

PAPER • OPEN ACCESS

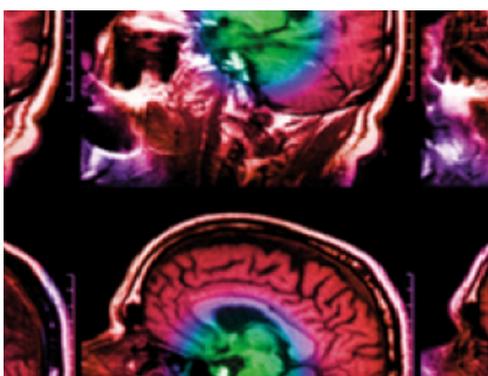
# A patient-specific hybrid phantom for calculating radiation dose and equivalent dose to the whole body

To cite this article: Erika Kollitz *et al* 2022 *Phys. Med. Biol.* **67** 035005

View the [article online](#) for updates and enhancements.

## You may also like

- [Inclusion of thin target and source regions in alimentary and respiratory tract systems of mesh-type ICRP adult reference phantoms](#)  
Han Sung Kim, Yeon Soo Yeom, Thang Tat Nguyen *et al.*
- [The UF family of reference hybrid phantoms for computational radiation dosimetry](#)  
Choonsik Lee, Daniel Lodwick, Jorge Hurtado *et al.*
- [Dosimetric impact of voxel resolutions of computational human phantoms for external photon exposure](#)  
Choonsik Lee, Andreu Badal, Yeon Soo Yeom *et al.*



**IPEM | IOP**

Series in Physics and Engineering in Medicine and Biology

Your publishing choice in medical physics,  
biomedical engineering and related subjects.

Start exploring the collection—download the  
first chapter of every title for free.



## PAPER

## A patient-specific hybrid phantom for calculating radiation dose and equivalent dose to the whole body

## OPEN ACCESS

## RECEIVED

28 September 2021

## REVISED

18 December 2021

## ACCEPTED FOR PUBLICATION

30 December 2021

## PUBLISHED

4 February 2022

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



Erika Kollitz<sup>1</sup> , Haegin Han<sup>2</sup> , Chan Hyeong Kim<sup>2</sup>, Marco Pinto<sup>1</sup> , Marco Schwarz<sup>3,4</sup> , Marco Riboldi<sup>1</sup>, Florian Kamp<sup>5</sup>, Claus Belka<sup>5</sup>, Wayne Newhauser<sup>6</sup>, George Dedes<sup>1,7</sup> and Katia Parodi<sup>1,7</sup> 

<sup>1</sup> Department of Medical Physics, Fakultät für Physik, Ludwig-Maximilians-Universität München (LMU Munich), D-85748 Garching bei München, Germany

<sup>2</sup> Department of Nuclear Engineering, Hanyang University, Seoul, Republic of Korea

<sup>3</sup> Azienda Provinciale per i Servizi Sanitari (APSS) Proton therapy Department, Trento, Italy

<sup>4</sup> TIFPA Trento Institute for Fundamental Physics and Applications, Italy

<sup>5</sup> Department of Radiation Oncology, University Hospital, LMU Munich, D-81377 Munich, Germany

<sup>6</sup> Department of Physics and Astronomy, Louisiana State University, Baton Rouge, Louisiana, United States of America

<sup>7</sup> These authors have contributed to this work equally.

E-mail: [E.Kollitz@physik.uni-muenchen.de](mailto:E.Kollitz@physik.uni-muenchen.de)

**Keywords:** computational phantoms, out-of-field dose, proton therapy, neutrons, equivalent dose

Supplementary material for this article is available [online](#)

## Abstract

**Objective.** As cancer survivorship increases, there is growing interest in minimizing the late effects of radiation therapy such as radiogenic second cancer, which may occur anywhere in the body. Assessing the risk of late effects requires knowledge of the dose distribution throughout the whole body, including regions far from the treatment field, beyond the typical anatomical extent of clinical computed tomography (CT) scans. **Approach.** A hybrid phantom was developed which consists of in-field patient CT images extracted from ground truth whole-body CT scans, out-of-field mesh phantoms scaled to basic patient measurements, and a blended transition region. Four of these hybrid phantoms were created, representing male and female patients receiving proton therapy treatment in pelvic and cranial sites. To assess the performance of the hybrid approach, we simulated treatments using the hybrid phantoms, the scaled and unscaled mesh phantoms, and the ground truth whole-body CTs. We calculated absorbed dose and equivalent dose in and outside of the treatment field, with a focus on neutrons induced in the patient by proton therapy. Proton and neutron dose was calculated using a general purpose Monte Carlo code. **Main results.** The hybrid phantom provided equal or superior accuracy in calculated organ dose and equivalent dose values relative to those obtained using the mesh phantoms in 78% in all selected organs and calculated dose quantities. Comparatively the default mesh and scaled mesh were equal or superior to the other phantoms in 21% and 28% of cases respectively. **Significance.** The proposed methodology for hybrid synthesis provides a tool for whole-body organ dose estimation for individual patients without requiring CT scans of their entire body. Such a capability would be useful for personalized assessment of late effects and risk-optimization of treatment plans.

## 1. Introduction

Radiation therapy is ubiquitous in cancer treatment with approximately 50% of cancer patients receiving some form of radiotherapy (Delaney *et al* 2005). A growing absolute cancer incidence combined with increased survivorship rates has expanded the population of cancer survivors. As an example, more than 16.9 million people have a history of cancer in the United States alone. Furthermore, this number is projected to increase over the next decade to 22.1 million (Miller *et al* 2019, Howlader *et al* 2021). With this increase in the survivor population comes increased interest not only in tumor control, but also in minimization of late-occurring

treatment side effects and improvement of patient quality of life after treatment (Oeffinger *et al* 2006, Newhauser *et al* 2016). Of greatest concern are potentially fatal secondary malignant neoplasms (SMNs), which are new primary cancers that are not a metastasis or recurrence of the original tumor. Including radiation and all other causes, SMNs constituted 19% of all cancer diagnoses in the US between 2005 and 2009 (Morton *et al* 2014). It is estimated that as many as 8% of all SMNs are related to the patient's radiation treatment (Berrington de Gonzales *et al* 2011). From this, and from the reported findings of Dracham *et al* (2018), approximately 1% of radiotherapy patients will develop a radiogenic SMN. This percentage is higher for pediatric patients (estimated at 3%), who have longer lifespans post-treatment compared to adult patients.

Concurrently, there is also increasing interest in proton therapy, especially for pediatric cases (Zhang *et al* 2013, Rombi and Timmermann 2014, Indelicato *et al* 2016), due to the tissue sparing properties of the Bragg peak (Wilson 1946). Reducing the dose to healthy tissue is the primary argument for proton therapy, however, there is a lot of discussion on the stray neutron dose produced in proton therapy which could contribute to an increased secondary cancer risk (Schneider and Hälgl 2015). A study examining a cohort of patients who underwent gamma, x-ray, and electron radiotherapy showed that SMNs most frequently appear on the border of the treatment field, but they can occur very far from the primary tumor site in the low dose far field (Diallo *et al* 2009). Clinical treatment planning systems are frequently incapable of accurately calculating dose outside of the treatment field (Jagetic and Newhauser 2015), of which neutrons are the primary component in proton therapy (Clasie *et al* 2010).

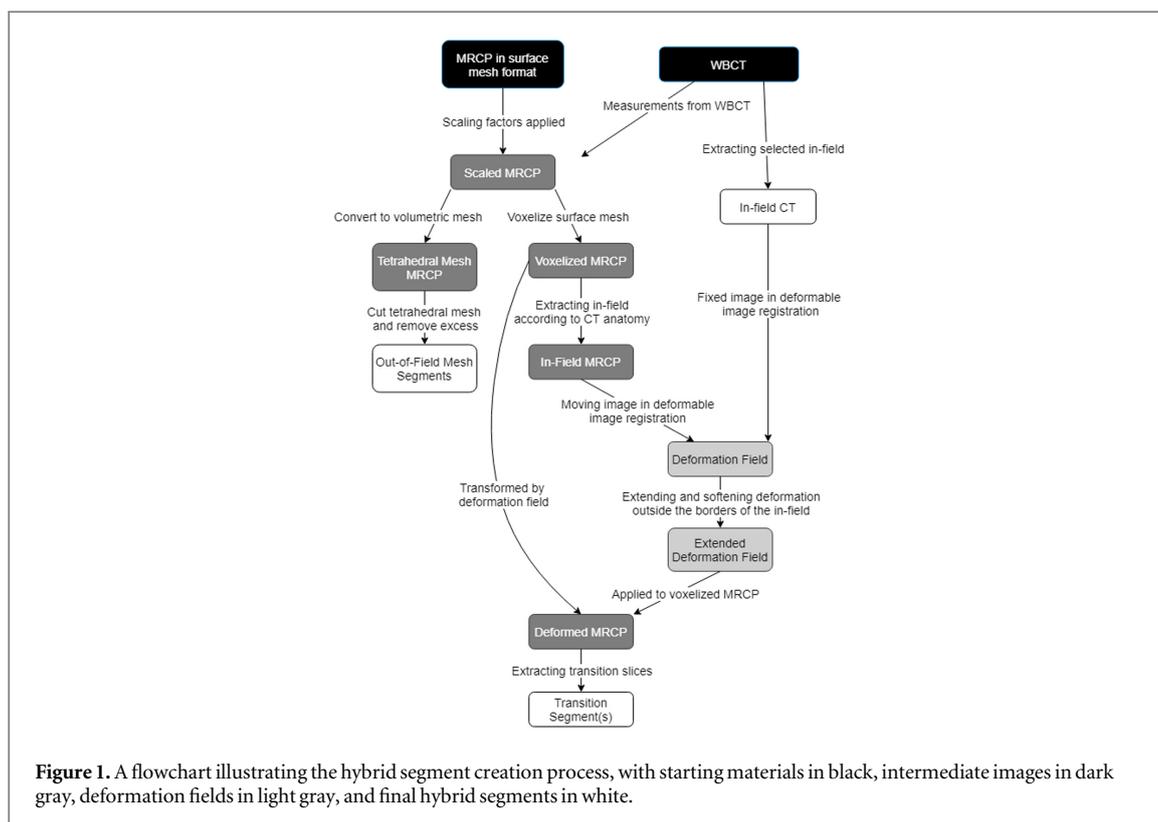
Neutrons are of particular interest in proton therapy, as the high energy protons can produce neutrons in the treatment head of a passive scattering machine (Fontenot *et al* 2008), within potential apertures or range shifters of active scanning machines, and generate them internally within the patient when using either scanning or scattering systems (Newhauser *et al* 2016, AAPM TG 158 Kry *et al* 2017). Some studies have found that even with the additional neutron contribution, photon and proton scattering therapies yielded more or less constant risk, while active scanning protons reduced risk over photon therapies (Schneider and Hälgl 2015) while others have suggested that both proton delivery systems yield much lower risk than IMRT (Newhauser *et al* 2009). However, neutrons dominate the out-of-field dose, and they also have high and variable relative biological effectiveness (RBE) and radiation weighting factors ( $w_R$ ) depending on their energy (ICRP 92 2003, Baiocco *et al* 2016). In recent years, research has revealed the basic physics of stray radiation exposures from proton and photon therapies, including measurements, Monte Carlo simulations, and even analytical models (Jagetic and Newhauser 2015, Newhauser *et al* 2017, Schneider *et al* 2019). A few studies have even demonstrated the feasibility of extending treatment planning studies with analytical models for routine out-of-field dose assessments (Eley *et al* 2015, Wilson *et al* 2020). However, very few studies address the issue of missing anatomical data. Specifically, most radiotherapy is localized (whole-body radiotherapy is rare) and computed tomography (CT) scans for treatment plans are limited to the anatomic region being treated. Thus, for the vast majority of patients, the CT scan contains incomplete anatomic information. One study by Kuzmin *et al* 2018 developed a method of integrating patient CTs into a reference non-uniform b-spline (NURBS) phantom from a library with good results, however this work was only demonstrated in chest-abdomen-pelvis (CAP) CT scans and used simple boundary registration at the CT-phantom edge. Nonetheless, it shows that a whole-body patient-specific CT-reference phantom hybrid is feasible and can provide benefits in Monte Carlo dose estimation.

This work aims to develop and verify a new method to create a bespoke whole-body computational phantom to predict out-of-field dose and equivalent dose for an individual patient. This hybrid phantom uses patient CT images of the treatment site and mesh-type gender-specific reference computational phantoms (MRCP) (Yeom *et al* 2014) which have been scaled to simple patient measurements. This study will additionally utilize deformable image registration to create a blending transition region between the CT and MRCP that matches both internal and external anatomy to form cohesive whole body patient representations that can be implemented into Monte Carlo simulations.

## 2. Methods

The methodology is divided into three portions: hybrid phantom design, Monte Carlo development, and hybrid phantom validation. The key objectives were to:

1. Design a flexible process which combines CT and mesh phantom geometries to form a cohesive whole-body computational phantom.
2. Simulate hypothetical treatment plans on the hybrid, the original MRCP, the MRCP scaled to patient measurements, and the original whole-body CT (WBCT) using a Monte Carlo radiation transport code.



3. Analyze the absorbed dose and equivalent dose delivered to various organs in the phantoms, comparing to the WBCT as the ground truth.

The MRCPs used in this study represent generic adult male and female patients. They were developed by the ICRP Task Group 103 under Committee 2 (Kim *et al* 2016, Kim *et al* 2018) and are modelled after and intended to update the ICRP 110 adult reference phantoms (ICRP 110 2009). For this study it is important to understand that these mesh phantoms come in two formats: triangular polygon surface meshes and tetrahedral volumetric meshes. The surface mesh consists of concentric hollow shells representing anatomical structures. These meshes are easily manipulated in modern 3D modelling software and can be scaled, posed, translated, etc On the other hand the tetrahedral mesh consists of solid volumetric anatomical structures formed by three dimensional tetrahedra. This format is less flexible, but it provides the parameters necessary to construct the geometry in Monte Carlo simulation platforms.

### 2.1. Hybrid phantom design

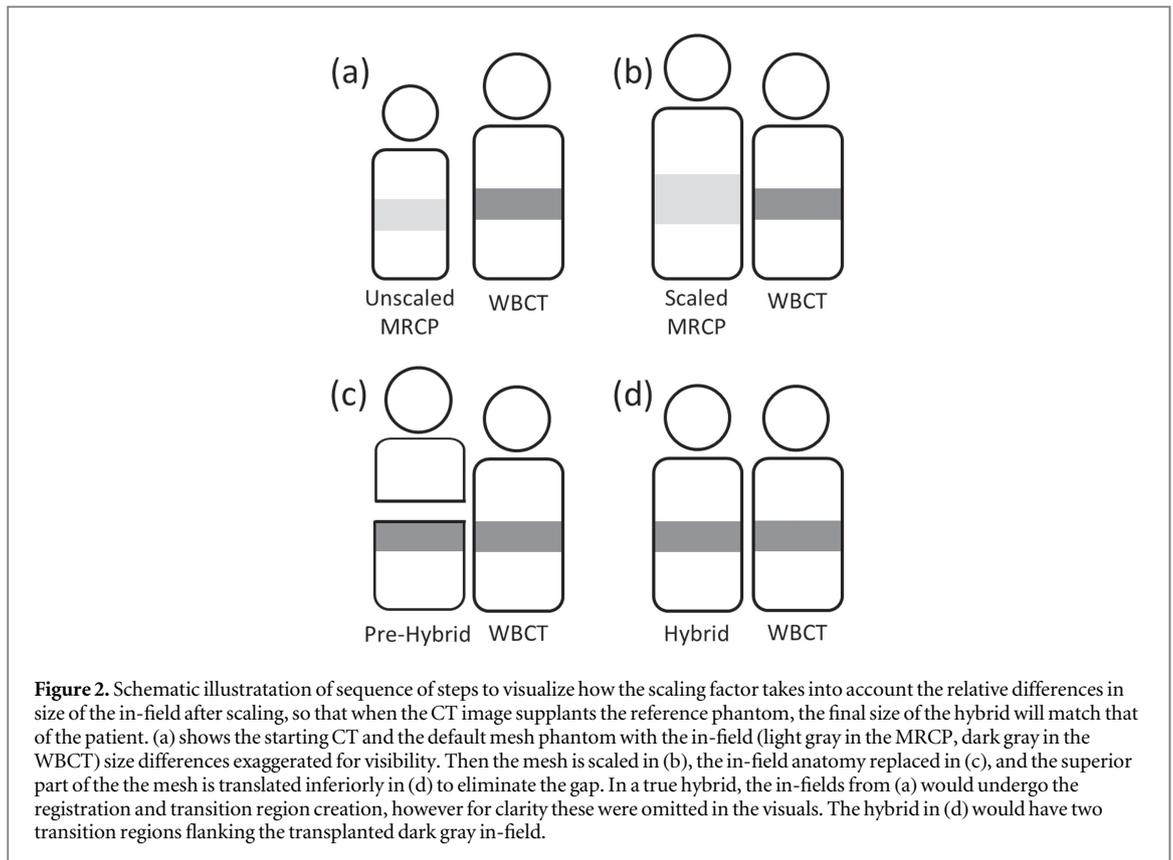
For each hybrid, the phantom is split into three region types with between 3 and 5 separate segments depending on which sides of the in-field are flanked by patient anatomy:

- In-field—Patient CT image corresponding to the treatment site.
- Out-of-field—Tetrahedral mesh regions formed by scaling the MRCP to basic patient measurements (discussed in section 2.1.1), then removing mesh anatomy covered by other segments.
- Transition—Voxelized region formed by first voxelizing the whole body scaled MRCP, then registering the scaled MRCP in-field anatomy to the CT in-field, and finally modifying the deformation field from that registration with an extended softening gradient is used to blend the in-field patient CT into the out-of-field mesh region(s) and applying it back to the whole body scaled MRCP.

A flowchart showing the simplified hybrid creation process from starting materials to final hybrid is shown in figure 1.

#### 2.1.1. Creating the out-of-field segments

The first step in the hybrid creation process was scaling the reference mesh phantom to patient specific measurements. A scaling factor in each dimension (width ( $x$ ), height ( $z$ ), and depth ( $y$ )) was calculated based on



bony anatomy measurements taken from both the unscaled MRCP and the patient WBCT. In clinical practice these measurements might be taken from the patient via physical measurements or surface scanning technology, without the need for a whole-body CT. However, as that information was not known for our ground truth WBCTs, the measurements were taken directly from the patient scans to adequately represent what physical measurements would be available. If no physical measurements or size data is present for a historical case, the methodology would need to be adapted to either not use a scaling factor or to make an educated guess about these physical measurements based on any available patient data.

The measurement points in the MRCP and the WBCT were defined based on the methods of Zvereva *et al* (2017) and adapted to use measurements taken directly from CT anatomy to mimic physical patient measurements which were not available. The measurements taken in each the MRCP and the WBCT were: the most anterior and posterior points of the pelvis (depth), the left and rightmost points of the pelvis (width), the most inferior and superior points of the spine (height), and the most inferior and superior points of the in-field anatomy (height scaling correction). For the height scaling correction, the CT in-field as well as the segment of the MRCP with the closest anatomical fit to the corresponding patient CT in-field was identified and measured in height.

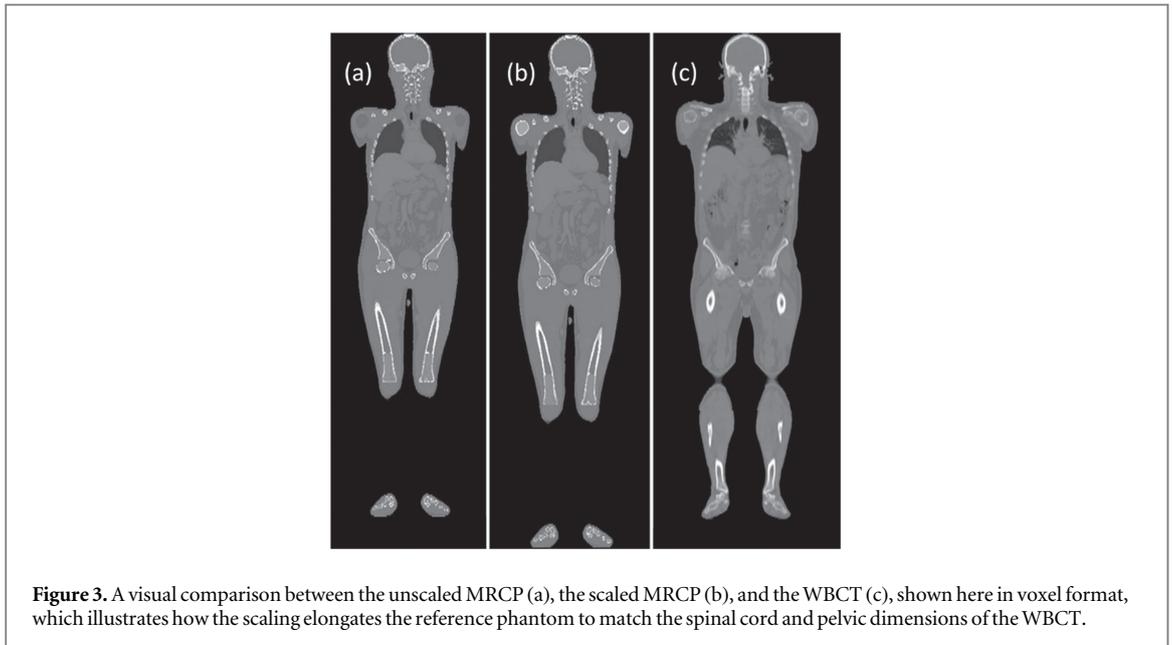
The addition of the in-field measurements is required specifically for adjusting the scaling factor in height. Due to the relative simplicity of using a single scaling factor for each dimension, it is possible that the scaled MRCP in-field and the WBCT in-field will have different sizes because of inherent proportional differences in anatomy. If this aspect is not considered, then the final hybrid may be shorter or taller in the torso than anticipated, which effects its ability to predict dose in out-of-field organs. A simple diagram illustrating this concept is shown in figure 2.

Once the measurements were taken, the scaling factors for width and depth were calculated in the following way:

$$SF = \frac{WBCT_{max} - WBCT_{min}}{MRCP_{max} - MRCP_{min}}, \quad (2.1)$$

where SF is the scaling factor,  $WBCT_{max/min}$  are the terminal measurements (left/right and interior/posterior points of the pelvis) taken in the WBCT, and  $MRCP_{max/min}$  are the same measurements in the MRCP.

The scaling factor for the height was calculated similarly, but with an adjustment to account for the size-change of the in-field after scaling.



**Figure 3.** A visual comparison between the unscaled MRCP (a), the scaled MRCP (b), and the WBCT (c), shown here in voxel format, which illustrates how the scaling elongates the reference phantom to match the spinal cord and pelvic dimensions of the WBCT.

$$SF = \frac{WBCT_{spine\ height} - WBCT_{in-field\ height}}{MRCP_{spine\ height} - MRCP_{in-field\ height}}, \quad (2.2)$$

where  $WBCT_{spine\ height}$  is the difference in the minimum and maximum measurement of the spine in the WBCT,  $WBCT_{in-field\ height}$  is the difference in the minimum and maximum measurement of the designated anatomical in-field of the WBCT, and the denominator are the corresponding measurements in the MRCP. These scaling factors were then applied to the MRCP in its triangular polygon surface mesh data format in the Blender 3D modelling software, exported as a surface mesh from Blender, and converted into both voxel format and a tetrahedral mesh format. The voxelized mesh was then used to create the transition region and the tetrahedral mesh format was used for simulating in Geant4 Monte Carlo geometries. Figure 3 shows how the application of this custom scaling factor visually effects the male MRCP in comparison to the patient WBCT. In particular, this figure reveals how the scaling factor in height is designed to match the torso dimensions, not the total height of the patient. We chose to scale to the height of the spine rather than the total height as leg length does not indicate any relevant information about the spatial properties of any organs of interest, and so its inclusion via use of total height might result in a poor scaling factor match. However, the spine is indicative of the height of the torso, which influences the size and placement of the organs contained within. In a retrospective case, where only total height is available, that information alone could be used for scaling, however it would likely result in a less accurate out-of-field anatomy than one which scales to the torso.

After scaling and exporting from Blender, the scaled MRCP was converted to a tetrahedral mesh prior to use in simulation since using a polygon geometry severely impacts the computational speed in Geant4. A code, utilizing Tetgen (Si 2015), initially developed by the ICRP Committee 2 team responsible for the mesh phantoms was then modified to match our implementation and used to tetrahedralize the scaled MRCP (Han *et al* 2020).

After the scaled MRCP was converted to a tetrahedral mesh, it was cut at each edge of the transition region and then the unneeded portion of the mesh was removed. For this purpose, we created an in-house code which uses the mathematical algorithm from Wang *et al* (2014). Specifically, the code cuts a tetrahedral structure along the intersecting edge of a surface mesh and re-calculates the tetrahedra at that edge to maintain the tetrahedral mesh structure. This algorithm was expanded to iterate through all mesh structures, include material information, remove the unwanted volume, and write the cut mesh in the tetrahedral data format.

As another note, as discussed previously when describing the scaling factors from equations (2.1) and (2.2), it is possible that the scaled MRCP in-field and the CT in-field might be of different heights. If this was the case, the mesh phantom was vertically translated such that, once cut, the anatomy was once again continuous with the supplanting CT in-field and the flanking transition region(s). The necessity for this out-of-field segment translation can be seen from figure 2, as shown previously.

### 2.1.2. Preparing the in-field CT

To facilitate the deformation used to create the transition region, the WBCT was first manually rigidly translated to align the WBCT and scaled MRCP in-field anatomy as closely as possible. Once aligned, the selected in-field image slices were extracted from the aligned WBCT. This in-field CT is used both as the fixed image in the

deformable registration needed for the transition region, and in the final hybrid as the in-field CT segment. In a typical clinical workflow, the in-field CT would be the only available image, and so this step would not be necessary. However, since this study utilized the WBCTs and their subsequent dose distributions as a ground truth, the in-field extraction is a required step.

### 2.1.3. Creating the transition region

The first step in creating the transition region was to voxelize the scaled MRCP. The conversion to a voxel format was done with an in-house Matlab code capable of converting the surface mesh into a voxelized array and assigning HU values to each voxel based on the mesh phantom organ material definitions and the same density to HU conversion curve as a research version of RayStation 7, which was the treatment planning system (TPS) used in this study. The voxel resolution was derived based on the resolution of the original ICRP Publication 110 phantoms (1.08 mm × 1.08 mm × 8 mm), but was adjusted from 8 mm slice thickness to 5 mm slice thickness in order to (1) reduce the impact of partial volume effects (2) minimize the computer memory required for finer resolutions and (3) match the superior–inferior (*Z*) dimension of the WBCT voxels for ease of assimilating the transition regions to the abutting regions of the hybrid phantom.

Once the voxelized scaled MRCP was completed, it was converted into a DICOM image format and the section which matches the anatomy of the CT in-field is excised. To facilitate a better registration, the in-field CT was resampled in the superior–inferior direction to match that of the MCRP. The in-field CT was resampled and not the scaled MRCP since the final deformation field is applied to the voxelized scaled MRCP, so keeping the deformation field at the native dimensions of the scaled MRCP avoided undue distortion caused by resizing the deformation field.

The scaled MRCP in-field was then deformably registered to the CT in-field using the Morphons algorithm (Knutsson and Andersson 2005), (Plumat *et al* 2009) using the openREGGUI Matlab extension (<https://openreggui.org/>). The subsequently created deformation field was then exported from Reggui as a 4D Matlab array. An in-house code was created which duplicates the terminal borders of the deformation field in the *Z* direction, smoothly reduces the magnitude of the deformation using a sigmoid function over the pre-determined distance of the transition region (10 cm in this study) and appends the extended borders back onto the original deformation field.

This modified deformation field was then applied to the whole-body voxelized scaled MRCP. The MRCP was exported as a DICOM image format and the image slices corresponding to the transition region were extracted. Since the modified deformation field is applied directly to the whole-body scaled MRCP, there is a potential for the two transition regions, while continuous within the scaled MRCP, to have a gap or overlap when the CT in-field is transplanted into the hybrid (such a case was described in section 2.1.1 and shown in figure 2). If this was the case for a given hybrid, then the transition slices were translated in the *Z* direction to align with the out-of-field segments and CT in-field.

## 2.2. Monte Carlo based dose calculations

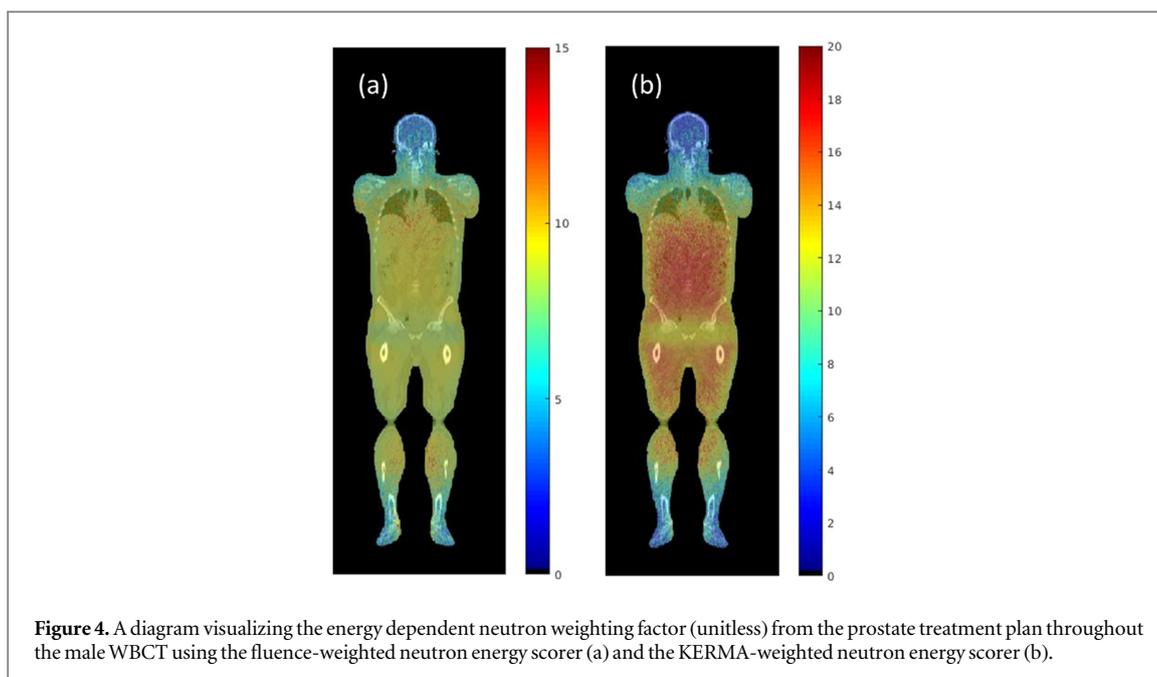
This study uses the Geant4 Monte Carlo platform (Agostinelli *et al* 2003). The base implementation was derived directly from the work of Schmid *et al* (2015) and updated to be compatible with version 10.05.p01. The original code was built to only process DICOM images using the HU-based materials list from Schneider *et al* (2000). The Monte Carlo physics models were configured as the default predefined QGSP\_BIC\_HP Geant4 physics list and were not altered during this study.

### 2.2.1. Implementing mesh and hybrid geometry

The tetrahedral mesh phantom and all tetrahedral mesh segments were built using code heavily derived from the methods developed by the ICRP Committee 2 team (Kim *et al* 2016). A combined list of materials was created in Geant4 which allows the tetrahedral mesh phantom segments to use their native tissue-specific material definitions while the voxel regions use the RayStation 7 CT conversion table to convert to the Schneider *et al* (2000) material definitions with an additional fine-tuning based on density using the methods of Schmid *et al* (2015).

### 2.2.2. Designing custom scorers

Equivalent dose is calculated by multiplying the absorbed dose by a radiation type by a radiation weighting factor specific to that type. To facilitate this calculation, custom scorers were required to separate the therapeutic proton contribution from the neutron contribution and to calculate the neutron radiation weighting factor following the recommendations of ICRP 92 (2003) as closely as possible. For this study, neutron dose was considered as all dose which at some point in its history originated from a neutron. This decision was made to mimic the energy dependent neutron weighting factor assumptions, which was designed to account for all the



potential dose depositions from an incident neutron field of a given energy. The neutron dose scorer assigned a tag to any neutron as it was created, and then propagated that tag as it produced secondaries and further. All the dose deposited by tagged particles were then summed to get the total neutron dose component. Following a similar justification regarding the ICRP 92 conditions required to use the proton weighting factor, the therapeutic proton dose was considered as all dose scored even if deposited by a different particle, except that which is already counted as neutron dose. It was calculated by scoring the total absorbed dose and then subtracting out the neutron component.

Two methods of scoring neutron energy were also created. The first weights the neutron energy by the relative fluence through the scored voxel, and the second weights by the relative amount of kinetic energy transferred from neutrons to charged particles in that voxel (KERMA). Due to how Geant4 tallies scored quantities, to get the average neutron energy per voxel weighted by either KERMA or fluence, two separate scorers were needed for each weighting. For example, to calculate the KERMA weighted neutron energy, one scorer was needed to tally the total energy transferred to charged particles by neutrons in the voxel, and the other was needed to score and sum in a voxel the energy of a given neutron multiplied by the specific amount of KERMA due to that neutron. Both the KERMA- and the fluence-weighted neutron energy were calculated with the intent to evaluate the impact of scoring choice on equivalent dose and ultimately risk prediction.

The scorers were applied as a voxel grid which covered the full extent of the simulated geometry. If voxelized geometry was present, the scoring grid was aligned with that pre-established voxel geometry grid. Although the scoring was done as a voxel grid, all geometries which utilized mesh segments built those segments as a tetrahedral mesh in the simulation geometry.

Figure 4 shows a comparison between the neutron weighting factor using the fluence-weighted and KERMA-weighted neutron energy scorer across the whole-body of the male patient in the pelvic treatment plan. Overall, compared to the fluence-weighting factor, a much larger portion of voxels were attributed a higher weighting factor ( $>15$ ) when using the KERMA-weighted neutron energy scorer. Both scorers exhibit a dip in neutron weighting factor where the primary treatment beams were located, and both scorers reach the minimum range of the neutron weighting factor (2.5 using the ICRP 92 model) at the most extreme ends of the patient.

### 2.2.3. Calibrating stopping power

To ensure consistent transport calculations from the TPS and Monte Carlo code, the proton stopping powers in the Monte Carlo code were calibrated to match those of the TPS. A series of box phantoms were created, each of which consisted of a singular HU value which together encompassed the range of all possible HU-based materials in Geant4. A single pencil beam was created in the TPS and simulated in Geant4 to be incident on the box, repeated at three energies which spanned the available energies on the chosen treatment machine. The ranges calculated by the TPS and simulated in Geant4 were compared at those three energies to ensure that the stopping power calibration was not energy dependent. Based on a comparison of the proton ranges (defined by the distal R80, the point at which dose is 80% of the maximum in the fall-off of the Bragg peak dose profile)

between the TPS and Geant4 in those box phantoms, the ionization energy of each of the HU-based materials were adjusted such that the R80s of the pencil beam in Geant4 matched that of the TPS. Across the entire range of material HUs, the average ionization energy adjustment was 8%, and for materials less than 1400 HU it was 4%. After the calibration, the average difference in R80 between the TPS and Geant4 was 0.04%.

### 2.3. Hybrid validation

With the hybrid methodology established, and the proper computational tools in place, the next stage was to measure and quantify the hybrid predictive abilities compared to the other phantoms with respect to the ground truth of the patient WBCT. To accomplish this, a series of treatment plans were created for selected in-fields and clinical indications. These treatment plans were then simulated on their respective WBCT as well as the corresponding hybrid phantom, scaled MRCP, and unscaled MRCP. The absorbed dose and equivalent dose across the whole-body were then calculated from the scored quantities, isolated per organ of interest, and compared.

#### 2.3.1. Hybrid selection

A selection of seven WBCTs were provided by the University Hospital of LMU Munich for use in the hybrid studies. These WBCTs do not have any solid tumors and were taken for the purpose of whole-body irradiation. From this group, the largest male and the smallest female were selected to represent the hypothetical patient from which we would derive the hybrids. These two specific WBCTs were chosen with the intent to push the scaling as far as possible to test the ability of the hybrid methodology to adapt to most extreme patient dimensions available. The heights of the patients were not given, and exact determination of total height from the WBCT was difficult due to the imaging couch lifting significantly beneath the knees. With the bent knees the male WBCT had a height of 176 cm and the female WBCT had a height of 149 cm. However, this height discrepancy in the legs did not affect the scaling efficacy due to the nature of the scaling factor.

Since the acquired WBCTs do not have any solid tumors, there was a relative freedom in choosing meaningful hypothetical treatment sites. The two anatomical locations chosen to test the hybrid were the head and the pelvis. These two sites were chosen to maximize the potential distance between the in-field and the farthest organ of interest. The specific oncological treatment sites chosen were: prostate cancer for the male pelvis, cervical cancer for the female pelvis, nasopharyngeal cancer for the male head, and meningioma for the female head. The indications were chosen based on a variety of factors including: ease of contouring, relevance to proton therapy (or external beam therapy in general), detailed examples in the literature (for referencing dose objectives, constraints, or beam setups), and availability of full treatment plans to reference.

#### 2.3.2. Treatment planning

Treatment planning was done in a research version of RayStation 7 using a pencil-beam scanning proton machine. For each treatment site the treatment plan was designed solely on the in-field CT to mimic the conditions in the clinic. In addition to the in-field CT alone, the WBCT and the in-field with transition slices were also imported into the TPS for out-of-field organ contouring for use in post-simulation dose analysis. All organ at risk (OAR) contouring was done according to atlases from the Radiation Therapy Oncology Group (RTOG) and NRG Oncology (Center for Innovation in Radiation Oncology 2019). The in-field had all structures contoured necessary for treatment planning. As a note, not all possible anatomical structures are represented by the mesh (for example there are no discrete structures for the brainstem, lobes of the brain, optic chiasm, etc). Because of this, besides the structures needed for treatment planning objectives, an additional set of organs which had complete mesh structures in the MRCP were contoured with the express purpose of post-simulation dose analysis of the hybrid, WBCT, and mesh phantoms.

Gross tumor volumes (GTVs) were contoured either based on some case studies available in the literature (Toita *et al* 2011, Romano *et al* 2016, Brouwer *et al* 2015) or based on the volumes observed in some anonymized treatment plans from the Proton Therapy Center in Trento, Italy. From there, appropriate margins for the clinical target volume (CTV) and planning target volume (PTV) were used again from information available in the literature and the Trento Proton Therapy Center's plans for each tumor type.

Similarly, the beam setups, fractionation, clinical objectives, and dose limits were either directly taken from the literature (Hashimoto *et al* 2016, Sasidharan *et al* 2019, Brodin and Tomé 2018) or from the Trento Proton Therapy Center's treatment plans. After planning was complete, the treatment plans were checked for clinical reasonability by a clinical medical physicist with experience in proton therapy. Some figures of the treatment plans used in this study can be found in Figures SM. 2–5 in the supplementary materials (available online at [stacks.iop.org/PMB/67/035005/mmedia](https://stacks.iop.org/PMB/67/035005/mmedia)).

### 2.3.3. Monte Carlo simulation

To simulate the treatment plan in Geant4, an in-house Python script was developed to convert the beam information from RayStation to a text file which could be executed in Geant4 as part of a user command file. The script converts the spot coordinates in beams-eye view from the TPS into a pencil beam particle origin coordinate and direction using gantry angle, couch angle, and isocenter coordinates. It also draws information regarding spot weight, beam shape, and energy spread directly from the TPS.

Once the treatment plan was designed and exported, the plan was simulated on the WBCT, the hybrid, the scaled MRCP, and the unscaled MRCP. The treatment plan was already naturally aligned in the hybrid and WBCT, but in the case of the mesh phantoms, where the anatomy and positioning are different, the treatment plan was translated to be as closely aligned as possible by using manual rigid translation using a distinct bony feature in the central region of the treatment field in both the CT and mesh phantom anatomies.

One million primary particles were simulated in each simulation job submitted in a computing cluster and was repeated with different random seeds for 500 simulations for a total of 500 million particles. These simulations were split and combined into 10 groups of 50 simulations each. The simulated particles within each group yields high statistics in the in-field (<1 mGy uncertainty for neutron dose <0.05 Gy or Sv for all other scorers). The 10 sampling groups enables a robust assessment of uncertainty by calculating the 25th and 75th percentiles of the simulated organ dose quantities. No optimization of the number simulated particles was performed to achieve specific statistical uncertainty in the out-of-field organs.

### 2.3.4. Organ dose analysis

In each simulation, total absorbed dose, neutron dose, and the four scorers for fluence and KERMA weighted neutron energy were tallied across the whole-body. To isolate the scored quantities for an individual organ, masks were constructed from individually voxelizing organs from the meshes to make binary arrays, imported contours from RayStation, or a combination of the two (for organs which extend into both the mesh and voxelized regions). To calculate the total and neutron dose to an organ, the organ mask was applied to the scoring array such that any voxel not containing the organ would not be considered. The dose in an organ was calculated as the mean dose from all voxels in the mask.

The subsequent weighted whole-body neutron energy arrays were used as an input into the ICRP 92 (2003) analytical model (reproduced here in equation (2.3)) for the energy-dependent neutron weighting factor

$$w_R = 2.5 \left[ 2 - e^{-4E} + 6e^{\frac{-\ln(E)^2}{4}} + e^{\frac{-\ln(\frac{E}{30})^2}{2}} \right]. \quad (2.3)$$

By the end of this process there was a whole-body voxel array of the neutron weighting factor which should be attributed to each voxel. This array was then multiplied voxel-wise by the neutron dose array to get the neutron equivalent dose to the whole-body. The proton equivalent dose was calculated by subtracting out the neutron absorbed dose from the total absorbed dose and then multiplying with the proton weighting factor of 2 recommended by the ICRP 92 (2003). Total equivalent dose for a given organ was calculated by adding the two components together and applying the relevant organ mask as described previously.

For a given organ in a given treatment site in a given patient representation, 10 values of each quantity were calculated corresponding to each of the 10 batches. These 10 samples were then plotted in a boxplot and compared against the same quantity for the same organ in the other patient representations (WBCT, hybrid, scaled and unscaled MRCP). The whiskers of the boxplot extend the full range of the samples excluding outliers, and the outer edges of the box represent the 25th and 75th percentile. Outliers are defined as points that are more than 1.5 times the interquartile range away from the 25th or 75th quartiles.

### 2.3.5. Quantification of performance

Hybrid performance was assessed for organs separately based on position relative to the in-field CT. In and near field organs were defined as those organs which were contained at least partially in the CT in-field region. For the male pelvis, these organs were the prostate, the bladder, and the colon. For the female pelvis they were the bladder, the colon, and the kidneys. For both the male and female head and neck treatment plans, the in and near field organs were defined as the brain, left and right eyes, and the thyroid. For a given organ in a given treatment site, each dose metric counted as a separate 'case' when calculating the fraction of organs which performed at a certain level. The number of in and near field cases for each patient representation is 56 (4 dose quantities, 3 organs for the male pelvis treatment plan, 3 for the female pelvis, and 4 each for the head and neck plans) and a similar calculation for the out-of-field organs yields a total of 48 cases.

For each of the cases, the phantom predictions were compared to the ground truth WBCT. Based on their predictions, each phantom's case was categorized as 'closest', 'intermediate', or 'farthest'. If two phantoms yielded estimates where the 25th and 75th percentile uncertainties overlapped, those phantoms were considered



**Figure 5.** Coronal and sagittal visual representation of the male pelvis hybrid (a), (b), the WBCT (c), (d), and the male head and neck hybrid (e), (f). Each hybrid is aligned with the corresponding anatomy in the WBCT so the relative out-of-field anatomy is easily compared. The mesh segments have been voxelized, and the whole-body have been cropped for visualization purposes. The horizontal lines in the hybrid phantoms delineate the different segment types.

to have performed equally. If two phantoms performed equally well over the third, the two would both be categorized as ‘Closest’ while the third would be categorized as ‘Farthest’ and no ‘Intermediate’ level would be assigned for that case. By analyzing the relative fractions of cases assigned to the three categories for a given phantom type (hybrid, scaled MRCP, default MRCP), we can assess how often the relative performance of each phantom across scorer, patient gender, and treatment site.

### 3. Results

This section presents the results of the hybrid creation process and the quantitative organ absorbed dose and equivalent dose analysis comparing the hybrid to the mesh phantoms and the WBCT.

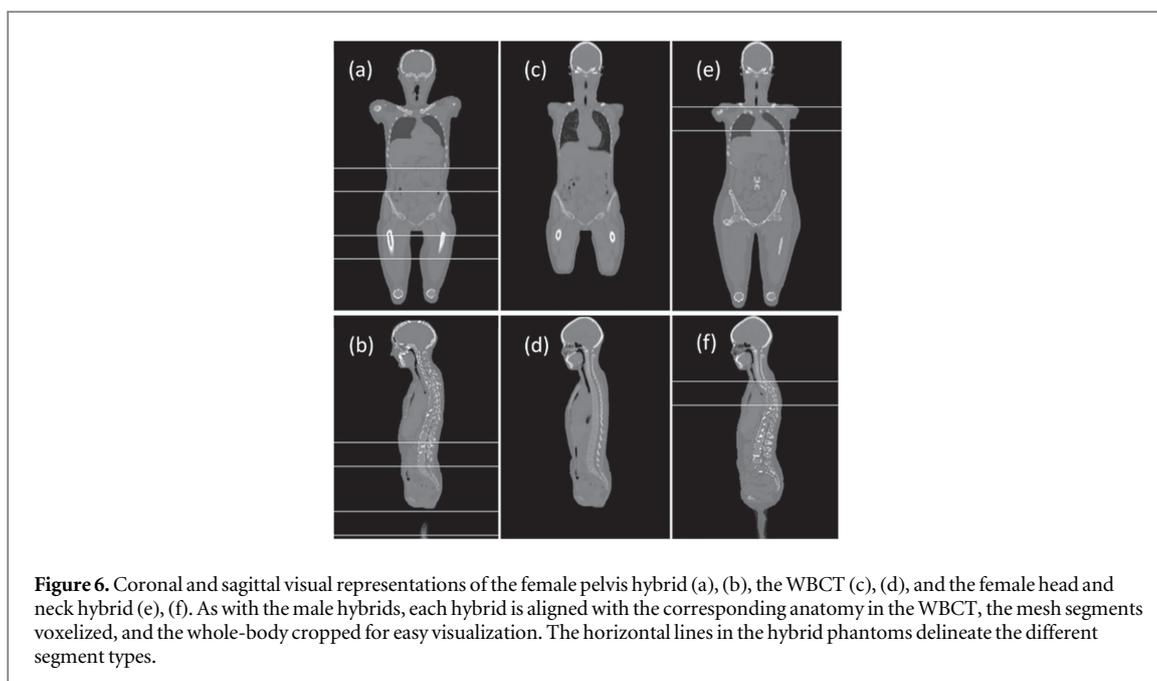
#### 3.1. Hybrid creation results

Figures 5 and 6 show the coronal and sagittal cross sections for all four hybrid phantoms created for this study compared against the original patient WBCT. Overall the hybrid creation process was able to create reasonably continuous phantoms. Other than some bony anatomy distortion in the transition region and some discontinuities at the skin-air border, the internal structures and relative HU values blended together well. There were some inherent positioning differences between the WBCTs (lying flat on a table gripping handholds with legs propped up at the knees) and the mesh phantoms (free-standing), and because of this there are some major distortions and discontinuities in the arms and hands when those regions lay in the transition. Despite this, those affected regions did not contain any vulnerable organs of interest for this study, and so these discontinuities were considered acceptable.

In addition to these considerations, in some instances at first glance a hybrid might appear to have some significant height differences when compared to the WBCT (for instance in figure 4 in sub figures (a) and (c)). However, the vertical scaling factor does not take total height into account, and its function is to match the spine height of the hybrid to the WBCT. By observing the positions of the inferior and superior points of the spine in figure 4, it can be seen that the scaling factor has indeed fulfilled its purpose and the spinal columns of the hybrid and WBCT are the same height.

#### 3.2. Hybrid performance results

Due to the large amount of raw data from the four dose quantities, four treatment sites, four patient representations, and six or seven organs per site, this section will present a summary of the trends observed in phantom performance. When necessary, this summary will be supported with specific and pertinent examples. The full extent of the results is given in the supplementary materials tables 1–4 for quantitative data and figures SM. 6–21 for comparative boxplots.



**Table 1.** Fraction of organs by dose quantity prediction performance for all phantoms across all treatment sites.

Organ location	Rank of prediction with respect to WBCT ground truth	Hybrid	Scaled MRCP	MRCP
In/near-field	Closest	0.98	0.09	0.04
	Intermediate	0.02	0.61	0.30
	Farthest	0.00	0.30	0.66
Out-of-field	Closest	0.56	0.54	0.44
	Intermediate	0.42	0.10	0.17
	Farthest	0.02	0.35	0.40

Table 1 shows the performance of the phantoms as a fraction of organs which gave the closest predicted dose quantity value as the WBCT. For the in-field, the hybrid provides the closest estimate to the WBCT by far. The only case where it achieved an intermediate rank over a closest rank was for the neutron dose in the colon for the male pelvis case. The scaled MRCP typically provided the next best estimate in the in-and near-field, followed by the default MRCP.

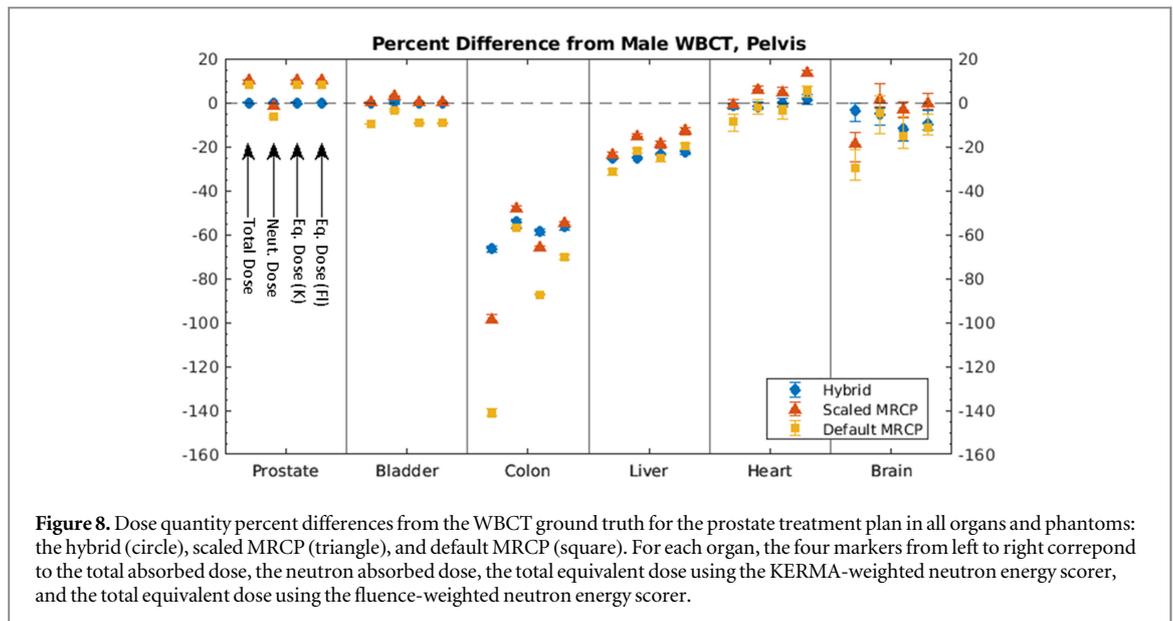
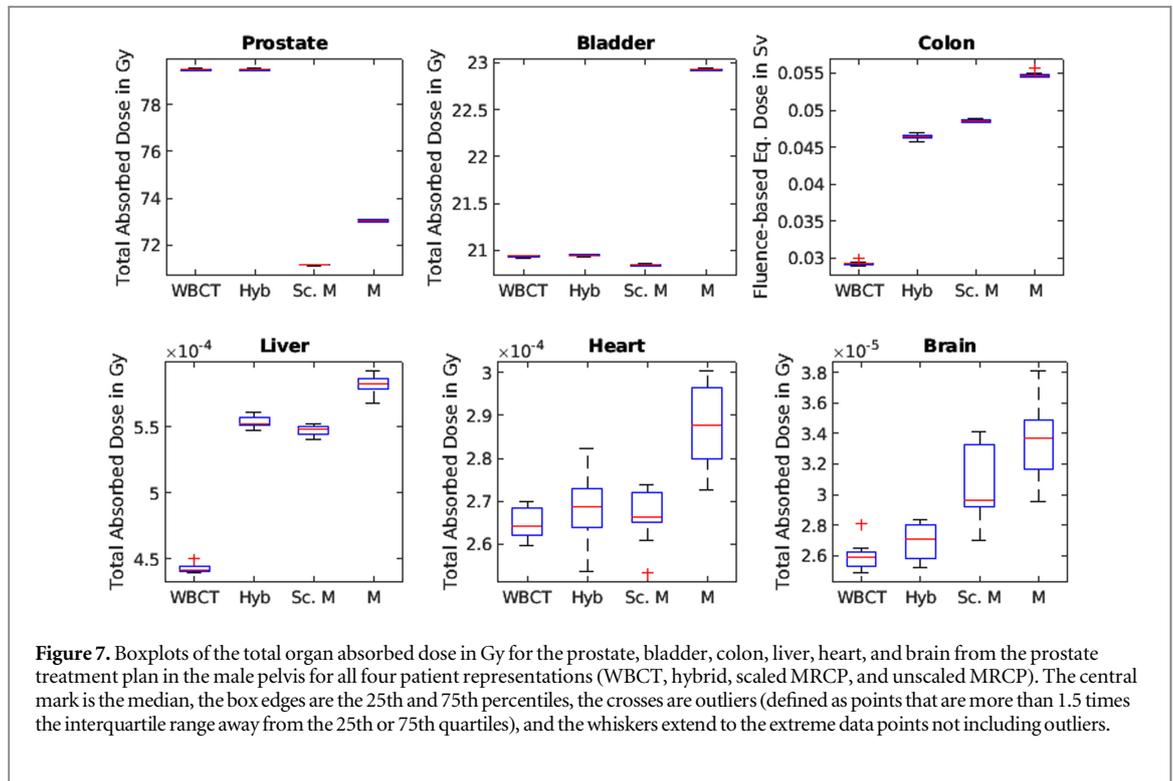
The out-of-field cases were a little more varied. The hybrid was by far the least likely to provide the farthest estimate of dose or equivalent dose, where there was only a single case: the liver in the male pelvis plan scoring neutron dose. That said, while the hybrid was outside of the 25th–75th percentiles defined as our uncertainties, in this case the hybrid was very close to matching the intermediate phantom with a 0.467 mGy/25% difference in the hybrid versus 0.455 mGy/22% in the intermediate default MRCP.

In the out-of-field the hybrid very frequently shared the closest rank with one or both of the other phantoms, reflecting the higher uncertainties and the shared/similar anatomy. For each of the phantoms, the most common ranking for cases in the out-of-field was closest, although the hybrid did have the most closest ranks over the scaled MRCP by one case. Unlike in the hybrid, for both the scaled MRCP and the default MRCP the next most common ranking after closest was farthest.

The scaled MRCP yielded the farthest estimate of the dose quantity in 35% of cases, and for the default MRCP it further increased to 40% of cases. In stark contrast, the hybrid had only a single case (2%) where it yielded the farthest estimate. Essentially, while all phantoms had more or less a similar effectiveness in providing the closest estimate of the ground truth scored dose quantities for the chosen organs in the out-of-field, if a phantom did not yield the closest estimate, the hybrid was by far the most likely to yield the next closest.

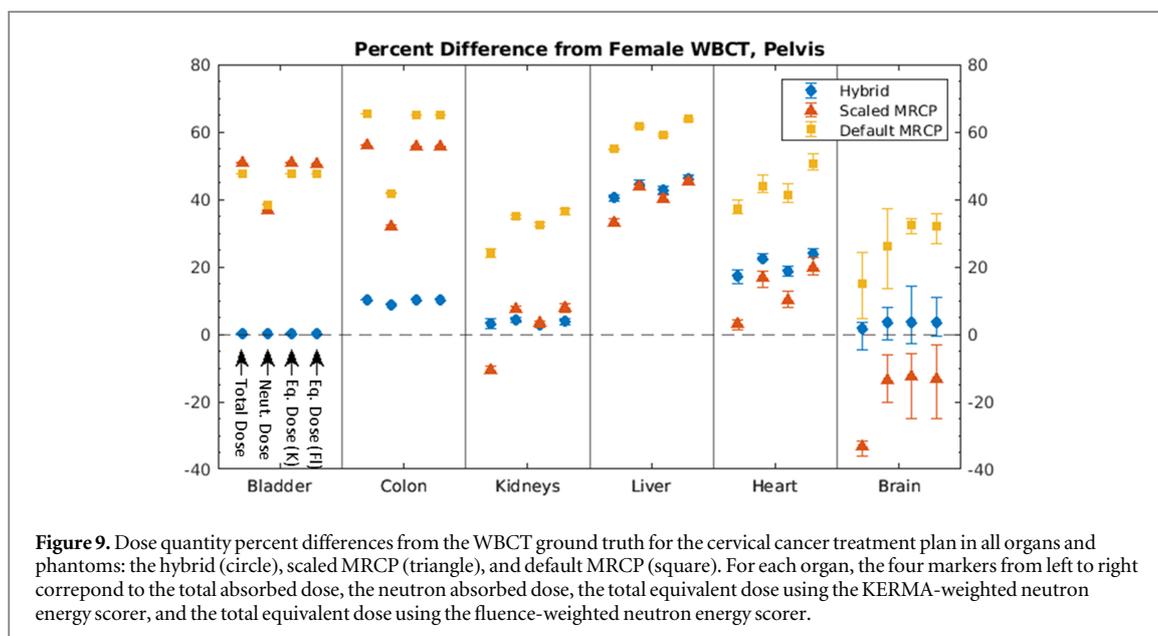
For each treatment site, boxplot figures were made showing the calculated value for each dose quantity for all examined organs. The boxplot for the total dose in the male pelvis treatment plan is given in figure 7. From this figure the trend described before can be clearly seen.

In and near the treatment field the hybrid provides the best estimate of total dose. In the out-of-field the uncertainties grow, and in the case of the heart, both the hybrid and the scaled MRCP provide equivalent predictions and would be classified together as being the closest phantoms to the WBCT ground truth.

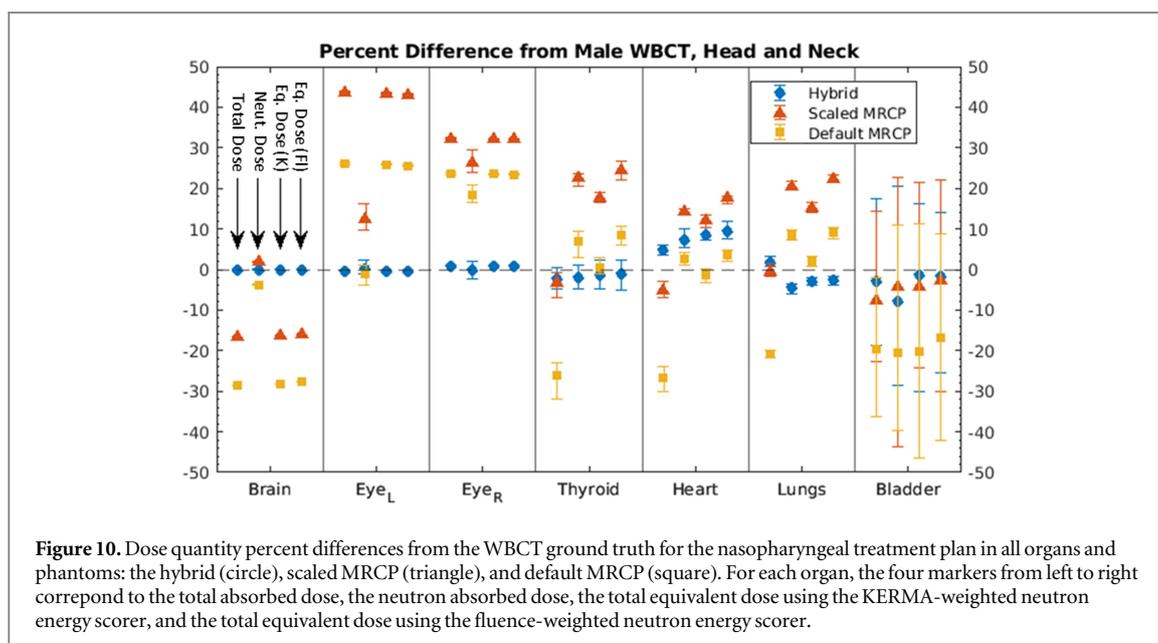


Furthermore in the liver it can be seen that the scaled MRCP actually provides a closer estimate to the ground truth than the hybrid. The boxplots for the other dose quantities for the male pelvis treatment plan as well as those for the other treatment sites are given in the supplementary materials figures SM. 6-21. For full quantitative information for all treatment sites and all dose quantities, see the supplementary materials tables 1–4.

Looking at the individual treatment sites, figures 8–11 shows the percent difference from the WBCT for all dose quantities and organs for a given treatment sites. In this way not only is the relative performance of the phantoms clearly visible, but also shows some information on the magnitude of difference from the WBCT and from each other. One thing which is clear through these figures is that while the hybrid is most frequently the best predictor of the chosen dose quantities, there are cases (particularly in the out-of-field) where the scaled MRCP or even the default MRCP outperforms the hybrid, similar to the liver case regarding the figure 7 boxplots.



**Figure 9.** Dose quantity percent differences from the WBCT ground truth for the cervical cancer treatment plan in all organs and phantoms: the hybrid (circle), scaled MRCP (triangle), and default MRCP (square). For each organ, the four markers from left to right correspond to the total absorbed dose, the neutron absorbed dose, the total equivalent dose using the KERMA-weighted neutron energy scorer, and the total equivalent dose using the fluence-weighted neutron energy scorer.

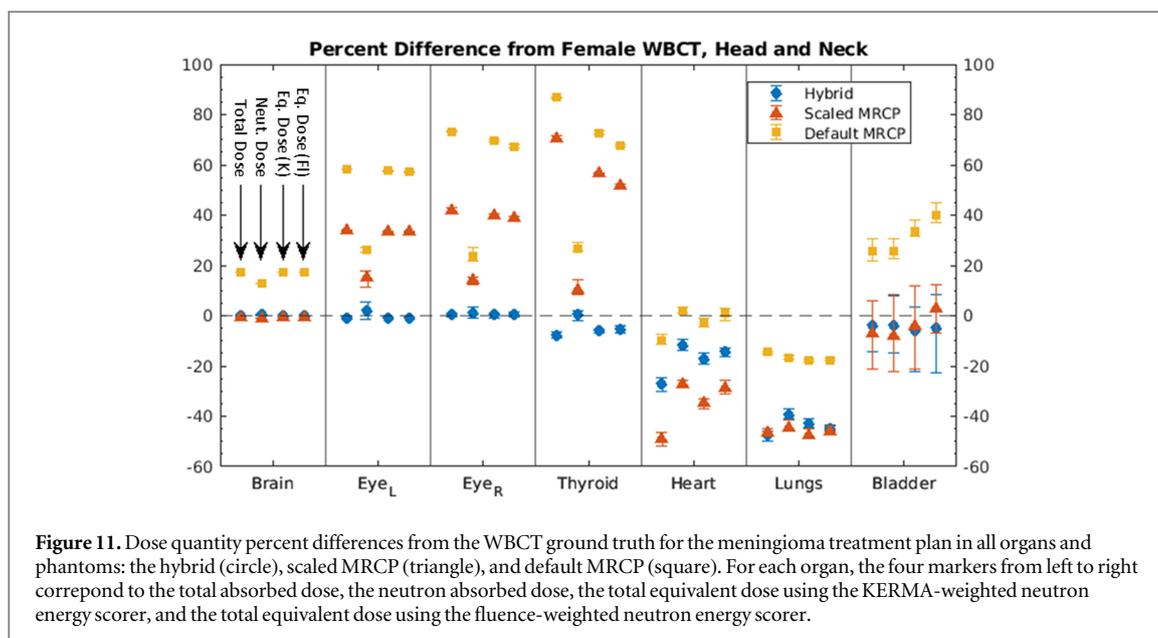


**Figure 10.** Dose quantity percent differences from the WBCT ground truth for the nasopharyngeal treatment plan in all organs and phantoms: the hybrid (circle), scaled MRCP (triangle), and default MRCP (square). For each organ, the four markers from left to right correspond to the total absorbed dose, the neutron absorbed dose, the total equivalent dose using the KERMA-weighted neutron energy scorer, and the total equivalent dose using the fluence-weighted neutron energy scorer.

Across all the treatment sites, if an organ was in or near the treatment field, then the equivalent dose tended to match the total dose in terms of the percent difference from the WBCT (visible for example in the MRCPs' performance in the brain and eyes in figures 10 and 11). This is likely due to the dominance of the therapeutic dose in that region, with a relatively small impact of the neutron dose. Conversely, in the out-of-field, the equivalent dose varies a lot more. A little fluctuation is present between the two weighting methods, but overall the equivalent dose tends to be closer to the percent difference of the neutron dose scorer than the total dose scorer (clearly visible in the brain in figure 9 and the heart and lungs in figure 10).

Another important thing to note, that while the percent differences in the out-of-field can be quite high (frequently between 20% and 50% difference from the WBCT) the distance from the treatment field can mean the absolute difference in dose or equivalent dose is not significant. For example, in the extreme distance case of the bladder in the male head and neck treatment plan, both the hybrid and the scaled MRCP had a prediction of less than 5% difference from the ground truth for fluence-weighted equivalent dose while the default MRCP had a 20% difference. However, in terms of absolute numbers the difference is quite small, with the hybrid, scaled MRCP, and default MRCP yielded equivalent dose estimates of  $99.3 \mu\text{Sv}$ ,  $102 \mu\text{Sv}$ , and  $115 \mu\text{Sv}$  respectively.

With respect to computation time, the hybrid geometry had a small impact, but was far outweighed by the effect of the relative size of the patients. A change in mesh or CT geometry might add 10–30 min whereas the difference between the male and female simulations was around 4–5 h.



## 4. Discussion

### 4.1. Conclusions from results

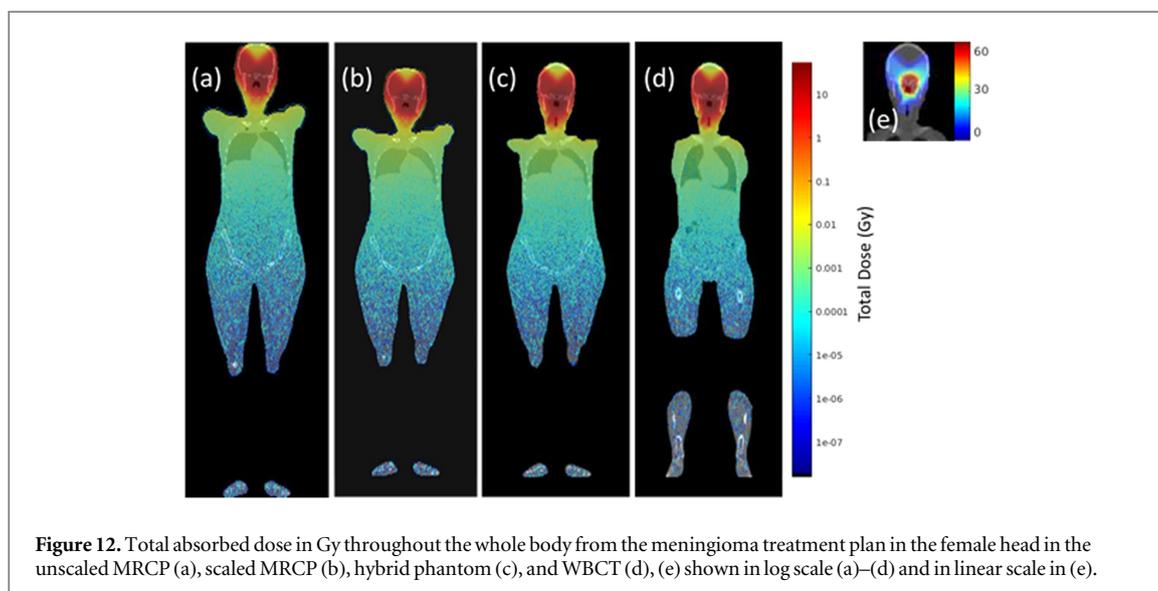
This study has developed a method of creating a whole-body computational patient representation which includes patient-specific images and measurements combined with segments of an adult mesh-type computational phantom. The methodology is flexible with respect to body size, patient sex, and treatment site. As further mesh phantoms are developed, the hybrid process can be refined even further to better fit patients of varied ages, pregnancy status, weight distributions, etc The hybrid geometry was able to be seamlessly constructed in Geant4 Monte Carlo and did not overly extend the computation time compared to the WBCT.

In all treatment sites, for all tested in-field organs for all scorers, the hybrid outperformed the generic mesh phantoms, both scaled and unscaled. This could be anticipated since for the in-field region the WBCT and the hybrid contain identical CT images. Organs partially in the treatment field or outside the radiation field but still within the designated in-field CT region also typically showed an advantage in the hybrid over the MRCPs. The only exception was the neutron dose in the colon in the male prostate treatment plan, which gave a slight 6%/0.17 mGy advantage to the scaled MRCP over the hybrid, and the KERMA-weighted equivalent dose in the same organ in the same treatment site which was within uncertainties of the best performing phantom.

Far outside of the treatment field, the largest factor in whether a given phantom predicted the same value as the WBCT was the distance of that organ from the treatment field. However, matching the organ distance of the WBCT cannot be guaranteed in the hybrid for every organ, since there are inherent relative anatomical differences between the mesh reference phantoms and an individual patient and the scaling factor is a linear adjustment in each dimension. Since the vertical scaling factor is calibrated to spine height, for both pelvic and head and neck treatment sites, the hybrid performed better than or equal to all other phantoms for the tested organs at the most extreme points of the anatomy (bladder for head and neck fields and brain for pelvic fields). The only exception was again in the male pelvis treatment site in the brain, where the scaled MRCP outperformed the hybrid by 9%/1.2  $\mu$ Gy.

In individual treatment sites, there was some observed fluctuation between the relative performance of the hybrid, scaled MRCP, and unscaled MRCP. However, there was only one scorer in one organ in one treatment site where the hybrid was the worst performing of the three phantoms (neutron dose scorer in the liver in the male pelvis). Even in this case, the spread in absolute dose difference between the worst and best performing phantom was 36–40  $\mu$ Gy where the ground truth was 371–373  $\mu$ Gy.

In most cases the hybrid is an improvement over using a generic or even a scaled reference phantom, but no given organ is a guaranteed improvement due to the inherent anatomical differences. This fluctuation in hybrid performance was most frequently observed in organs which are far outside the treatment field, but still part of the central torso. This can be explained by the relative anatomy and placement of the organs in the mesh phantoms versus the CT. For example: in the female head and neck treatment, after scaling down the MRCP, the scaled MRCP and hybrid have a heart placement which is overall closer to the treatment field than in the CT. This results in a higher dose to the heart than in the CT. However, in the unscaled MRCP, while the heart has the same shape and relative anatomical positioning as the hybrid and scaled MRCP, the heart is farther away from



the treatment field than in the hybrid or scaled MRCP. This results in a lower average dose to the organ which happens to align with the expected WBCT value. This relative reduction in out-of-field organ dose was also observed in the female pelvis treatment site, which utilized a similar scaling factor. Figure 12 shows how the dose distribution changes across the whole-body in all patient representations for this treatment site. The reader is referred to figure SM.1 in the supplementary materials for a diagram illustrating this concept. This is an inherent limitation of the simplistic scaling factor and of attempting to create a facsimile of the patient's out-of-field without having knowledge of the unimaged organ's positioning. Without knowing this, there is no reasonable way to predict which organs the hybrid will misrepresent. Furthermore, if this region is known, then it would be better served by being directly included as part of the known in-field, rather than being mimicked by a mesh phantom. The hybrid serves its purpose precisely because this information is absent.

One way the hybrid does exhibit consistent improvement over the mesh phantoms is in predicting the relative composition of the dose distribution. Throughout the body, across all treatment sites (except the heart and lungs in the female head and neck treatment site), the ratio of neutron dose to total dose was best estimated by the hybrid over either of the MRCPs. This ratio can be observed in supplementary materials tables 1–4 and it indicates that, in cases where the hybrid was outperformed, that some combination of organ shape, size, and/or positioning happened to give a better numerical estimate of dose or equivalent dose in the mesh-only phantoms. However this apparent improvement does not necessarily represent an accurate reflection of the dose components distributed in that organ. For example, in the heart in the male head and neck treatment site, the unscaled MRCP overestimates the total dose to the heart by 27%, so while the numerical value of the neutron dose is the closest of the phantoms to the WBCT, it is likely due to an overall excess of dose to the heart such that the neutron component coincidentally aligns with the ground truth.

As an additional observation, due to the resampling required when performing some image translations, the in-field contours/organ masks of the male and female head and neck treatment plan were slightly altered between the WBCT in-field and the hybrid in-field (which contain precisely the same CT image data). This resulted in a slight change in the dose quantities scored in the in-field organs. However, this change does not alter the assessment of the hybrid effectiveness, and it is only relevant when a single WBCT is used for creating hybrids for multiple treatment sites, which would not be the case in a real clinical case.

The neutron equivalent dose was within what could be expected from a scanning system based on the literature (Hälg and Schneider 2020), although in organs far from the treatment field ( $>20$  cm) the simulations in this study typically underestimated the neutron equivalent dose when compared to the results of Hälg and Schneider. For example, in the male head and neck treatment plan, The thyroid at a distance of  $\sim 13.5$  cm from the treatment field yielded a neutron equivalent dose of  $\sim 800 \mu\text{Sv Gy}^{-1}$  which is within the expected range of about  $100\text{--}1000 \mu\text{Sv Gy}^{-1}$  based on the scanning measurements at an angle of  $90^\circ$  from the treatment beam (Hälg and Schneider 2020). At a much greater distance of  $\sim 72$  cm, the bladder in the same treatment plan yielded a neutron equivalent dose of  $\sim 0.8 \mu\text{Sv Gy}^{-1}$ , when according to Hälg and Schneider at this distance the neutron equivalent dose is expected to be slightly less than  $10 \mu\text{Sv Gy}^{-1}$ .

However, there are some conditions which influence the comparison with the work of Hälg and Schneider, for example this work focuses solely on the stray radiation produced within the patient, without the presence of a treatment room or spot scanning machine setup. The lack of this geometry in the simulation environment could

result in the omission of a significant component of stray radiation. From the work of Englbrecht *et al* (2021) it was shown that a position 2 m from an active scanning field at a 90° angle yielded a significant contribution of neutrons in the 0.1–10 MeV energy range from the gantry components and some minor contributions from the outer walls and floor. These neutrons energies coincide with the peak of the ICRP 92 neutron weighting factor model, indicating that this contribution could potentially contribute significantly to the out of field equivalent dose. Additionally, since our study utilized clinically reasonable treatment plans, there were multiple beams at varied angles, which reduces the viability of a comparison with measurements taken at precise angles with respect to a single field.

For the purposes of this work, it was considered sufficient to neglect the influence of the machine and treatment room because (1) the treatment plan utilized a pencil-beam scanning system so the neutron dose from leakage/scattering would not be as large as in a passive scattering system (Hälg and Schneider 2020) and (2) evaluating the relative performance between the WBCT, the hybrid, and the mesh phantoms was of greatest importance.

Overall, in the out-of-field the scaled MRCP and the hybrid were both more likely to provide the closest approximation of the patient WBCT ground truth than the unscaled MRCP. Although some individual organs may be better predicted by one over the other, in most cases the difference between them was small, and typically in the favor of the hybrid. Even in the cases of the most extreme percentile differences between the hybrid and the WBCT in the out-of-field, the absolute difference in the simulated organ dose was quite small. For example a 47.5% difference from the ground truth in total dose in the lungs in the female head and neck was a 0.49 mGy absolute difference, and a 44.4% difference in neutron dose in the liver in the female pelvis was a 0.95 mGy absolute difference. These results fit well with the results of Kuzmin *et al* (2018), which similarly found good agreement in the CT region and variable, though favorable, agreement outside the CT region (specifically referencing figure 7 of Kuzmin *et al* 2018). Also with respect to their work, it is unlikely that our use of deformable image registration in the transition would in general result in any significant dosimetric impact compared to the boundary registration used by Kuzmin *et al* (2018). Deformable image registration could show a potential benefit over boundary registration in cases where there is a significant dose gradient within an organ located in the transition region, where the internal deformation could serve to more accurately reflect the ground truth organ shape. However, this has yet to be investigated.

#### 4.2. Impact and future directions

With the use of the proposed hybrid, whole-body dose and equivalent dose distributions can be simulated via Monte Carlo platforms. The output from these simulations will provide custom estimates of out-of-field quantities while preserving the integrity of the known in-field. This information can be used to predict secondary cancer risk directly from a given treatment plan or alternatively inform organ dose limits in a patient presenting for re-irradiation. It gives the most benefit for organs partially contained in the patient CT which are relatively close to the treatment field. Not only does the hybrid have the most impact from an anatomical/geometric perspective over the mesh phantoms alone, but also this is the region most likely to experience a secondary malignancy (Diallo *et al* 2009). Furthermore, while the hybrid has less patient specificity in the far-field, it is capable of maintaining the accuracy of the treatment planning CT while also providing a realistic whole-body dose distribution data from a single simulation geometry.

While only the original male and female adult MRCPs were used for this study, more mesh phantoms are in development, including pediatric phantoms (Choi *et al* 2021) and adult phantoms with varied body types (Choi *et al* 2020). The hybrid phantom methodology developed in this study will be compatible to any of these new mesh phantoms with only a few adjustments, allowing for both intelligent selection to match body type, and a custom scaling factor to further customize the reference phantom to match patient anatomy.

Next steps include creating hybrids from a larger sample of source WBCTs to further verify hybrid efficacy in different patient anatomies, particularly for patients of higher BMI which might impact the efficacy of the scaling factor and deformable image registration. Additionally, multiple secondary cancer risk models will be implemented to measure not only hybrid risk predictive ability, but also sensitivity to risk model uncertainties and the impact of specific treatment planning decisions on out-of-field secondary cancer risk. The ultimate goal would be to use the hybrid to create risk-optimized treatment plans which minimizes secondary cancer risk while maintaining clinical objectives, as in the work of Wilson and Newhauser (2021) but extended to consider the whole-body using the hybrid phantom.

One hurdle to address in coming studies is the scaling factor implementation. For the hybrids which utilize WBCT source images, the scaling factor is calculated directly from points in the CT. However, when trying to make hybrids from real solid tumor cancer patients, some other method would have to be developed, for example taking physical measurements such as sitting height and waist circumference, or using surface scanning technology to pull measurements from.

A further consideration is that this study utilized Monte Carlo simulations extensively, which increase the accuracy of the results, but at the cost of computational time. Depending on the complexity of the simulation geometry (size of the patients, etc), the time to complete a single batch of simulations ranged from 3.5 to 9 h. Depending on the need for high statistical accuracy in the far-field, it might be possible to combine Monte Carlo simulations in the in and near field with an analytical model such as those examined in the review article by Newhauser *et al* (2017) to get a full picture of the dose without extensive computing times.

Additionally, in future studies the simulation geometry could be further customized by adding a relevant beamline and treatment vault model. While not strictly necessary for comparing phantom performance, without this geometry it is impossible to accurately predict the of out-of-field dose and subsequently the out-of-field risk of secondary cancer for an individual patient. Such a model should be benchmarked against physical measurements to make more absolute calculations of secondary cancer risk for a particular clinical measurement.

For some applications it might be beneficial to consider using the voxelized scaled MRCP in the out-of-field rather than the mesh structure, as the results indicate that the small change in anatomy due to the voxelization is not significant outside the treatment field. However, the mesh should be used when high anatomical fidelity is needed, or when dosimetric data is necessary for structures too fine to voxelize (such as layers of the skin or digestive tract lining).

## Acknowledgments

This work was funded through the Deutsche Forschungsgemeinschaft (DFG) as part of the Research Training Group Graduiertenkolleg (GRK) 2274 joint initiative for Advanced Medical Physics for Image-Guided Cancer Therapy. We would like to additionally thank and acknowledge the support from Erik Traneus and Rasmus Nilsson of RaySearch Laboratories AB (Stockholm, Sweden) for technical assistance and access to RayStation for research purposes.

## ORCID iDs

Erika Kollitz  <https://orcid.org/0000-0002-6331-6171>  
Haegin Han  <https://orcid.org/0000-0002-7635-1611>  
Marco Pinto  <https://orcid.org/0000-0001-6835-2561>  
Marco Schwarz  <https://orcid.org/0000-0001-6964-7732>  
Katia Parodi  <https://orcid.org/0000-0001-7779-6690>

## References

- Agostinelli S *et al* 2003 Geant4—a simulation toolkit *Nucl. Instrum. Methods Phys. Res. A* **506** 250–303
- Baiocco G *et al* 2016 The origin of neutron biological effectiveness as a function of energy *Sci. Rep.* **6** 34033
- Berrington de Gonzalez A, Curtis R E, Kry S F, Gilbert E, Lamart S, Berg C D, Stovall M and Ron E 2011 Proportion of second cancers attributable to radiotherapy treatment in adults: a cohort study in the US SEER cancer registries *Lancet. Oncol.* **12** 353–60
- Brodin N P and Tomé W A 2018 Revisiting the dose constraints for head and neck OARs in the current era of IMRT *Oral Oncol.* **86** 8–18
- Brouwer *et al* 2015 CT-based delineation of organs at risk in the head and neck region: DAHANCA, EORTC, GORTEC, HKNPCSG, NCIC CTG, NCRI, NRG Oncology and TROG consensus guidelines *Radiother. Oncol.* **117** 83–90
- Center for Innovation in Radiation Oncology 2019 Contouring atlases, templates & tools *NRG Oncology*
- Choi C *et al* 2021 Development of skeletal systems for ICRP pediatric mesh-type reference computational phantoms *J. Radiol. Prot.* **41** 139
- Choi C, Yeom Y S, Lee H, Han H, Shin B, Nguyen T T and Kim C H 2020 Body-size-dependent phantom library constructed from ICRP mesh-type reference computational phantoms *Phys. Med. Biol.* **65** 125014
- Clasie B, Wroe A, Kooy H, Depauw N, Flanz J, Paganetti H and Rosenfeld A 2010 Assessment of out-of-field absorbed dose and equivalent dose in proton fields *Med. Phys.* **37** 311–21
- Delaney G, Jacob S, Featherstone C and Barton M 2005 The role of radiotherapy in cancer treatment: estimating optimal utilization from a review of evidence-based clinical guidelines *Cancer* **104** 1129–37
- Diallo I *et al* 2009 Frequency distribution of second solid cancer locations in relation to the irradiated volume among 115 patients treated for childhood cancer *Int. J. Radiat. Oncol. Biol. Phys.* **74** 876–83
- Dracham C B, Shankar A and Madan R 2018 Radiation induced secondary malignancies: a review article *Radiat. Oncol. J.* **36** 85–94
- Eley J, Newhauser W, Homann K, Howell R, Schneider C, Durante M and Bert C 2015 Implementation of an analytical model for leakage neutron equivalent dose in a proton radiotherapy planning system *Cancers* **7** 427–38
- Englbrecht F S, Trinkl S, Mares V, Rühm W, Wielunski M, Wilkens J J, Hillbrand M and Parodi K 2021 A comprehensive Monte Carlo study of out-of-field secondary neutron spectra in a scanned-beam proton therapy gantry room *Z. Med. Phys.* **31** 215–28
- Fontenot J, Taddei P, Zheng Y, Mirkovic D, Jordan T and Newhauser W 2008 Equivalent dose and effective dose from stray radiation during passively scattered proton radiotherapy for prostate cancer *Phys. Med. Biol.* **53** 1677–88
- Hälg R A and Schneider U 2020 Neutron dose and its measurement in proton therapy—current State of Knowledge *Br. J. Radiol.* **93** 1107
- Han H, Yeom Y S, Choi C, Moon S, Shin B, Ha S and Kim C H 2020 POLY2TET: a computer program for conversion of computational human phantoms from polygonal mesh to tetrahedral mesh *J. Radiol. Prot.* **40** 962

- Hashimoto S, Shibamoto Y, Iwata H, Ogino H, Shibata H, Toshito T, Sugie C and Mizoe J E 2016 Whole-pelvic radiotherapy with spot-scanning proton beams for uterine cervical cancer: a planning study *J. Radiat. Res.* **57** 524–32
- Howlader N et al 2021 *SEER cancer statistics review (CSR) 1975-2018* National Cancer Institute ([https://seer.cancer.gov/csr/1975\\_2018/](https://seer.cancer.gov/csr/1975_2018/))
- ICRP 2003 Relative biological effectiveness (RBE), quality factor (Q), and radiation weighting factor (wR). ICRP publication 92 *Ann. ICRP* **33** 1–117 (<https://icrp.org/publication.asp?id=ICRP%20Publication%2092>)
- ICRP 2009 Adult reference computational phantoms. ICRP Publication 110 *Ann. ICRP* **39** 1–165 (<https://icrp.org/publication.asp?id=ICRP%20Publication%20110>)
- Indelicato D et al 2016 *Int. J. Radiat. Oncol. Biol. Phys.* **96** 387–92
- Jagetic L J and Newhauser W D 2015 A simple and fast physics-based analytical method to calculate therapeutic and stray doses from external beam, megavoltage x-ray therapy *Phys. Med. Biol.* **60** 4753–75
- Kim C H et al 2016 The reference phantoms: voxel vs polygon *Ann. ICRP* **45** 188–201
- Kim C H et al 2018 New mesh-type phantoms and their dosimetric applications, including emergencies *Ann. ICRP* **47** 45–62
- Knutsson H and Andersson M 2005 Morphons: segmentation using elastic canvas and paint on priors *IEEE Int. Conf. on Image Processing, ICIP 2005.*, (2:11–1226–9, Sept. 2005) (<https://doi.org/10.1109/ICIP.2005.1530283>)
- Kry S F, Bednarz B, Howell R M, Dauer L, Followill D, Klein E, Paganetti H, Wang B, Wu C - S and George Xu X 2017 AAPM TG 158: Measurement and calculation of doses outside the treated volume from external-beam radiation therapy *Med. Phys.* **44** e391–429
- Kuzmin G A, Mille M M, Jung J W, Lee C, Pelletier C, Akabani G, Lee C and Novel A 2018 Method to extend a partial-body CT for the reconstruction of dose to organs beyond the scan range *Radiat. Res.* **189** 618–26
- Miller K D, Nogueira L, Mariotto A B, Rowland J H, Yabroff K R, Alfano C M, Jemal A, Kramer J L and Siegel R L 2019 Cancer treatment and survivorship statistics *CA Cancer J. Clin.* **69** 363–85
- Morton L M, Onel K, Curtis R E, Hungate E A and Armstrong G T 2014 The rising incidence of second cancers: patterns of occurrence and identification of risk factors for children and adults *Am. Soc. Clin. Oncol. Edu. Book* **34** e57–67
- Newhauser W D et al 2009 The risk of developing a second cancer after receiving craniospinal proton irradiation *Phys. Med. Biol.* **54** 2277–91
- Newhauser W D, Berrington de Gonzalez A, Schulte R and Lee C 2016 A review of radiotherapy-induced late effects Research after advanced technology treatments *Frontiers Oncol.* **6** 13
- Newhauser W D, Schneider C, Wilson L, Shrestha S and Donahue W 2017 A review of analytical models of stray radiation exposures from photon- and proton-beam radiotherapies *Radiat. Prot. Dosim.* **180** 245–51
- Oeffinger K et al 2006 Chronic health conditions in adult survivors of childhood cancer *New Engl. J. Med.* **355** 1572–82
- Plumat J, Andersson M, Janssens G, Orban de Xivry J, Knutsson H and Macq B 2009 Image registration using the Morphon algorithm: an ITK implementation *The Insight Journal* (<https://doi.org/10.54294/kmysv2>)
- Romano K D, Trifiletti D M and Read P W 2016 ARROCase Nasopharynx Cancer Assoc. Residents *Radiat. Oncol.* (<https://astro.org/Affiliate/ARRO/Resident-Resources/Educational-Resources/ARROcase/Head-Neck-and-Skin>)
- Rombi B and Timmermann B 2014 Proton beam therapy for pediatric chordomas: state of the art *Int. J. Part. Ther.* **1** 368–85
- Sasidharan B K, Aljabab S, Saini J, Wong T, Laramore G, Liao J, Parvathaneni U and Bowen S 2019 Clinical Monte Carlo versus Pencil Beam Treatment Planning in Nasopharyngeal Patients Receiving IMPT *Int. J. Part. Ther.* **5** 32–40
- Schmid S, Landry G, Thieke C, Verhaegen F, Ganswindt U, Belka C, Parodi K and Dedes G 2015 Monte Carlo study on the sensitivity of prompt gamma imaging to proton range variations due to interfractional changes in prostate cancer patients *Phys. Med. Biol.* **60** 9329–47
- Schneider C, Newhauser W D, Wilson L and Kapsch R P 2019 A physics-based analytical model of absorbed dose from primary, leakage, and scattered photons from megavoltage radiotherapy with MLCs *Phys. Med. Biol.* **64** 185017
- Schneider U and Hälgl R 2015 The impact of neutrons in clinical proton therapy *Front. Oncol.* **5** 235
- Schneider W, Bortfeld T and Schlegel W 2000 Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions *Phys. Med. Biol.* **45** 459–78
- Si H 2015 Tetgen, a Delaunay-based quality tetrahedral mesh generator *ACM Trans. Math. Soft.* **41** 2–11
- Toita T et al 2011 A consensus-based guideline defining clinical target volume for primary disease in external beam radiotherapy for intact uterine cervical cancer *Japan. J. Clin. Oncol.* **41** 1119–26
- Wang Y T, Jiang C F F, Schroeder C and Teran J 2014 An Adaptive virtual node algorithm with robust mesh cutting *ACM SIGGRAPH/Eurographics Symp. on Computer Animation (SCA)* pp 77–85
- Wilson L J and Newhauser W D 2021 *Phys. Med. Biol.* **66** 065007
- Wilson L J, Newhauser W D, Schneider C W, Kamp F, Reiner M, Martins J C, Landry G, Giussani A, Kapsch R P and Parodi K 2020 Method to quickly and accurately calculate absorbed dose from therapeutic and stray photon exposures throughout the entire body in individual patients *Med. Phys.* **47** 2254–66
- Wilson R R 1946 Radiological use of fast protons *Radiology* **47** 487–91
- Yeom Y S, Jeong J H, Han M C and Kim C H 2014 Tetrahedral-mesh-based computational human phantom for fast Monte Carlo dose calculations *Phys. Med. Biol.* **59** 3173
- Zhang R, Howell R M, Giebel A, Taddei P J, Mahajan A and Newhauser W D 2013 Comparison of risk of radiogenic second cancer following photon and proton craniospinal irradiation for a pediatric medulloblastoma patient *Phys. Med. Biol.* **58** 807–23
- Zvereva A, Schlattl H, Zankl M, Becker J, Petoussi-Hens N, Yeom Y S, Kim C H, Hoeschen C and Parodi K 2017 Feasibility of reducing differences in estimated doses in nuclear medicine between a patient-specific and a reference phantom *Phys. Med.: Eur. J. Med. Phys.* **39** 100–12