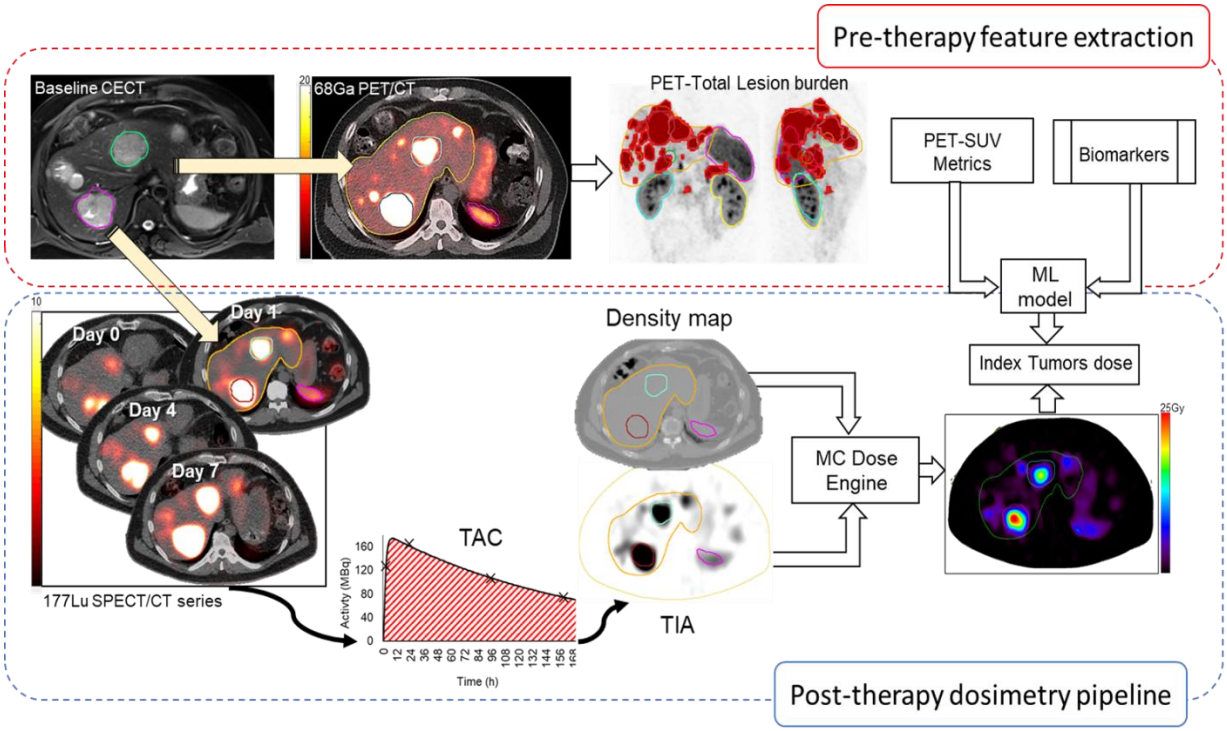


Figure 1. (Top panel) baseline diagnostic images (contrast-enhanced CT/MRI) were used to define target lesions, which were then co-registered to pre-therapy 68Ga-DOTATATE PET/CT and post-therapy 177Lu-DOTATATE SPECT/CT images. (bottom panel) dosimetry pipeline included four timepoints registration of SPECT images to generate TIA that is fed into MC-based dose engine.

Self and cross-correlation of all baseline features compared to tumor absorbed dose and 4 other dose-related parameters (parameters that directly contribute to absorbed dose) is presented in Figure 2 (Spearman-correlation (ρ), q-value<0.05). The dose-related parameters are the scale factor C of the time-activity curve normalized by tumor volume (C_{vol}), TIA normalized by tumor volume (TIA_{vol}) and T_{eff}. We expect a physics-informed correlation between dose and TIA_{vol}, according to the assumption of local-energy-deposition for 177Lu-labeled agents [23], and hence a correlation with TIA_{vol} components (C_{vol} and T_{eff}). SUV_{mean} shows a strong correlation with dose parameters (dose: ρ=0.64, TIA: ρ =0.39, C_{vol}: ρ =0.56), followed by $TLSUV_{mean}$ (dose: ρ =0.45, TIA_{vol}: ρ =0.51, C_{vol}: ρ =0.63) and SUV_{peak} (dose: ρ =0.41, TIA_{vol}: ρ =0.54, C_{vol}: ρ =0.56). $TotLiverSUV_{mean}$ shows a correlation only with C_{vol} (ρ =0.44). A significant but moderate correlation between T_{eff} and the pre-PRRT number of systemic treatments (#Systemic therapy) (ρ =-0.31), Capecitabine/ temozolomide (ρ =-0.35) and bilirubin (ρ =0.33) is observed.

Figure 3 illustrates the intra-patient variability of the index tumor absorbed doses among the study population. The intra-patient tumor dose variability in terms of coefficient of variation (CoV) was within the range of 0.04-0.78 (Median=0.38); this is comparable to the variation within the whole tumor-set, which had a CoV of 0.69.

For our predictive models, we excluded lesions smaller than 4 mL in order to reduce the dose calculation uncertainties owing to mis-registrations and partial volume effects (>4mL, N=80). Also, two lesions with highest doses (P_22, P_25 in Figure 3) were considered outliers and excluded from model building because of their exceptionally high uptake in 177Lu-SPECT, despite 68Ga-PET uptake in a similar range compared to the other analyzed lesions (both in the same patient and other patients).

The association of dose and different SUV parameters were evaluated using univariate analysis (linear least-square regression). SUV_{mean} (coefficient-of-determination: R^2=0.28), compared to SUV_{peak} (R^2=0.07), and SUV_TNRs ($R^2 \leq 0.12$) showed a better performance in prediction of therapeutic dose (Figure 4). The majority of studied tumors were found in the liver (75/90), while 11 lesions were lymph node metastases. Three primary pancreas tumors and one chest tumor were also included. No significant differences of absorbed doses or SUV-parameters were observed based on tumor volume or localization.

Table 2. Patient population clinicopathological biomarkers. The variability of the tumor dose and SUVmean with respect to the dichotomized predictors is illustrated. The forest-plot represents the range of dose values in the selected predictor's group while dots represent mean dose values. P-value was obtained from ANOVA test.

Parameters	N (%)	Mean ± std	SUVmean	Dose (Gy/GBq)	p-val
Volume					
≤52 mL	70	15.07 ± 10.6	20.74 ± 8.6		0.13
>52 mL	30	177.07 ± 217.5	17.93 ± 6.7		
SUV_CoV					
≤ 45 %	68	0.30 ± 0.1	19.51 ± 7.9		0.02
> 45 %	32	0.55 ± 0.1	20.70 ± 8.7		
TLV (mL)					
≤ 150	10	60.87 ± 42.6	23.30 ± 8.6		<0.001
> 150	90	624.30 ± 664.4	19.52 ± 8.0		
TL-SSE (SUV.mL**)**					
≤ 2864	30	1683 ± 763	16.30 ± 7.5		0.001
> 2864	70	12005 ± 9895	21.44 ± 7.9		
White blood cells					
≤ 5.4	35	3.57 ± 0.8	21.26 ± 10.1		0.08
> 5.4	65	8.16 ± 2.1	19.16 ± 6.8		
Lymphocytes					
≤ 1.6	64	0.91 ± 0.4	19.73 ± 8.5		0.04
> 1.6	36	1.95 ± 0.3	20.19 ± 7.5		
Absolute neutrophil					
≤ 2	18	1.62 ± 0.2	21.91 ± 9.8		0.03
> 2	82	4.54 ± 1.5	19.47 ± 7.7		
Hemoglobin					
≤ 12	20	11.24 ± 0.9	18.71 ± 10.1		0.01
> 12	80	13.70 ± 1.0	20.20 ± 7.6		
Platelet					
≤ 190	29	167.81 ± 19.3	19.40 ± 9.4		0.008
> 190	71	255.90 ± 59.4	20.10 ± 7.6		
eGFR					
≤ 52	12	40.26 ± 10.6	16.47 ± 7.3		0.04
> 52	88	82.46 ± 14.0	20.39 ± 8.2		
Creatinine					
≤ 1.2	81	0.89 ± 0.2	20.67 ± 8.1		0.03
> 1.2	19	1.39 ± 0.3	16.56 ± 7.5		
Bilirubin					
≤ 0.9	81	0.49 ± 0.2	20.18 ± 8.0		0.08
> 0.9	19	1.13 ± 0.1	18.67 ± 8.9		
Albumin					
≤ 4.5	68	4.04 ± 0.3	20.86 ± 7.8		<0.001
> 4.5	32	4.88 ± 0.3	17.91 ± 8.5		
Alkaline phosphatase					
≤ 88	20	74.57 ± 13.7	23.21 ± 8.2		
> 88	80	149.91 ± 40.3	19.07 ± 7.9		<0.001
CgA					
≤ 500	28	146.60 ± 140.6	25.91 ± 8.1		0.001
> 500	72	3434.12 ± 7070.4	17.62 ± 6.9		
Grade					
G=1	33		22.16 ± 9.8		0.3
G ≥ 2	67		18.69 ± 7.0		
Bone Met					
No	51		20.85 ± 7.5		0.12
Yes	49		18.90 ± 8.7		
Nodal Met					
No	29		23.22 ± 8.5		0.14
Yes	71		18.56 ± 7.6		
# Systemic therapy					
0	52		17.90 ± 5.5		0.54
1 or 2	42		20.99 ± 9.2		
≥ 2	5		31.56 ± 11.3		
Primary tumor site					
Midgut	61		19.40 ± 7.8		0.79
Pancreas	28		22.84 ± 9.2		
Other	11		15.43 ± 4.0		
				0 2 4 6	

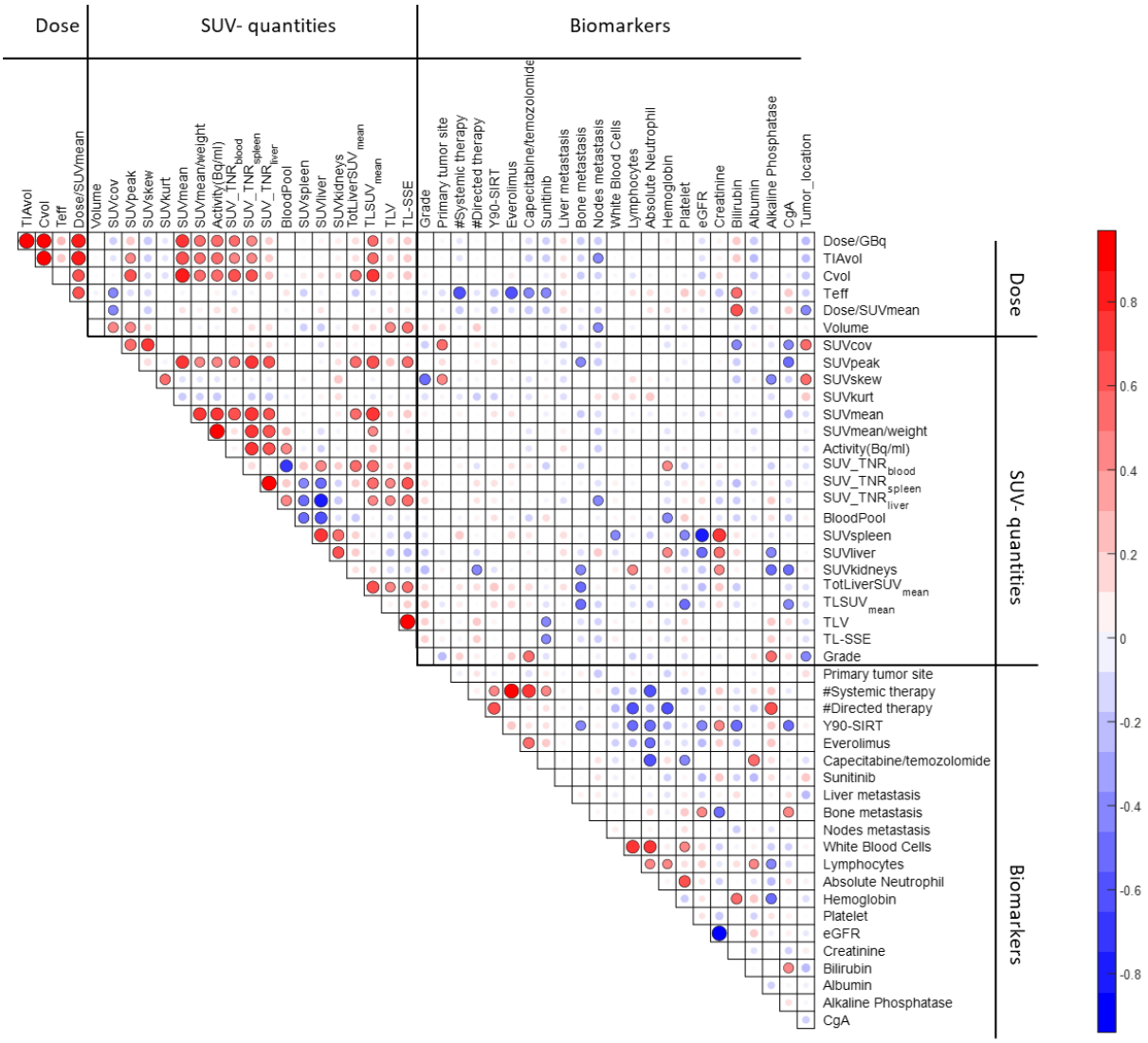


Figure 2. Spearman rank self and cross correlation between absorbed dose-related parameters (Dose, TIA_{vol}, C_{vol} and T_{eff}) and PET-SUV parameters along with biomarkers. The color-code and size of spheres show the correlation magnitude. The insignificant correlations (q-value>0.05) are plotted as faded spheres.

According to the proposed hierarchical feature-selection strategies (supplemental-figure 5), linear univariate regression model picked SUV_{mean} as the anchor predictor with R^2=0.28 and MAE=1.08 Gy/GBq. We compared a cross-combination of all features with SUV_{mean} to evaluate the second and third important features in dose prediction (supplemental-Figure 5 and 6). In bivariate model, $TotLiverSUV_{mean}$ and $TLSUV_{mean}$ were the most effective predictors in terms of R^2 and MAE (R^2=0.61 and 0.48, MAE=0.82 and 0.88 Gy/GBq, respectively) from Ensembled Tree (Ens-Tree) models. The best prediction performance was achieved from a trivariate Ens-Tree algorithm consisting of SUV_{mean}, $TotLiverSUV_{mean}$ and $TLSUV_{mean}$ with R2=0.64 and MAE=0.73 Gy/GBq (Table 3). The predicted dose compared with the measured absorbed dose from different algorithms is illustrated in Figure 5.

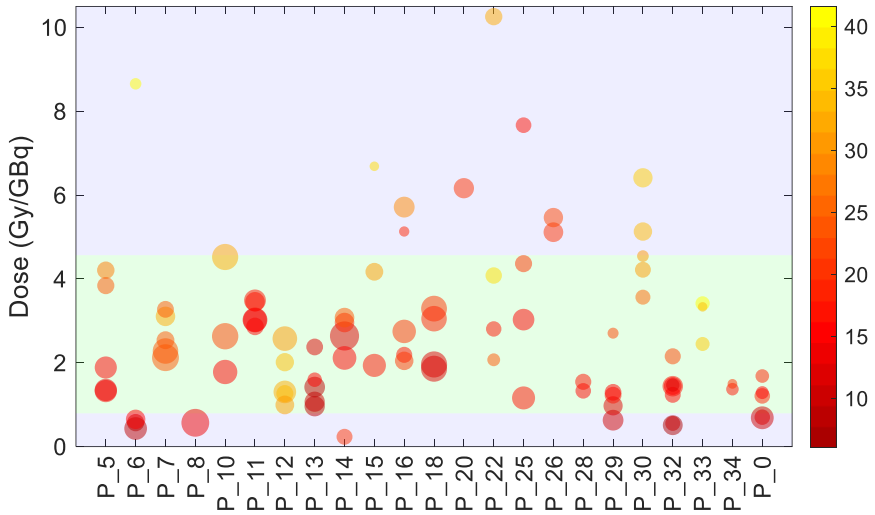


Figure 3. Intra-patient variability of tumor doses for all patients. The sphere color indicates SUV_{mean} and background color shows the margins of standard deviation of tumor dose values. The size of spheres depicts the volume of tumors in logarithmic form (4-1039 mL).

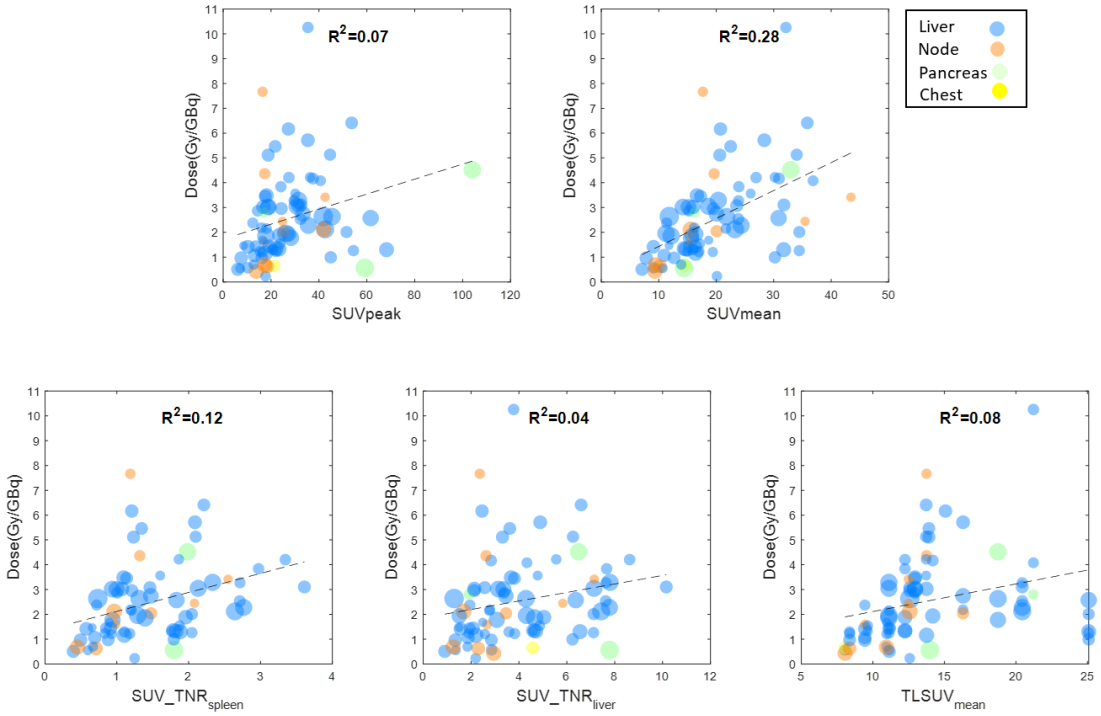


Figure 4. Tumor absorbed dose plotted vs. tumor PET-SUV quantities, where the color shows the tumor location. The size of spheres depicts the volume of tumors in logarithmic form (4-1039 mL).

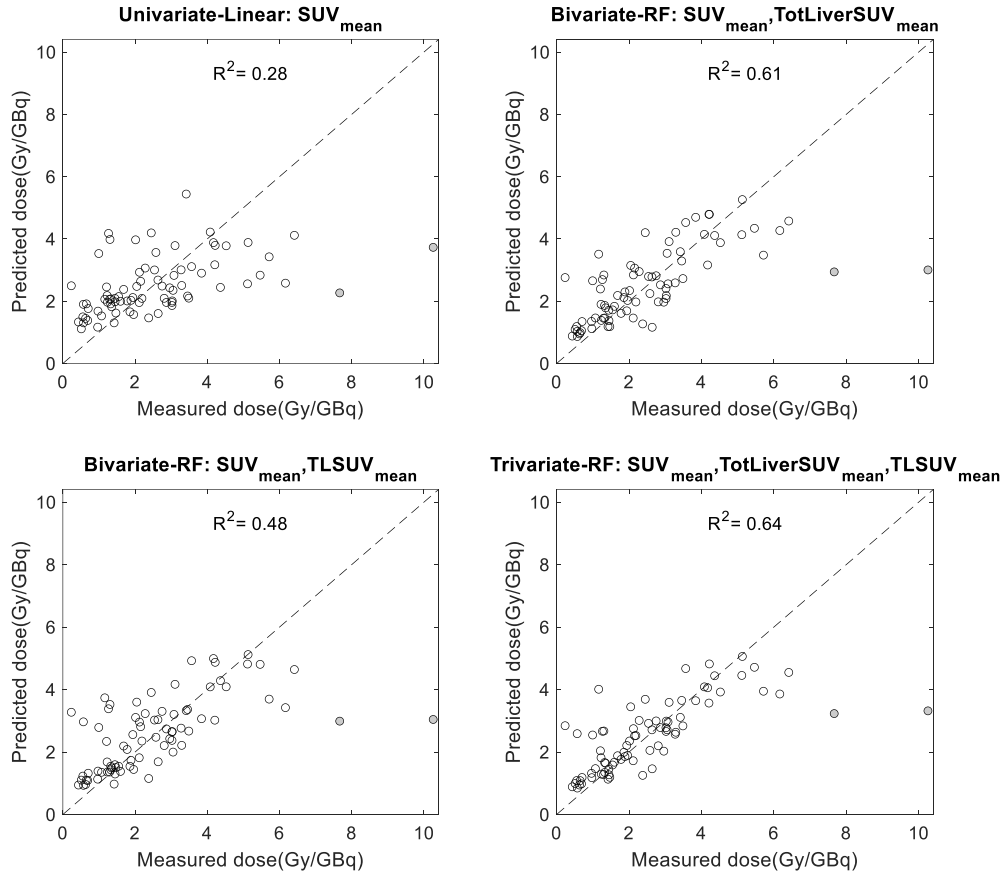


Figure 5. Pre-therapy predicted dose using univariate linear model and random forest (RF) bi/tri-variate models of Table 2 vs. the delivered dose measured from Lu-177 SPECT/CT (the filled gray dots represent the 2 outliers).

Table 3. Model performance of the selected prediction algorithms using 68Ga-PET SUV metrics. The quantitative metrics are reported as mean (95% CI) calculated from nested CV. The MAE quantile range is reported based on the averaging over 10-outerloop CV point prediction.

Model		Features	R2* 10-fold	Median MRAE	MAE (Gy/GBq)	MAE Quantile (0.05-0.95)
Univariate	Linear					
		SUV_{mean}	0.28(0.00)	0.38(0.00)	1.08(0.00)	0.14-2.8
Bivariate	Tree_Ens					
		SUV_{mean}, $TotLiverSUV_{mean}$	0.61(0.01)	0.26(0.01)	0.82(0.01)	0.10-2.29
		SUV_{mean},$TLSUV_{mean}$	0.48(0.03)	0.26(0.01)	0.88(0.02)	0.05-2.66
Trivariate	Tree_Ens					
		SUV_{mean}, $TotLiverSUV_{mean}$, $TLSUV_{mean}$	0.64(0.02)	0.20(0.01)	0.73(0.02)	0.02-2.46

* Two outliers are excluded from R2

The sensitivity and specificity of the best-performing model (trivariate Ens-Tree from Table 3 and Figure 5), using a threshold of 25 Gy/cycle for response, was calculated as 0.76 and 0.94, respectively (Figure 6). Again, this threshold-level was chosen to mirror previously reported dose-cut-offs for response following 177Lu-PRRT [5, 6].

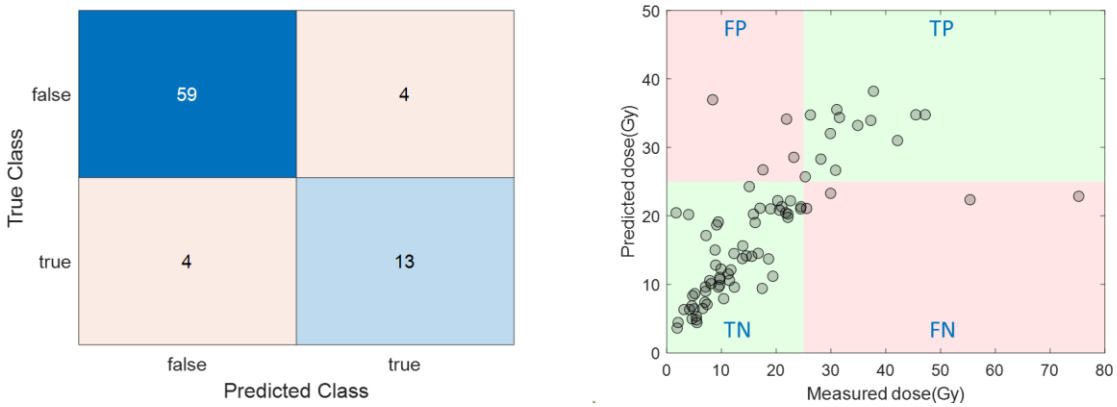


Figure 6. Considering the threshold dose for responders of 25 Gy/cycle, confusion matrix of predicted dose from trivariate Ens-Tree model compared to the measured absorbed dose (left). Sensitivity ($\frac{TP}{TP+FN}$) and specificity ($\frac{TN}{FP+TN}$) visualization of the predcition model (right). TP: True-Positive, FP: False-Positive, TN: True-Negative, FN: False-Negative.

IV. Discussion

Accurate and early prediction of therapeutic absorbed dose in NETs is important information that can be used to guide appropriate patient selection and treatment alterations for PRRT, potentially helping to distinguish between patients likely to undergo effective versus futile treatments. To date, 68Ga-PET derived quantitative metrics have appeared promising as a measure of SSTR2 density in neuroendocrine tumors [24]; however, studies assessing correlation between SUV features and absorbed dose/treatment outcomes remain scarce, and further investigation is necessary to establish conclusive relationships.

Prediction of tumor and organ-absorbed doses may help optimize treatment efficacy prior to therapy by enabling an individualized treatment plan, administering variable doses of PRRT that maximize tumor irradiation while minimizing organ exposure. According to a recent study indicating the decline of 177Lu-DOTATATE tumor uptake over therapy cycles, an individualized dose escalation strategy may be more effective in the first cycle [12]. A clinical trial reported on personalized 177Lu-DOTATATE PRRT guided by the prediction of renal toxicity based on eGFR and patient-surface-area prior to the therapy [5]. In our ongoing research with standard dose 177Lu-DOTATATE, we have already developed a predictive model for kidney absorbed dose based on pretherapy PET-SUV metrics and biomarkers (i.e. eGFR) estimating posttherapy renal dose within 18% accuracy [11]. In the current study, we further evaluated the predictive power of 68Ga-PET SUV metrics with readily available baseline biomarkers to develop machine learning models for tumor absorbed dose prediction.

The relationship between baseline PET-derived features and delivered absorbed dose is not straightforward. First, there are notable differences in the pharmacokinetics and biodistribution of 68Ga/177Lu-DOTATATE theragnostic pairs [2], influenced by variable masses and chemical structures of administered radiopharmaceuticals, patient behavior [4], radioactive metabolites [25], medication effects, etc. Second, the static 68Ga-PET acquisition (~60 min post-injection) potentially only depicts the SSTR2 density distribution, while the absorbed dose quantity is related to dynamic physiologic circulation and accumulation of the radiopharmaceutical. In the other word, dose quantity is proportional to the multiplication of C_{vol} (scale factor of the time-activity curve normalized by tumor volume) and T_{eff} (retention half-life).

In this context, we observed a significant correlation of PET-SUV metrics with C_{vol} (Figure 2-3, SUV_{mean}: ρ=0.63), while no correlation with T_{eff} (Supplemental-Figure 3-4). Therefore, it can be concluded that the observed correlation between PET-SUV parameters and the tumor absorbed dose quantity (SUV_{mean}: ρ=0.62) stems from

the correlation between 68Ga-tumor-uptake and 177Lu-tumor-uptake. There is a body of literature that indicated a significant correlation between 68Ga-SUV and 177Lu-induced tumor dose [9, 26-28]. Ezziddin et al. reported a strong correlation between 68Ga-DOTATOC SUV-metrics with 177Lu-Octreotate absorbed dose (SUVmean: ρ=0.72; SUVmax: ρ=0.71) [26]. Hänscheid et al. showed that PET-based SUVmax significantly correlates (ρ=0.76) with the maximum dose delivered to tumor in meningioma patients [29]. However, one group, Singh et al, found no significant correlation between SUVs and the tumor dose from 177Lu-DOTATATE therapy in metastatic-NETs [24].

In previous studies, tumor-to-normal organ rations (SUV_TNRs) were suggested as potential factors that might reduce the inter-patient and inter-acquisition variability associated with tumor SUV by using physiological uptake in normal organs as an individualized reference [30-32]. We compared the correlation of tumor SUV, SUV_TNRs and activity concentration with respect to absorbed dose, but SUV_{mean} outperformed other metrics in terms of strength of the correlation (Figure 4). We have previously noted discordance using TNR between the 68Ga PET and the 177-Lu-PRRT dosimetry SPECT/CT, with significantly higher SUV TNR on 177Lu SPECT compared with 68Ga PET [2]. This phenomenon may be related to temporal differences in DOTATATE uptake and internalization in tumor as compared to normal organs, further accentuated by differences in image timing (60 minute PET vs. >4 hr SPECT/CT) [2].

We evaluated the correlation of inter-patient PET-derived total lesion burden metrics, including total lesion volume (TLV), average SUV of the total lesion volume ($TLSUV_{mean}$), and total lesion somatostatin expression (TL-SSE= TLV× $TL\text{-}SUV_{mean}$), all compared to the index tumor absorbed dose (Figure 2). $TLSUV_{mean}$ showed a strong correlation with dose components (C_{vol}: ρ>0.63), while no significant correlation was observed regarding TLV and TL-SSE. This association is reasonable from a physiologic standpoint, given that greater overall PET tracer avidity may correlate to increased PRRT binding and dose deposition by a similar theragnostic pair. Accordingly, a recent paper notably found correlation of $TLSUV_{mean}$ with survival in NET patients treated by 177Lu-DOTATATE, implicitly showing correlation of $TLSUV_{mean}$ with tumor absorbed dose and accordingly therapy-response [33]. Furthermore, we found a strong correlation between SUV_{mean} of the total liver volume ($TotLiverSUV_{mean}$) with dose components (C_{vol}: ρ=0.45). We used $TotLiverSUV_{mean}$ as a surrogate for extent of hepatic metastatic disease involvement:

$$TotLiverSUV_{mean} = \frac{(Tumor\_liver\_SUV_{mean} \times Tumor\_volume) + (healthy\_liver\_SUV_{mean} \times healthy\_liver\_volume)}{(healthy + tumor)\ liver\ volume}$$

By expanding a univariate analysis showing the predictive value of SUV_{mean}, we built bi/tri-variate models to enhance prediction accuracy. The best model performance achieved by a trivariate model composed of only PET-SUV metrics: SUV_{mean}, $TotLiverSUV_{mean}$ and $TLSUV_{mean}$. All three metrics showed strong correlation with radiopharmaceutical-uptake-related dose component (C_{vol}), illustrated in supplemental-Figure 7. A bivariate model only using SUV_{mean} and $TotLiverSUV_{mean}$ likewise showed a good predictive performance (R^2=0.61, MAE=0.82 Gy/GBq). These results illustrate that the extent of liver tumor involvement, via $TotLiverSUV_{mean,}$ is predictive of dose. The main advantage of using this variable is that it is readily calculated from PET images without any complicated computation: it is merely the SUV_{mean} of entire liver segmented volume, which can be simply performed through machine learning models from CT images.

In addition to Ens_Tree models, we evaluated bi- and tri-variate linear models, where SUV_{mean} combined with bilirubin and albumin improved the prediction performance (R^2=0.47, MAE=0.87 Gy/GBq). Bilirubin and prior systemic treatment showed significant correlations with T_{eff} (ρ=-0.33 and ρ=0.3, respectively) that can indirectly imply the impact of retention half-life on absorbed dose values. These findings may suggest that prior treatments or underlying hepatic dysfunction may alter tumor behavior and potentially the degree of PRRT tumor uptake and metabolism. Our linear model, built upon the features selected by ElasticNet (7 variables), also showed some improvement compared to trivariate models (R2=0.57, MAE=0.8 Gy/GBq), but due to a higher number of model-variables, it is prone to spurious correlations in a small-size dataset. The features selected by PRFvI algorithm align with those from the hierarchical algorithm; however, compared to trivariate decision tree, the model performance did not show any improvement (supplemental-Figure 8).

Tumor absorbed dose in PRRT is likely influenced by multiple biological factors, both individual patient characteristics and specific tumor features (i.e. proliferation rate, heterogeneity, intrinsic radio-sensitivity). The intra-patient tumor dose variability of our dataset is comparable with inter-patient variability of the whole set (0.38 vs. 0.69), therefore, we treated each individual tumor independently, while the biomarkers and some PET

features, such as TL-SSE and $TLSUV_{mean}$ were calculated in the patient-level, feeding inter-patient information to our models.

The primary limitations of our study are its small sample size and lack of independent multi-center validation and back-testing of the models, relying instead on nested cross validation. Although we followed the recommended rules for generalizability and interpretability of the models [20], further investigation is warranted. An inherent limitation of tumor dosimetry studies relates to the uncertainties associated with quantitative imaging (i.e. scatter/attenuation correction, segmentation, and partial volume correction) and multi-timepoint serial imaging to determine kinetics (i.e. time-series registration) [16]. In addition, simplification in post-therapy imaging such as using SPECT-planar hybrid imaging or reduced time-points; or approximation in particle transport algorithms can introduce extra uncertainties into dosimetry process [34]. To the best of our knowledge, this is the first study of predictive dosimetry using complete four-timepoint posttherapy 3D SPECT/CT imaging, radiologist-defined lesion contours, and a validated Monte Carlo-based dosimetry workflow that reduces some of these uncertainties in the measured absorbed dose and hence help to build a more precise model. As post-PRRT imaging is increasingly used as part of routine clinical protocols at some centers, we expect more data to be available in the future to independently validate and improve the proposed model.

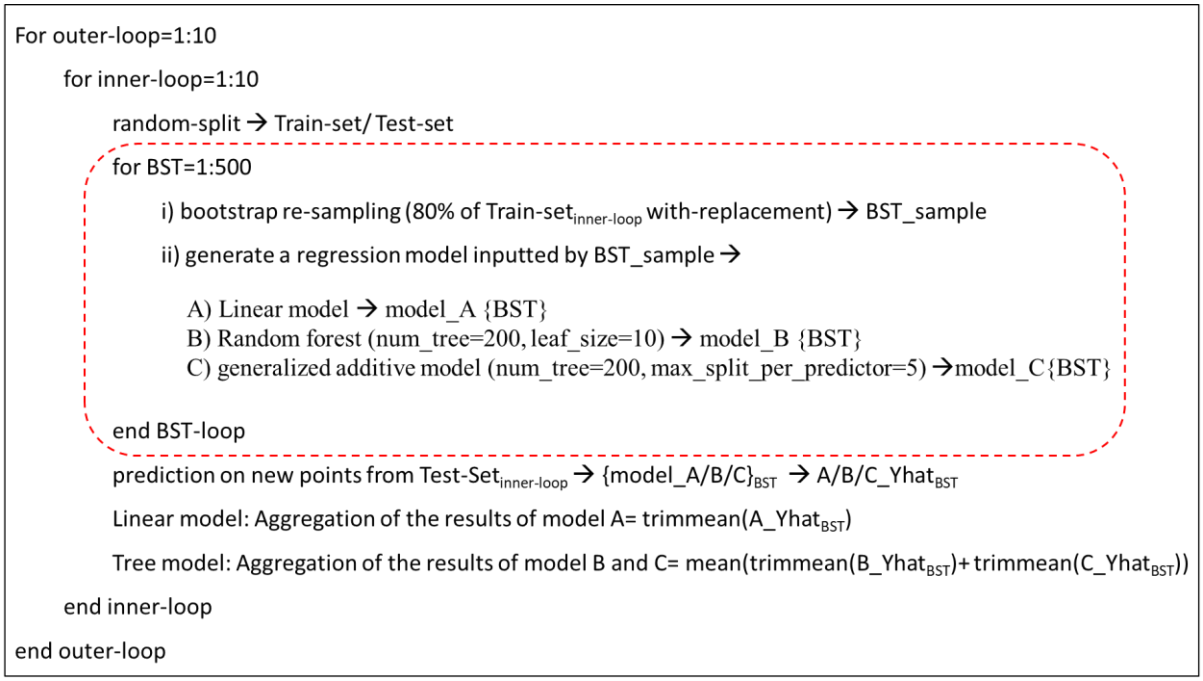
V. Conclusion

To explore dose-response relationships in PRRT, we investigated the predictive value of using 68Ga-PET-based SUV metrics along with biomarkers to estimate the tumor absorbed dose with 177Lu-DOTATATE therapy. We showed that tumor SUV_{mean}, $TotLiverSUV_{mean}$ and average SUV of the total lesion volume ($TLSUV_{mean}$) are capable of predicting the 177Lu-PRRT delivered tumor absorbed dose with an accuracy of MAE=0.71 Gy/GBq (R^2=0.64) in nested cross validation. We hope to further test the proposed models on multi-center data, to eventually provide a validated decision-support tool for clinicians to improve patient-selection and thus optimize treatment outcomes. Developing such precise quantitative metrics establishes a greater role of 68Ga-PET for patient stratification, as well as prognostication and assessment of the therapeutic response modeling.

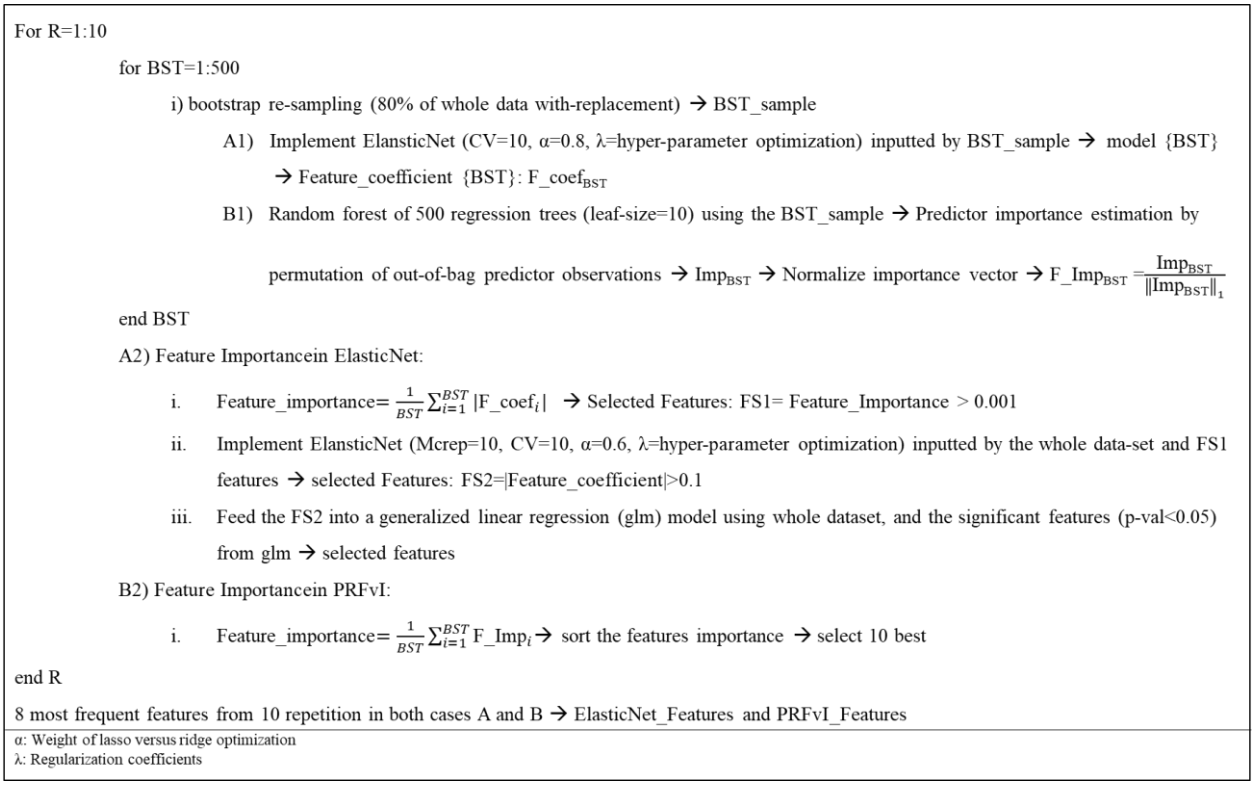
Acknowledgement

This work was supported by grants R01CA240706 and P30CA046592 from the National Cancer Institute; and the Euratom research and training programme 2019-2020 Sinfonia project under grant agreement No 945196. Open access funding provided by University of Geneva.

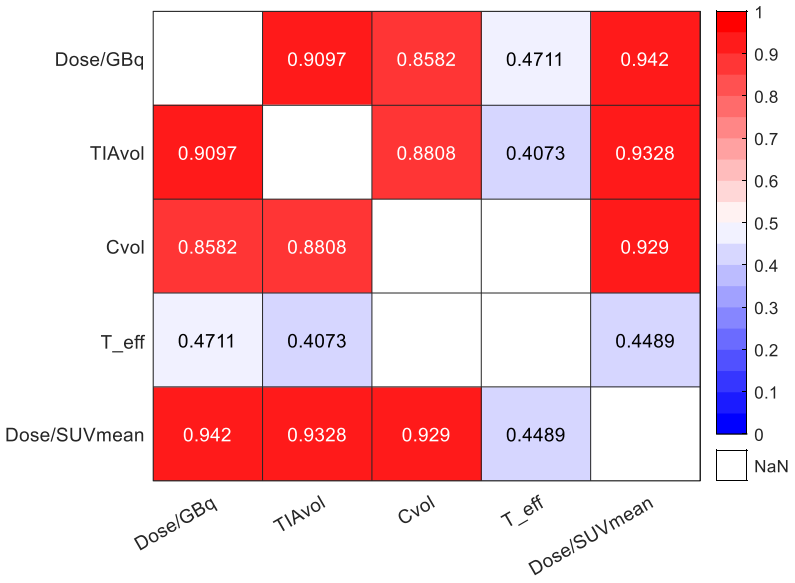
Supplemental material



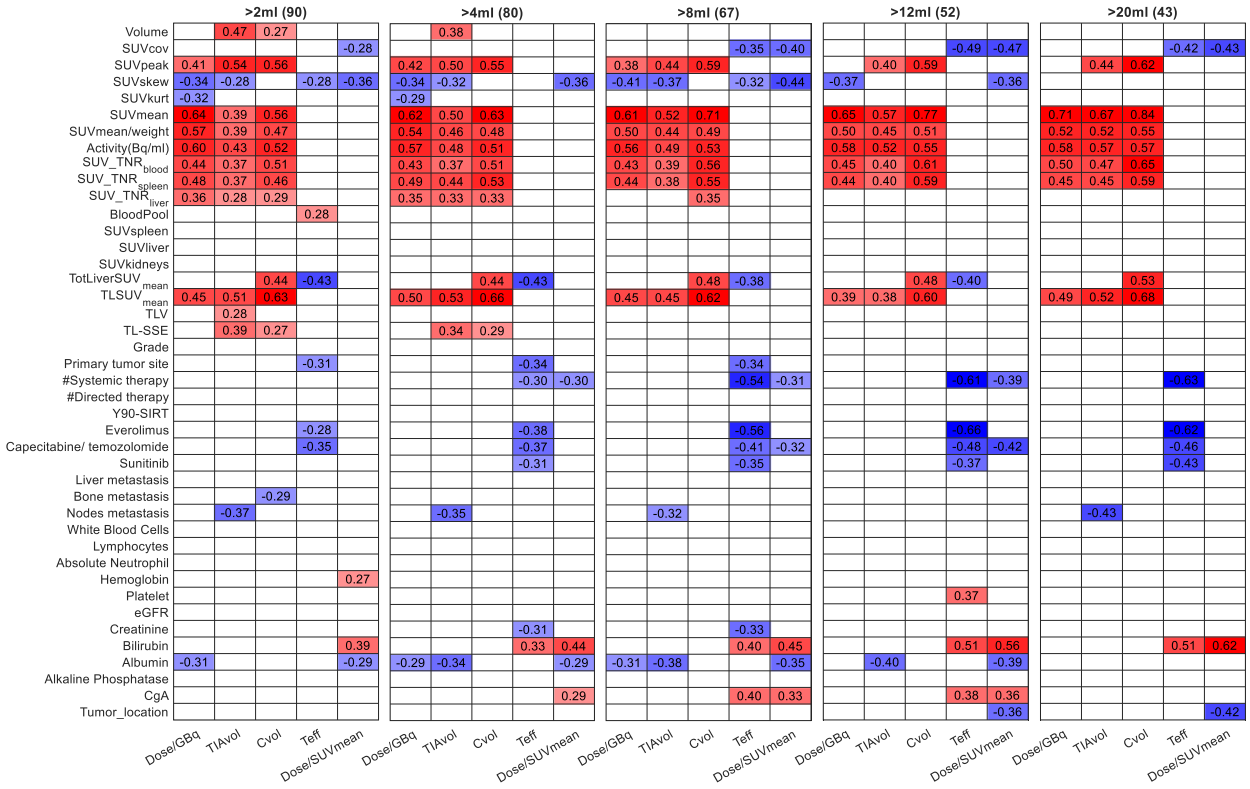
Supplemental-Figure 1. Nested CV in validation process for evaluating both linear and decision tree models in bootstrap aggregation strategy.



Supplemental-Figure 2. Flowchart representing feature selection using two strategies: 1. ElasticNet and Permutation-based Random Forest variable-Importance (PRFvI).



Supplemental-Figure 3. Self-correlation between absorbed dose related parameters (Spearman, q-value <0.05).

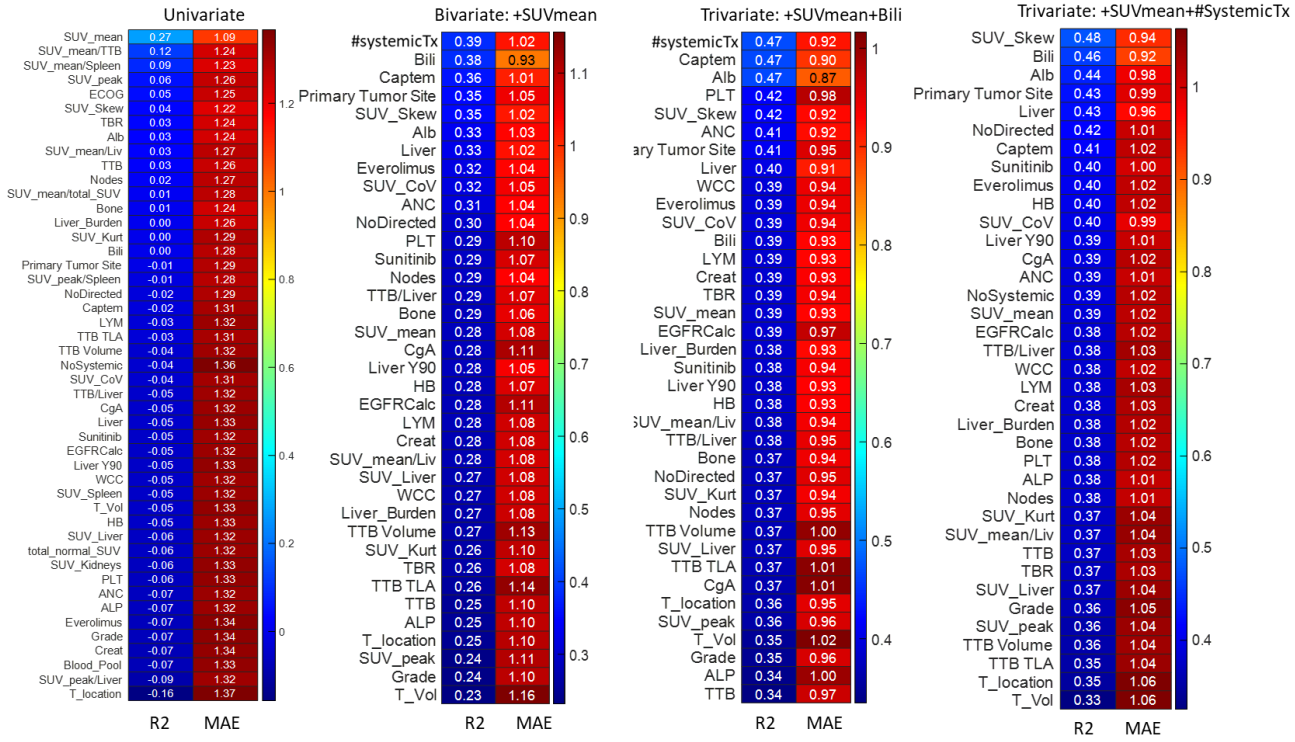


Supplemental-Figure 4. Spearman rank correlation between dose parameters and SUV quantities along with biomarkers in different subsamples of tumor volume (q-value<0.05).

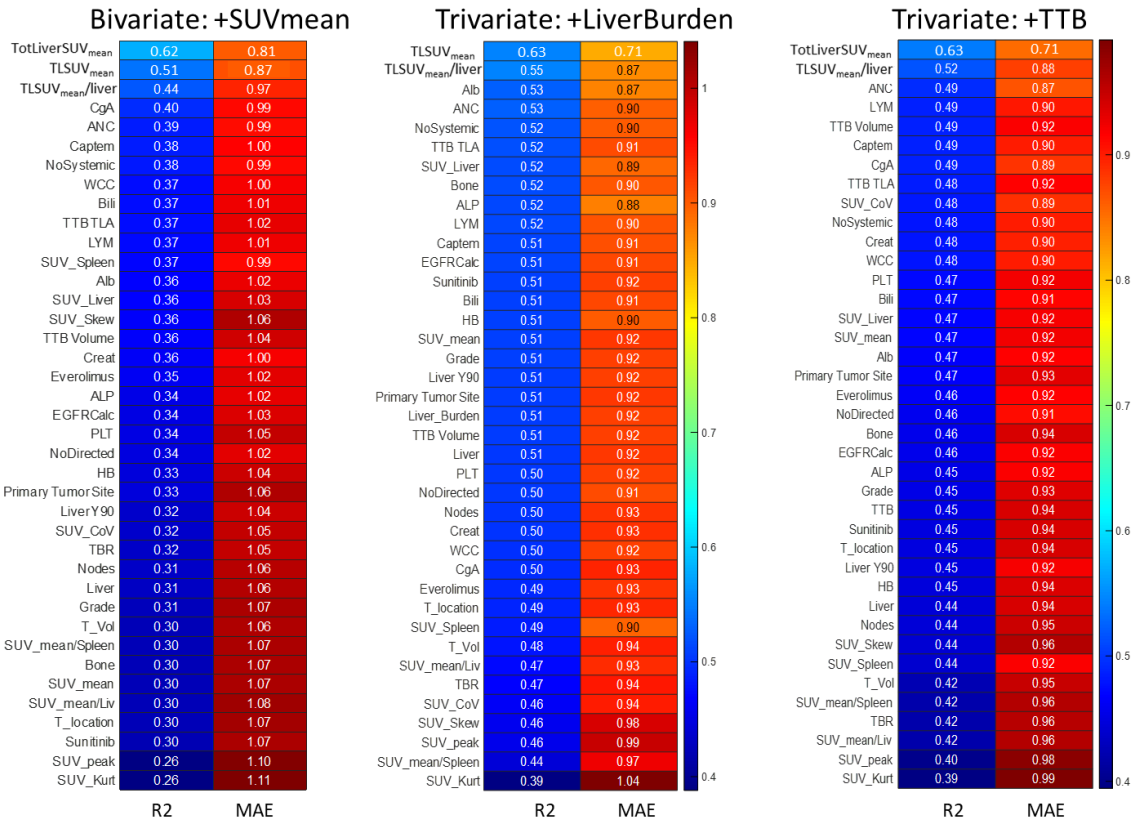
We first examined the correlation between dose-related parameters, i.e. TIA and C normalized by tumor volume (TIA_{vol}, C_{vol}), and T_{eff}. Presuming that local energy deposition is a valid approximation for 177Lu-labeled radionuclides' dosimetry [23], we expect a physics-informed correlation between dose and TIA_{vol} and further with C_{vol} and T_{eff}. Physically, the self-dose is composed of these parameters in addition to an independent noise level owing to the mathematical process. The term "noise" refers to the non-linear transformation applied on TIA during the MC-based dose calculation that degrades the linear correlation between TIA and dose. In this regard, we designed two strategies to test the strength of correlation between predictors and dose.

1. Comparing the correlations between predictors and three dose-related parameters, i.e. dose, TIAvol and C_{1vol}.

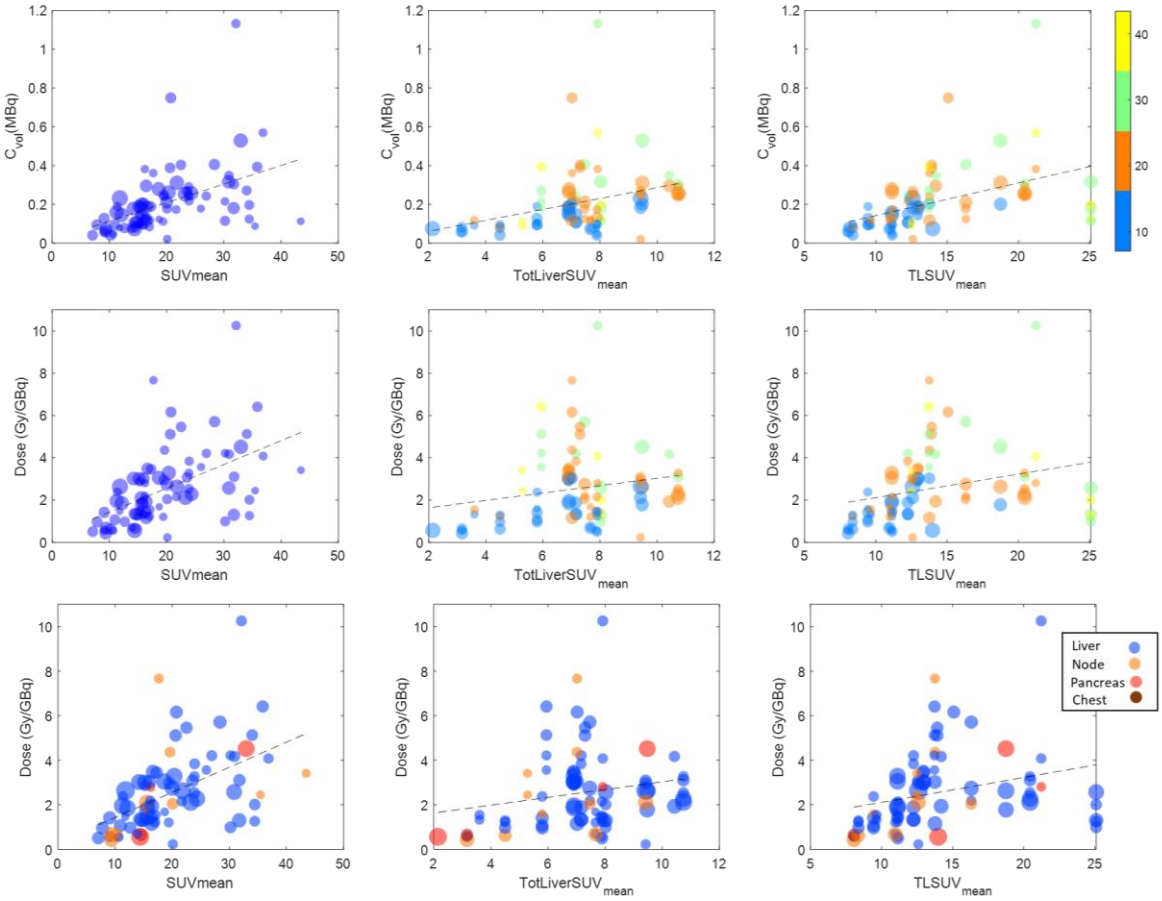
2. Comparing the correlations between predictors and dose parameters in multiple subsamples of the data filtered by tumor volume.



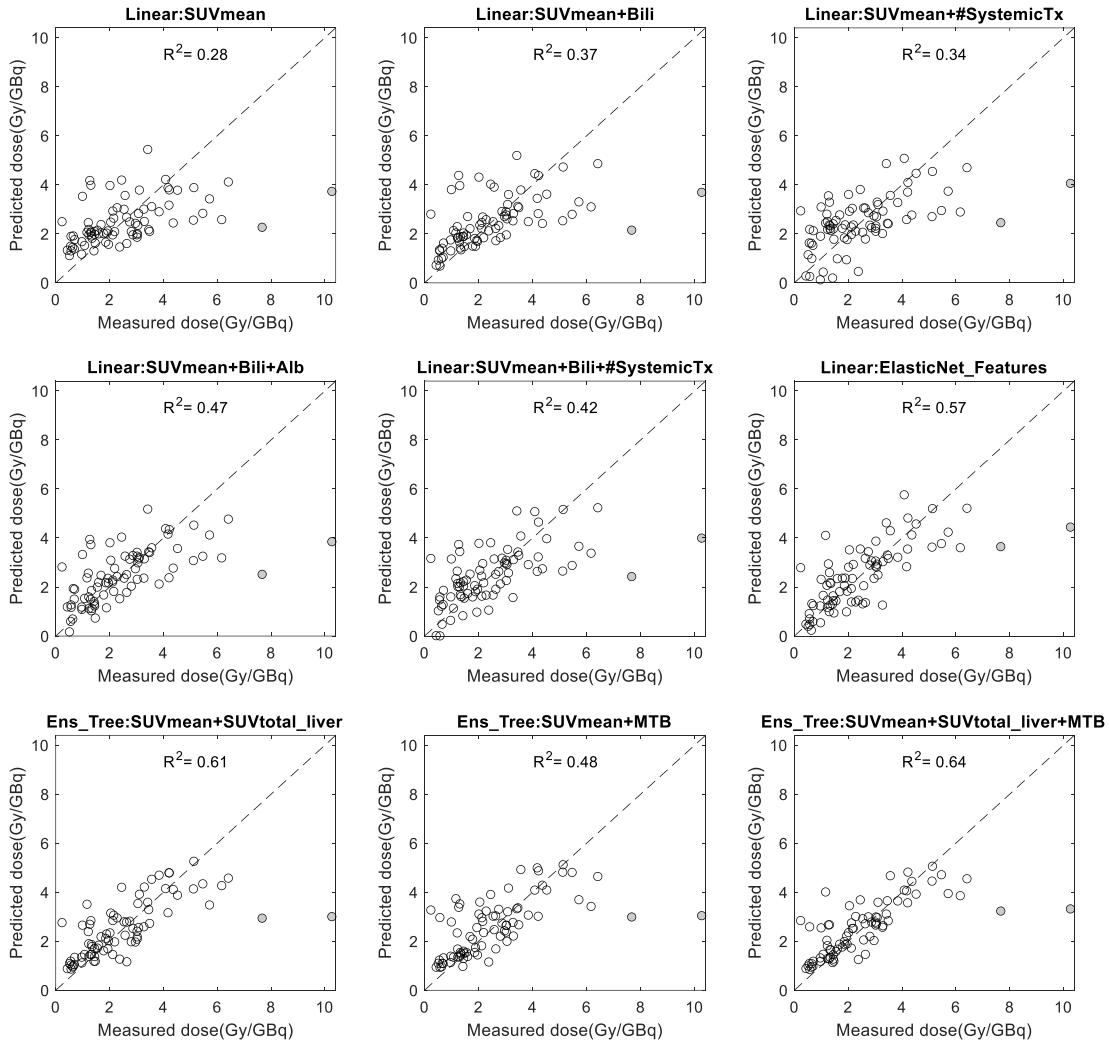
Supplemental-Figure 5. Model performance in hierarchical feature selection approach: Linear Univariate, Bivariate and Trivariate analysis (10-fold, Bootstrap aggregation models).



Supplemental-Figure 6. Model performance in hierarchical feature selection approach: Ens-Trees, Bivariate and trivariatate analysis (10-fold, Bootstrap aggregation models).



Supplemental-Figure 7. Tumor absorbed dose and C_{vol}, plotted vs. tumor SUV_{mean}, $TotLiverSUV_{mean}$ and $TLSUV_{mean}$, where the color bar shows the tumor SUV_{mean}; (bottom row) Tumor absorbed dose plotted vs. tumor SUV_{mean}, $TotLiverSUV_{mean}$ and $TLSUV_{mean}$, where the color bar shows the tumor location. The size of spheres depicts the volume of tumors in logarithmic form (4-1039 mL).



Supplemental-Figure 8. Pre-therapy predicted dose using different models of Table 2 vs. the delivered dose measured from Lu-177 SPECT/CT (the filled gray dots represent the 2 outliers).

Supplemental-Table 1. Patient demographic information

Demographics	Median	Range
Age (y)	63.7 (9.5)	38-76
Weight (kg)	85.5	51-129
Gender	**N (%)**	
Female	44	
Male	66	
Diabetes	20	
Hypertension	60	
ECOG (score: 0-5)		
S=0	44	
S=1	44	
S=2	12	
Krenning score		
S=3	12	
S=4	88	
Administered Lutathera		
~ 7.3 (GBq)	88	7.1-7.4
~ 3.7 (GBq)	12	3.7-3.8
68Ga PET/CT scanner		**Reconstruction**
Biograph mCT (Siemens Healthineers)	88	3D-OSEM, TOF, 3 iteration (21subsets), 5-mm Gauss filter
Biograph TruePoint (Siemens Healthineers)	4	
Discovery MI (GE Healthcare)	4	
and Discovery STE (GE Healthcare)	4	

ECOG: Eastern Cooperative Oncology Group

Supplemental-Table 2. Model performance of different prediction algorithms using PET-SUV_{mean} with/ without clinical factors. The quantitative metrics are reported as mean (95% CI) calculated from nested CV. The MAE range is reported based on the averaging over 10-outerloop CV point prediction.

Model		Features	R2* 10-fold	Median MRAE	MAE (Gy/GBq)	MAE Quantile (0.05-0.95)
Univariate	Linear					
		SUV_{mean}	0.28(0.00)	0.38(0.00)	1.08(0.00)	0.14-2.8
Bivariate	Linear					
		SUV_{mean} +bilirubin	0.37(0.01)	0.31(0.01)	0.95(0.01)	0.05-2.94
		SUV_{mean} + #Systemic therapy	0.34(0.02)	0.40(0.02)	1.06(0.01)	0.10-2.61
Trivariate	Linear					
		SUV_{mean} + bilirubin +albumin	0.47(0.00)	0.26(0.00)	0.87(0.00)	0.04-2.63
		SUV_{mean} + bilirubin +#Systemic therapy	0.42(0.02)	0.34(0.02)	0.94(0.01)	0.04-2.68
Multi-variate ElasticNet features	Linear	SUV_{mean}+SUV_{Liver}+SUV_{kidney}+ #Systemic therapy +#Directed therapy + bilirubin +albumin	0.57(0.02)	0.26(0.01)	0.80(0.02)	0.06-2.56
Bivariate	Tree_Ens					
		SUV_{mean} + $TotLiverSUV_{mean}$	0.61(0.01)	0.26(0.01)	0.82(0.01)	0.10-2.29
		SUV_{mean} +$TLSUV_{mean}$	0.48(0.03)	0.26(0.01)	0.88(0.02)	0.05-2.66
Trivariate	**Tree_Ens**					
		SUV_{mean}+ $TotLiverSUV_{mean}$+ $TLSUV_{mean}$	**0.64(0.02)**	**0.20(0.01)**	**0.73(0.02)**	**0.02-2.46**
Multi-variate PRFvI features	Tree-Ens	SUV_{mean} + $TLSUV_{mean}$ + SUV_{Spleen} + SUV_{Skew} + $TotLiverSUV_{mean}$ + SUV_{Liver} + SUV_{peak} + ALP	0.46(0.02)	0.26(0.01)	0.90(0.02)	0.06-2.59

***** Two outliers are excluded from R2

References

[1] R. P. Baum and H. R. Kulkarni, "THERANOSTICS: From Molecular Imaging Using Ga-68 Labeled Tracers and PET/CT to Personalized Radionuclide Therapy - The Bad Berka Experience,"*Theranostics,* vol. 2, no. 5, pp. 437-47, 2012.

[2] K. K. Wong *et al.*, "Differences in tumor-to-normal organ SUV ratios measured with 68Ga-DOTATATE PET compared with 177Lu-DOTATATE SPECT in patients with neuroendocrine tumors,"*Nuclear Medicine Communications,* vol. 43, no. 8, pp. 892-900, 2022.

[3] J. Strosberg *et al.*, "Phase 3 Trial of (177)Lu-Dotatate for Midgut Neuroendocrine Tumors,"*N Engl J Med,* vol. 376, no. 2, pp. 125-135, 2017.

[4] C. Miller, J. Rousseau, C. F. Ramogida, A. Celler, A. Rahmim, and C. F. Uribe, "Implications of physics, chemistry and biology for dosimetry calculations using theranostic pairs,"*Theranostics,* vol. 12, no. 1, pp. 232-259, 2022.

[5] M. Del Prete *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,* vol. 46, no. 3, pp. 728-742, 2019.

[6] A. Sundlöv *et al.*, "Phase II trial demonstrates the efficacy and safety of individualized, dosimetry-based (177)Lu-DOTATATE treatment of NET patients,"*Eur J Nucl Med Mol Imaging,* vol. 49, no. 11, pp. 3830-3840, 2022.

[7] E. Ilan *et al.*, "Dose response of pancreatic neuroendocrine tumors treated with peptide receptor radionuclide therapy using 177Lu-DOTATATE,"*J Nucl Med,* vol. 56, no. 2, pp. 177-82, 2015.

[8] J. Violet *et al.*, "Dosimetry of (177)Lu-PSMA-617 in Metastatic Castration-Resistant Prostate Cancer: Correlations Between Pretherapeutic Imaging and Whole-Body Tumor Dosimetry with Treatment Outcomes,"*J Nucl Med,* vol. 60, no. 4, pp. 517-523, 2019.

[9] S. Xue *et al.*, "Application of machine learning to pretherapeutically estimate dosimetry in men with advanced prostate cancer treated with (177)Lu-PSMA I&T therapy,"*Eur J Nucl Med Mol Imaging,* vol. 49, no. 12, pp. 4064-4072, 2022.

[10] S. M. B. Peters *et al.*, "[(68)Ga]Ga-PSMA-11 PET imaging as a predictor for absorbed doses in organs at risk and small lesions in [(177)Lu]Lu-PSMA-617 treatment,"*Eur J Nucl Med Mol Imaging,* vol. 49, no. 4, pp. 1101-1112, 2022.

[11] A. Peterson *et al.*, "Lu-177-DOTATATE Theranostics: Predicting Renal Dosimetry from Pre-therapy Ga-68-DOTATATE PET and Clinical Biomarkers,"vol. (in press), 2022.

[12] D. Roth *et al.*, "Dosimetric Quantities in Neuroendocrine Tumors over Treatment Cycles with (177)Lu-DOTATATE,"*J Nucl Med,* vol. 63, no. 3, pp. 399-405, 2022.

[13] M. Lassmann and U. Eberlein, "The Relevance of Dosimetry in Precision Medicine,"*J Nucl Med,* vol. 59, no. 10, pp. 1494-1499, 2018.

[14] M. Öksüz *et al.*, "Peptide receptor radionuclide therapy of neuroendocrine tumors with (90)Y-DOTATOC: is treatment response predictable by pre-therapeutic uptake of (68)Ga-DOTATOC?,"*Diagn Interv Imaging,* vol. 95, no. 3, pp. 289-300, 2014.

[15] V. Ambrosini *et al.*, "Prognostic Value of 68Ga-DOTANOC PET/CT SUVmax in Patients with Neuroendocrine Tumors of the Pancreas,"*J Nucl Med,* vol. 56, no. 12, pp. 1843-8, 2015.

[16] Y. K. Dewaraja *et al.*, "A Pipeline for Automated Voxel Dosimetry: Application in Patients with Multi-SPECT/CT Imaging After (177)Lu-Peptide Receptor Radionuclide Therapy,"*J Nucl Med,* vol. 63, no. 11, pp. 1665-1672, 2022.

[17] J. Tran-Gia and M. Lassmann, "Characterization of Noise and Resolution for Quantitative (177)Lu SPECT/CT with xSPECT Quant,"*J Nucl Med,* vol. 60, no. 1, pp. 50-59, 2019.

[18] S. J. Wilderman and Y. K. Dewaraja, "Method for Fast CT/SPECT-Based 3D Monte Carlo Absorbed Dose Computations in Internal Emitter Therapy,"*IEEE Trans Nucl Sci,* vol. 54, no. 1, pp. 146-151, 2007.

[19] S. Cui, Y. Luo, H. H. Tseng, R. K. Ten Haken, and I. El Naqa, "Combining handcrafted features with latent variables in machine learning for prediction of radiation-induced lung damage,"*Med Phys,* vol. 46, no. 5, pp. 2497-2511, 2019.

[20] P. Peduzzi, J. Concato, A. R. Feinstein, and T. R. Holford, "Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates,"*J Clin Epidemiol,* vol. 48, no. 12, pp. 1503-10, 1995.

[21] L. Breiman, "Random forests,"*Machine learning,* vol. 45, no. 1, pp. 5-32, 2001.

[22] Y. Lou, R. Caruana, and J. Gehrke, "Intelligible models for classification and regression," in *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining,* 2012, pp. 150-158.

[23] K. Sjögreen Gleisner *et al.*, "EANM dosimetry committee recommendations for dosimetry of 177Lu-labelled somatostatin-receptor- and PSMA-targeting ligands,"*European Journal of Nuclear Medicine and Molecular Imaging*, vol. 49, no. 6, pp. 1778-1809, 2022.

[24] B. Singh, V. Prasad, C. Schuchardt, H. R. Kulkarni, and R. P. Baum, "Can the Standardized Uptake Values derived from Diagnostic 68 Ga-DOTATATE PET/CT Imaging Predict the Radiation Dose delivered to the Metastatic Liver NET Lesions on 177 Lu-DOTATATE Peptide Receptor Radionuclide Therapy?,"*Journal of Postgraduate Medicine, Education and Research*, vol. 47, pp. 7-13, 2013.

[25] M. Lubberink *et al.*, "In Vivo Instability of (177)Lu-DOTATATE During Peptide Receptor Radionuclide Therapy,"*J Nucl Med*, vol. 61, no. 9, pp. 1337-1340, 2020.

[26] S. Ezziddin *et al.*, "Does the pretherapeutic tumor SUV in 68Ga DOTATOC PET predict the absorbed dose of 177Lu octreotate?,"*Clin Nucl Med*, vol. 37, no. 6, pp. e141-7, 2012.

[27] Y.-i. Kim, J. Oh, C. Yoo, B.-Y. Ryoo, and J.-S. Ryu, "Prediction of absorbed dose by tumors and critical organs after Lu-177-DOTATATE therapy using pretherapeutic Ga-68-DOTATOC PET/CT," ed: Soc Nuclear Med, 2021.

[28] R. Bruvoll, J. Blakkisrud, L. Mikalsen, J. Connelly, and C. Stokke, "Correlations between [68Ga] Ga-DOTA-TOC uptake and absorbed dose from [177Lu] Lu-DOTA-TATE," 2022.

[29] H. Hänscheid *et al.*, "PET SUV correlates with radionuclide uptake in peptide receptor therapy in meningioma,"*Eur J Nucl Med Mol Imaging*, vol. 39, no. 8, pp. 1284-8, 2012.

[30] E. Ilan, I. Velikyan, M. Sandström, A. Sundin, and M. Lubberink, "Tumor-to-Blood Ratio for Assessment of Somatostatin Receptor Density in Neuroendocrine Tumors Using (68)Ga-DOTATOC and (68)Ga-DOTATATE,"*J Nucl Med*, vol. 61, no. 2, pp. 217-221, 2020.

[31] A. Kroiss *et al.*, "68Ga-DOTA-TOC uptake in neuroendocrine tumour and healthy tissue: differentiation of physiological uptake and pathological processes in PET/CT,"*Eur J Nucl Med Mol Imaging*, vol. 40, no. 4, pp. 514-23, 2013.

[32] Y. I. Kim *et al.*, "Tumour-to-liver ratio determined by [(68)Ga]Ga-DOTA-TOC PET/CT as a prognostic factor of lanreotide efficacy for patients with well-differentiated gastroenteropancreatic-neuroendocrine tumours,"*EJNMMI Res*, vol. 10, no. 1, p. 63, 2020.

[33] E. Pauwels *et al.*, "[(68)Ga]Ga-DOTATATE-avid tumor volume, uptake and inflammation-based index correlate with survival in neuroendocrine tumor patients treated with [(177)Lu]Lu-DOTATATE PRRT,"*Am J Nucl Med Mol Imaging*, vol. 12, no. 5, pp. 152-162, 2022.

[34] A. Akhavanallaf, I. Shiri, H. Arabi, and H. Zaidi, "Whole-body voxel-based internal dosimetry using deep learning,"*Eur J Nucl Med Mol Imaging*, vol. 48, no. 3, pp. 670-682, 2021.

Chapter 12

Conclusion and future perspectives

I. Summary of the research and achievements

The use of ionizing radiation in medical imaging and radiotherapy increased significantly over the past three decades globally. Medical imaging examinations, especially CT and molecular imaging, carry small risks of radiation hazards that should be justified and furthermore optimized by "*keeping the exposure of patients to the minimum necessary to achieve the required diagnostic or interventional objective*"[1]. Current radiopharmaceutical therapy regimens are in a transition phase from one-size-fits-all concept to a personalized approach by increasing the radiation dose to the target while minimizing the absorbed dose to healthy tissues. Therefore, establishing a practical framework for patient-specific dosimetric data estimation can be used in the optimization of radiation-involved medical procedures to ensure the minimum radiation dose necessary, while improving the efficacy of the medical task at hand.

This dissertation aimed at evaluating and developing unified frameworks that were developed to offer accurate and practical solutions for the key challenges and major limitations of personalized dosimetry in nuclear medicine procedures. In Part 1, we mainly focused on the two components of personalized dosimetry: computational phantoms and radiation transport algorithms. Toward patient-specific computational models, we started with the development of a library of habitus-dependent computational phantoms and extended our methodology to construct patient-specific phantoms through the registration of reference computational models on patient's CT images. However, taking advantage of the emergence of deep learning algorithms in the domain of computer vision and image processing, nowadays, the construction of patient-specific computational models from CT images is a practical task in clinical setting. In addition, we have developed and validated different MC simulators for personalized dosimetry from various medical exposure scenarios and further addressed the limitation of MC-based dose engines in terms of computational cost by implementing deep learning in dose reconstruction. We developed a unified framework for the deployment of deep learning algorithms for fast dose calculation in nuclear medicine, brachytherapy and external CT exposure with a reasonable accuracy compared to MC-based dose maps serving as standard of reference (ground truth). In addition to our studies in dose monitoring, we conducted a study on radiation dose optimization of CT acquisition protocols by means of deep learning and demonstrated the feasibility of ultra-low dose imaging in chest CT examinations with acceptable clinical diagnosis accuracy. Following our research on dosimetry in medical imaging, we extended the concepts and tools that we previously developed to theragnostic dosimetry. In this context, we developed the theragnostic dosimetry workflow for both ^{90}Y-SIRT and ^{177}Lu-DOTATATE therapy. As the result, a stusy was conducted on 177lu-DOTATATE radiopharmaceutical therapy that confirmed the predictive value of 68Ga- DOTATATE (theragnostic pair) for patient-stratification and personalized planning in RPT.

The main contributions achieved during this dissertation are summarized in the following paragraphs:

1. To cope with inter-subject variability of anatomical features, a habitus-dependent library of computational phantoms has been developed, covering the diversity of organ masses along with the morphometric parameters by adjusting voxel-based ICRP adult reference phantoms. The move toward patient-specific phantoms is a major improvement taking advantage of the availability of habitus-dependent phantoms associated with anthropomorphic and anatomical diversities and classified in different somatotypes.
2. We quantified the dosimetric characteristics of patient-specific computational models in CT dose estimation. Although using habitus-specific phantom series is feasible for dosimetry in clinical setting, the estimated organ dose may considerably differ from the ground truth (up to 36%). If, however, patient CT images are available, a reference computational phantom can be matched to the patient data to construct a patient-specific computational model through deformable registration, thus improving the accuracy of organ dose estimation.
3. We further adapted and validated our Monte Carlo simulation tool developed for dose calculation associated with CT examinations on a Siemens scanner. An experimental setup using an anthropomorphic physical phantom and TLDs was designed to evaluate the accuracy of MC-based personalized organ-level dosimetry. Using the validated CT dosimetry simulator, patient-specifics dosimetry, and moreover, optimization of CT technologies and scanning protocols would be feasible. In our study, we also assessed the dosimetric impact of input parameters in organ-level dose simulation. It can be concluded that, when the information from the

[1] IAEA, Safety Standards Series No. SSG-46

CT raw projection data is not available, the simulation results could be acceptable if the longitudinal tube current modulation (with <6 mm interval in z direction), is implemented in the simulation.

4. We proposed a unified methodology for patient-specific voxelwise whole-body internal dosimetry using deep learning algorithms. The comparison of the proposed approach with standard of reference MC simulations revealed very good accuracy with a mean relative absolute error of 2.6%. Our technique also outperformed conventional voxel-level and organ-level MIRD-based formalisms. Future work will focus on exploiting the current model to generate whole-body voxelwise dose maps for the purpose of using active learning to improve model performance and further extend this model to a faster pipeline through straightforward prediction of whole-body dose maps from hybrid PET/CT images.
5. We developed and validated a unified pipeline for MC-based dosimetry in high-dose-rate brachytherapy that has been used to provide an accurate set of MC simulations on a large retrospective cohort. We further developed a deep learning model to provide an alternative solution for accurate personalized dose distribution estimation in brachytherapy to overcome the computational burden of MC simulations. The proposed algorithm achieved good agreement with MC calculations while outperforming the conventional TG-43-based formalism. Future work will focus on extending the core idea to different radioactive seeds for various disease sites.
6. We proposed a deep learning model for construction of whole-body dose maps from CT scans with reasonable accuracy at the voxel level with excellent performance achieved for organ-level dose estimation. The whole process, including pre-processing and model inference on a new dataset, can be performed within seconds, which makes personalized dosimetry with an acceptable accuracy a feasible option in clinical setting. The main advantage of our model is its ability to construct accurate and personalized dose maps for a wide range of acquisition parameters.
7. Ultra-low-dose chest CT imaging of COVID-19 patients would result in the loss of critical information about lesion types. However, the results presented in this work indicated that ResNet is an optimal algorithm for denoising ultra-low-dose CT images for COVID-19 diagnosis. Future work will focus on a more accurate low-dose CT simulation algorithm along with the extension of the model to whole-body clinical CT studies to further validate our model.
8. Keeping the dose-response relationship in mind, the current study investigated the predictive values of ^{68}Ga-based SUVs along with biomarkers to predict the therapeutic tumor dose. We showed that tumor SUV_{mean}, $SUV_{total\text{-}liver}$ and SUV_{mean} of MTV is capable of predicting the ^{177}Lu-PRRT induced tumor dose with an accuracy in terms of MAE of 0.71 Gy/GBq (R^2=0.63). We foresee to further benchmark the proposed models on a multi-center study to provide a validated decision-support tool for clinicians to improve patient-selection and thus optimize treatment outcome. Developing these quantitative metrics forms the ground for the role of ^{68}Ga-PET not only for patient stratification but also for prognostication and assessment of therapeutic response modeling.

II. Future perspectives

This dissertation focused mainly on personalized dosimetry in diagnostic and therapeutic nuclear medicine examinations. On the basis of advancements in quantitative molecular imaging technologies, computational modeling of the human body, and sophisticated radiation transport techniques, nowadays, patient-specific dosimetry is feasible in routine clinical setting. Through the advent of artificial intelligence and in particular deep learning, in the area of computer vision and image processing, the task of segmentation, one of the main challenges in personalized medicine, has been addressed. Currently, robust automatic segmentation is a hot topic in medical image analysis for organs segmentation from structural images (such as CT and MRI) and the detection and delineation of tumoral tissues from different medical images such as PET, contrast-enhanced CT, MRI, pathological slides, … etc. Logically, prospective research in this field would deal with further development, evaluation and commercialization of the proposed segmentation algorithms in real-world clinical setting. To be specific in the field of personalized dosimetry, real time organ dosimetry from external exposure is becoming feasible thanks to fast organ segmentation algorithms.

In the case of CT dose optimization, deep learning-based low-dose CT image denoising is one of the promising approaches. While a novel technical approach could be real-time prediction of the 3D voxel density map of the

patient from 2D CT localizer (scout image) and accordingly modulating the tube current based on the predicted density map.

In radiopharmaceutical dosimetry, simplified radiation transport methods, like the MIRD formalism and local energy deposition provide clinically acceptable results for therapeutic radiotracers with short range particles (i.e beta and alpha) in soft-tissue regions, while in heterogenous medium, such as bone marrow dosimetry, the simplified approaches provide some uncertainties compared to MC simulations. In the case of imaging radiotracers labelled with longer travelling range radiation (i.e. photons) compared to therapeutic agents, simplified radiation transport approaches introduce considerable errors compared to patient-specific MC dosimetry. Therefore, according to the promising results achieved by AI-based internal dose construction, this could be a potential option for clinical dosimetry in a reasonable computational time. However, fast GPU-based Monte Carlo simulations using high performance computers and the next generation of quantum computers seems more reliable for use in clinical setting. In addition, patient-specific biokinetics of imaging radiotracers cannot be extracted from static image acquisitions. It would be interesting to build a simple data-driven model for the prediction of time-integrated activity from dynamic scans similar to studies in RPT dosimetry with single time points.

In theragnostic dosimetry, there are some gaps in image quatification because of motion that has not been adequately addressed yet. Respiratory motion degrades image quality and impacts image quantification from attenuation correction step to the construction of density maps for radiation transport. One potential solution for attenuation correction is using deep learning-based attenuation correction algorithms, as these methods showed promising results to compensate the impact of mismatch between CT and ^{18}F-FDG PET examinations resulting from respiration motion.

In addition, a better understanding of radiobiology in molecular radionuclide therapy is needed. Radiobiology has been a key factor for establishing optimal treatment regimens for external beam radiotherapy. Nowadays, there is some evidence that the extrapolation of radiobiology of external beam radiotherapy to molecular radionuclide therapy is not straightforward, because of dose-rate effects and more importantly owing to the different molecular and cellular signalling pathways. Therefore, there is a need for restablishment of specific radiobiological models in RPT.

In the current clinical scenario, there is minimal evidence on patient eligibility for RPT, in particular for ^{177}Lu-DOTATATE and PSMA, and no standardized criteria has been established yet. Thus, there is still room for outcome prediction of these type of treatments with encouraging responses and well tolerable side effects that can aid clinicians to further optimize clinical trial designs and individual patient management. Furthermore, potential prognostic models are useful tools in designing randomized clinical trials for patients' selection and stratification of patients to responder and non-responder groups. Considering the fact that the developed predictive model should not bias well-informed clinical decisions, instead, it has to be employed as a complementary decision-support tool. In addition, using dual tracer (e.g. ^{68}Ga-PSMA and ^{18}F-FDG) in RPT prognostic modelling is a hot topic that needs more clinical trials and investigations to build a conclusive understanding of the predictive value of these models.

Acknowledgements

I would like to express my appreciation and gratitude to my supervisor, Prof. Habib Zaidi, for his unwavering support, guidance, and mentorship throughout my PhD journey. His expertise, encouragement, and constructive feedback have been invaluable in shaping my research and academic development.

I extend my sincere appreciation to Prof. Yuni Dewaraja, who gave me an invaluable opportunity to work in her lab at the University of Michigan as a visiting scholar. I would also like to thank the members of my thesis committee, Prof. Wesley Bolch, Prof. Manuel Bardies, Prof. Arman Rahmim and Prof. Karsten Kruse, for their valuable feedback, constructive criticism, and insightful suggestions that helped me improve the quality of my work.

My sincere thanks go to my PinLab colleagues and Umich colleagues, the faculty members and staff in the department of nuclear medicine at Geneva University Hospital, who have provided me with a stimulating academic environment, invaluable resources, and supportive community that have enriched my academic and personal growth.

Words cannot express my gratitude to my nurturing parents for their love and support during all steps of my life from my childhood up to now. I also thank my brother and sister who always encouraged me and motivated me through my life.

I also would like to express my deepest appreciation to my lovely husband who have supported me throughout my PhD journey, listened to my ideas, soothed my concerns, and moreover tolerated and stood with me, non-stoply. Your encouragement, love, and unwavering belief in me have sustained me through this journey. Last but not least, my adorable twins who made me a responsible mother while trying to be a PhD student! They were more than Bless to my life and specifically to my PhD journey!

Thank you all for contributing to my success, and I hope this thesis will make a meaningful contribution to the improvement of patient care.

Azadeh Akhavanallaf

7th April 2023

List of publications

Peer-reviewed journals

1. **Akhavanallaf A**, Fayad H, Salimi Y, Aly A, Kharita H, Al Naemi H and Zaidi H, An update on computational anthropomorphic anatomical models, Dig Health Vol. 8, pp 20552076221111941 (2022).
2. **Akhavanallaf A***, Xie T*, and Zaidi H, Development of a library of adult computational phantoms based on anthropometric indexes, IEEE Trans Rad Plasma Med Sci, Vol. 3, No. 1, pp 65-75 (2019).
3. Xie T* , **Akhavanallaf A***, and Zaidi H, Construction of patient-specific computational models for organ dose estimation in radiological imaging, Med Phys, Vol. 46, No.4, pp 1519-1522 (2019).
4. **Akhavanallaf A**, Xie T, and Zaidi H, Assessment of uncertainties associated with Monte Carlo-based personalized dosimetry in clinical CT examinations, Phys Med Biol, Vol. 65, No. 4, pp 045008 (2020).
5. Arabi H, **Akhavanallaf A**, Shiri I, Sanaat A, Zaidi H, The promise of artificial intelligence and deep learning in PET and SPECT imaging, Phys Med, Vol. 83, pp 122-137 (2021).
6. Shiri I*, **Akhavanallaf A***, Sanaat A, et al, Ultra-low dose chest CT imaging of Covid-19 patients using deep neural networks, Eur Radiol, 31(3), 1420–1431 (2021).
7. Salimi Y, **Akhavanallaf A**, Mansouri Y, Shiri I and Zaidi H, Real-time, acquisition parameter-free, voxelwise patient-specific Monte Carlo dose reconstruction in total-body CT scanning using deep neural networks, Eur Radiol (2022) *under revision*.
8. **Akhavanallaf A**, Shiri I, Arabi H, and Zaidi H, Whole-body voxel-based internal dosimetry using deep learning" Eur J Nucl Med Mol Imaging Vol. 48, No. 3, pp 670-682 (2021).
9. **Akhavanallaf A**, Mohammadi R, Shiri I, Salimi Y, Arabi H, and Zaidi H, Personalized brachytherapy dose reconstruction using deep learning, Comput Biol Med, Vol. 136, pp 104755 (2021).
10. **Akhavanallaf A,** Peterson A, Fitzpatrick K, Roseland M, Wong KK, El Naqa I, Zaidi H and Dewaraja Y. The predictive value of ^{68}Ga-PET and baseline biomarkers in ^{177}Lu-DOTATATE tumor dosimetry, Eur J Nucl Med Mol Imaging (2023) *under revision.*
11. Yixuan Jia, Zongyu Li, **Azadeh Akhavanallaf**, Jeffrey A. Fessler, Yuni K. Dewaraja, 90Y SPECT Scatter Estimation and Voxel Dosimetry in radioembolization Using a Unified Deep Learning Framework" EJNMMI Phys (2023) *submitted*.
12. Riveira M, **Akhavanallaf A**, Bianchetto Wolf N, Salimi Y, Mainta I, Garibotto V, Lopez Medina AM and Zaidi H "Predictive Value of 99mTc-MAA-based Theragnostic Dosimetry in Personalized 90Y-SIRT planning of Liver Malignancies" Eur J Nucl Med Mol Imaging (2023) *submitted*.
13. Shiri I, Sadr AV, **Akhavanallaf A**, Salimi Y, Sanaat A, Amini M, Razeghi B, Saberi A, Arabi H, Ferdowsi S, Voloshynovskiy S, Gündüz D, Rahmim A and Zaidi H "Decentralized collaborative multi-institutional PET attenuation and scatter correction using federated deep learning" Eur J Nucl Med Mol Imaging (2023) *in press*.
14. Salimi Y, Shiri I, **Akhavanallaf A**, Mansouri Z, Arabi H and Zaidi H "Fully automated accurate patient positioning in computed tomography using anterior-posterior localizer images and a deep neural network: A dual-center study" Eur Radiol (2023) *in press*.
15. Asadi A, **Akhavanallaf A**, Hossein SA, Vosoughi N and Zaidi H "Development and validation of an optimal GATE model for proton pencil-beam scanning delivery" Z Med Phys (2023) *in press*.
16. Salimi Y, Shiri I, **Akhavanallaf A**, Mansouri Z, Sanaat A, Pakbin M, Hossein MG, Arabi H, and Zaidi H "Deep learning-based calculation of patient size and attenuation surrogates from localizer image: Toward personalized CT protocol optimization for thorax imaging" Eur J Radiol Vol. 157, pp 110602 (2022).
17. Asadi A, Hossein SA, **Akhavanallaf A**, Vosoughi N and Zaidi H "Comparative study of passive scattering and active scanning proton therapy techniques using Monte Carlo simulations" J Instrum Vol. 17, No. 9, pp P09008 (2022).

*Authors contributed equally

18. Sanaat A, **Akhavanallaf A**, Shiri I, Salimi Y, Arabi H and Zaidi H "Deep-TOF-PET: Deep learning-guided generation of time-of-flight (TOF) PET from non-TOF brain PET images in the image and projection domains" Hum Brain Mapp Vol. 43, No. 16, pp 5032-5043 (2022).
19. Shiri I, Mostafaei S, Avval AH, Salimi Y, Sanaat Y, **Akhavanallaf A**, Arabi H, Rahmim A and Zaidi H "High-dimensional multinomial multiclass severity scoring of COVID-19 pneumonia using CT radiomics features and machine learning algorithms" Sci Rep Vol. 12, No. 1, pp 14817 (2022).
20. Shiri I, Sadr AV, Amini M, Salimi Y, Sanaat A, **Akhavanallaf A**, Razeghi B, Ferdowsi S, Saberi A, Arabi H, Becker M, Voloshynovskiy S, Gündüz D, Rahmim A and Zaidi H, "Decentralized distributed multi-institutional PET image segmentation using a federated deep learning framework" Clin Nucl Med Vol. 47, No 7, pp 606-617 (2022).
21. Shiri I, Arabi H, Salimi Y, Sanaat A, **Akhavanallaf A**, Hajianfar G, Askari D, Moradi S, Mansouri Z, Pakbin M, Sandoughdaran S, Abdollahi H, Radmard AR, Rezaei-Kalantari K, Oghli MG, and Zaidi H"COLI-NET: Fully automated COVID-19 lung and infection pneumonia lesion detection and segmentation from chest CT images" Int J Imaging Syst Technol Vol. 32, No. 1, pp 12-25 (2022).
22. Salimi Y, Shiri I, **Akhavanallaf A**, Mansouri Z, Saberi A, Sanaat A, Pakbin M, Askari D, Sandoughdaran S, Sharifipour E, Arabi H, and Zaidi H "Deep learning-based fully automated Z-axis coverage range definition from scout scans to eliminate overscanning in chest CT scanning" Insights Imaging Vol. 12, No. 1, pp 162 (2021).

Conferences proceedings and abstracts

23. Akhavanallaf A, Xie T and Zaidi H "Development of a novel library of adult anthropomorphic phantoms with more realistic scaling of organ masses" Proc. of the 6th International Workshop on Computational Human Phantoms, Annapolis (MD), USA, 27-30 August 2017, pp 120-121.
24. Akhavanallaf A, Xie T and Zaidi H "Assessment of uncertainties associated with Monte Carlo simulation of clinical CT examinations" Proc. of the 7th International Workshop on Computational Human Phantoms, Neuherberg, Germany, 22-24 July 2019, pp 1-2.
25. Shiri I, Akhavanallaf A, Sanaat A, Askari D, Salimi Y, Mansouri Z, Shayesteh SP, Rezaei-Kalantari K, Salahshour A, Sandoughdaran S, Abdollahi H, Arabi H and Zaidi H "Ultra-low dose chest CT imaging of COVID-19 patients using deep neural networks" IEEE Nuclear Science Symposium & Medical Imaging Conference, Boston, MA, USA, 31 October - 7 November 2020.
26. Akhavanallaf A, Shiri I, Arabi H and Zaidi H "Deep learning-assisted whole-body voxel-based internal dosimetry" IEEE Nuclear Science Symposium & Medical Imaging Conference, Boston, MA, USA, 31 October - 7 November 2020. IEEE Nuclear Science Symposium & Medical Imaging Conference, Boston, MA, USA, 31 October - 7 November 2020.
27. Akhavanallaf A, Salimi Y, Shiri I, Arabi A, Rahmim A and Zaidi H "Deep learning-based dosimetry in radionuclide therapy: Is it worth the effort?" IEEE Virtual Nuclear Science Symposium & Medical Imaging Conference, Yokohama, Japan, 16-23 October 2021.
28. Salimi Y, Akhavanallaf A, Hajianfar G, Sanaat A, Shiri I, Arabi H and Zaidi H "Deep learning-based fully automated scan range detection in chest CT imaging" IEEE Virtual Nuclear Science Symposium & Medical Imaging Conference, Yokohama, Japan, 16-23 October 2021.
29. Salimi Y, Akhavanallaf A, Mansouri Z, Shiri I and Zaidi H "Real-time voxelwise patient-specific Monte Carlo dose reconstruction in whole-body CT scan images using deep neural network" IEEE Nuclear Science Symposium & Medical Imaging Conference, Milan, Italy, 5-12 November 2022.
30. Akhavanallaf A, Mohammadi R, Shiri I, Salimi Y, Arabi H and Zaidi H "Personalized brachytherapy dose reconstruction using deep learning" 2021 American Brachytherapy Society Virtual Annual Conference, USA, 22–25 June 2021, Brachytherapy, Vol. 20, Issue 3, pp S21.
31. Riveira M, Akhavanallaf A, Bianchetto Wolf N, Salimi Y, Mainta I, Garibotto V, Lopez Medina AM and Zaidi H "Predictive value of dose metrics from 99mTc-MAA compared to 90Y SPECT/CT in dosimetry-guided personalized SIRT of hepatocellular carcinoma" 68th SNMMI Annual Meeting, Vancouver, Canada, 11–14 June 2022, J Nucl Med, Vol. 63, Suppl 2, pp 2807.