Virtual cone-beam computed tomography simulator with human phantom library and its application to the elemental material decomposition

Taisei Shimomura [a,b], Daiyu Fujiwara [a], Yuki Inoue [a], Atsushi Takeya [a], Takeshi Ohta [b], Yuki Nozawa [b], Toshikazu Imae [b], Kanabu Nawa [b], Keiichi Nakagawa [b], Akihiro Haga [a,*]

[a] *Graduate School of Biomedical Sciences, Tokushima University, Tokushima 770-8503, Japan*
[b] *Department of Radiology, The University of Tokyo Hospital, Bunkyo, Tokyo 113-8655, Japan*

ARTICLE INFO

ABSTRACT

Purpose: The purpose of this study is to develop a virtual CBCT simulator with a head and neck (HN) human phantom library and to demonstrate the feasibility of elemental material decomposition (EMD) for quantitative CBCT imaging using this virtual simulator.
Methods: The library of 36 HN human phantoms were developed by extending the ICRP 110 adult phantoms based on human age, height, and weight statistics. To create the CBCT database for the library, a virtual CBCT simulator that simulated the direct and scattered X-ray on a flat panel detector using ray-tracing and deep-learning (DL) models was used. Gaussian distributed noise was also included on the flat panel detector, which was evaluated using a real CBCT system. The usefulness of the virtual CBCT system was demonstrated through the application of the developed DL-based EMD model for case involving virtual phantom and real patient.
Results: The virtual simulator could generate various virtual CBCT images based on the human phantom library, and the prediction of the EMD could be successfully performed by preparing the CBCT database from the proposed virtual system, even for a real patient. The CBCT image degradation owing to the scattered X-ray and the statistical noise affected the prediction accuracy, although these effects were minimal. Furthermore, the elemental distribution using the real CBCT image was also predictable.
Conclusions: This study demonstrated the potential of using computer vision for medical data preparation and analysis, which could have important implications for improving patient outcomes, especially in adaptive radiation therapy.

1. Introduction

Cone-beam computed tomography (CBCT) is essential in image-guided radiation therapy (IGRT) for accurate patient setup and dose distribution [1]. Three-dimensional (3D) information obtained from CBCT ensures precise patient positioning and treatment quality [2]. Reconstructing the actual dose distribution and monitoring differences from the planned dose in target areas and organs-at-risk (OAR) can be achieved using treatment planning with CBCT and machine logfiles [3,4]. Adaptive radiation therapy (ART) relies on the replanning on CBCT to optimize dose distribution based on changes in patient weight, target size, and anatomical deformations [5]. During CBCT-based replanning for ART, scattered X-rays cause image quality degradation, hindering quantitative dose evaluation. Various studies have explored hardware (e.g., anti-scattering grid [6], lead-strip blockers [7]) and

software (e.g., 3D scattering kernel modeling [8], fast Monte Carlo simulation [9]) approaches to address this issue. Deep learning (DL) methods have also been investigated, including CBCT-to-CT image conversion and scattering map removal on the flat panel detector (FPD) [10–12].

Improving CBCT image quality enhances the accuracy of electron density (ED) estimation using CT-to-ED lookup tables or the stoichiometric method [13,14]. While ED is sufficient for dose calculation in X-ray radiation therapy, in hadron therapy, elemental information of anatomies is valuable due to dose uncertainty arising from nuclear fragmentation reactions [15,16]. However, most current dose calculation approaches rely on ED or, at most, stopping-power ratio (SPR) to water, leading to prevalent uncertainties [17]. Ideally, incorporating elemental information into dose calculation would be desirable.

This study aimed to develop a virtual CBCT system and demonstrate

---

its usefulness by applying a DL-based material decomposition (MD) model. The development of the virtual CBCT system consists of two parts: a human phantom library part and a projection-reconstruction part. In this study, the developed human phantom library comprised H, C, N, O, P, and Ca, based on deformation data from the International Commission on Radiological Protection (ICRP) 110 adult female/male phantoms. These data were virtually scanned, and CBCT images were reconstructed in the virtual simulator. The proposed simulator was compared with the existing CBCT model through Monte Carlo simulations (MCGPU) [18].

The reconstructed CBCT images were then utilized in the MD process. In a clinical setting, creating a large database containing both CBCT images and corresponding material information is unrealistic. Additionally, there is inherent uncertainty in the ground truth, particularly in terms of material fraction and density, which is inevitable in MD cases. The virtual simulator addresses these issues. Despite several studies on MD from CBCT [19,20], which employ MD models based on dual- or multi-energy CBCT images, the proposed approach is markedly different. Specifically, we introduced a generative model for training datasets based on the virtual system. In other words, the database of CBCT images was generated using a virtual CBCT system with a human phantom library that includes material information.

The feasibility of the clinical application was validated based on the results of the experimental data obtained with real CBCT machine equipped on the radiotherapy system. Despite the study on the effect of the scattering in the MD method [20], its impact on the performance of elemental MD (EMD) has never evaluated. Therefore, considering the influence of the scattered X-ray on the image quality of input CBCT, we also investigated the impact of the removal of the scattered X-ray on the estimation accuracy of EMD.

2. Materials and methods

Fig. 1 provides an overview of this study, illustrating the development of the virtual CBCT system and its application to EMD. The ICRP 110 reference human phantoms were utilized to generate a human library with various voxel-based human phantoms following the humanity, and the corresponding CBCT images with/without scattered X-ray were produced by the CBCT simulator. The CBCT database labeled with the elemental material information was then used to develop the DL-based MD model. Thereafter, the developed model was validated using the testing datasets and applied to the experimental and clinical CBCT data for Rando phantoms that acquired in an image-guidance system, respectively.

This study specifically targeted the head and neck (HN) region, which represents one of the most prevalent cancer types, with approximately 75% of patients undergoing radiation therapy [21,22]. With the significance of this region in clinical practice, focusing on the HN region

allowed us to assess the feasibility of EMD from CBCT images effectively. In the following sections, we have provided separate descriptions of the methods employed for the virtual CBCT system and the application of EMD.

2.1. Virtual CBCT image generation

2.1.1. Human phantom generation

In this study, a human library was developed considering the human characteristics, such as, body size (weight and height) and the age, and the deviation of material-density fraction, which was also summarized in the Appendix. Two ICRP 110 reference adult human phantoms (adult female (AF) and adult male (AM)) were employed as the base phantoms. Because the spatial resolution in these two phantoms is relatively low ($1.775 \times 1.775 \times 4.84$ mm^3 for female and $2.137 \times 2.137 \times 8.0$ mm^3 for male) [23], both phantoms were enhanced with the higher resolution of $1.0 \times 1.0 \times 2.0$ mm^3 before developing the human library; the nearest neighbor method was applied in the x-y plane, whereas the deformed vector field (DVF) method using elastix [24] was applied in the z direction for the higher resolution production. The latter is a method used to generate intermediate slices via the linear transform of the DVF produced from the deformable image registration (DIR) between the adjacent two slices of original ICRP phantoms. For instance, in the AM phantom, where 3 slices must be inserted to achieve the targeted resolution, the DVF produced from DIR of two adjacent slices was multiplied by factors of 1/4, 1/2, and 3/4, and then the deformed vectors were operated to one of the slices to generate the intermediate slices. Table 1 summarizes the property of high-resolution AM and AF phantoms developed in this study.

Based on the high-resolution ICRP reference phantoms, various human phantoms were further generated by following the method proposed in another study [25]. In this study, only the CBCT images with the HN region were used. Therefore, we focused on the generation of our human library used for the HN region, where the height and the weight

Table 1
Information of the high-resolution computational phantoms extended from the International Commission on Radiological Protection (ICRP) 110 adult phantoms [23].

Property	Adult female	Adult male
Height [m]	1.63	1.76
Weight [kg]	60	73
Age [year]	43	38
Slice thickness (voxel height) [mm]	2	2
Voxel in plane resolution [mm]	1	1
Number of columns	543	531
Number of rows	271	243
Number of slices	838	879
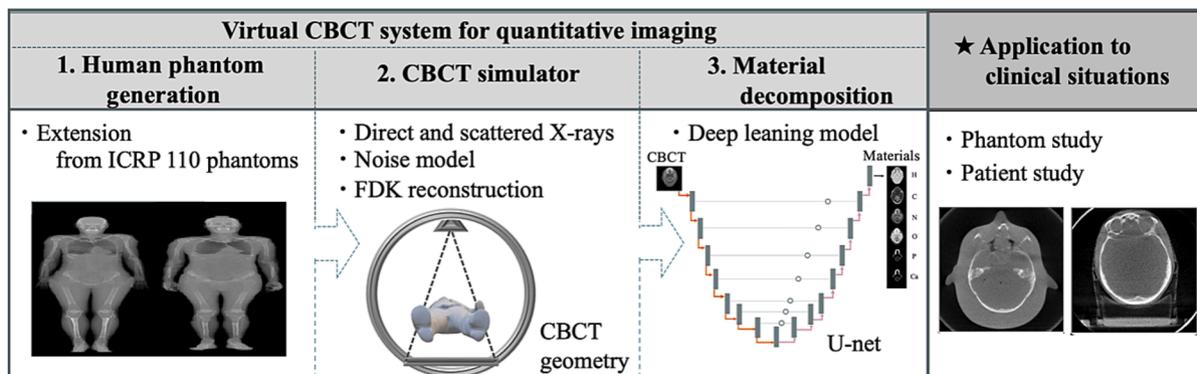
Fig. 1. Overview of this study: a human phantom library and a cone-beam computed tomography (CBCT) simulator are developed to create the CBCT database. With this database, the deep learning model for material decomposition is then trained. Experimental CBCT images are used in clinical application.

were first sampled to evaluate the head circumference *HC* [cm] as follows:

$$HC = 41.02 + 0.08673*H + w_{sd}*sd, \text{(for female)}, \quad (1)$$

$$HC = 42.40 + 0.08673*H + w_{sd}*sd, \text{ (for male)}, \quad (2)$$

where H is the height [cm] and sd is the standard deviation (this study set $sd = 1.41$ [cm]) [26]. Using HC, the high-resolution ICRP phantoms were then rescaled by HC/HC_ref, where HC_ref is the head circumference obtained from the ICRP phantoms. Further, w_{sd} is the confidence coefficient of the statistical distribution model and in this study, the normal distribution of the weight [kg] was employed; that is, 25%, 50%, and 75% of the weight yielded w_{sd} as -0.67, 0, and $+0.67$, respectively. The percentiles including 25%, 50%, and 75% in height and weight statistics extracted from the NHANES database [27] for 60-, 70-, and 80-year-old groups were considered to generate the human phantoms. Thus, in total, 36 human phantoms (18 each for female and male) were generated. Fig. 2 shows the frontal and rotated views of the representative phantoms for female (left) and male (right), where the age of 60 years, 75% height, and 25% weight were employed.

In the phantom generation, a non-rigid deformation was adopted to the brain, where the DVF was artificially generated (See Appendix). The physical and elemental densities of each anatomy was assigned by considering the deviation from the referenced value provided in ICRP 110 report according to the age, height, and weight without considering the difference of gender. In this study, we randomly added a 10% Gaussian deviation to the elemental composition. However, for the density value of medium, Gaussian deviation of 5% was randomly added except in case of bone and teeth. In teeth, multiplied factor of 0.9–1.0 was uniformly sampled, whereas in bone, considering the variation over aging based on gender [28], the factor of the aging for each female and male was multiplied with the density value.

In this study, additional human phantoms, which are independent of the above human library, were prepared by 3D DIR from the ICRP 110 human phantom for real HN CT images collected from The Cancer Imaging Archive (TCIA) [29]. Briefly, the process is as follows. 3D CT images were downloaded and the sequential preprocessing such as removing the couch and supporter was performed. Then, 3D DIR was performed as the moving and the targeting 3D images were the 3D density images of ICRP 110 phantoms and the preprocessed TCIA 3D images, respectively. However, as the deformed 3D images generate an inconsistent geometrical connection in certain anatomies, these were manually modified. Thus, an additional human phantom was created. In this study, this phantom was only used as the test data, which have been shown in the Appendix.

2.1.2. CBCT simulator

In this subsection, the virtual CBCT image production using the aforementioned human library is described (Fig. 3). The implemented CBCT machine geometry was referred to from an Elekta X-ray Volumetric Imaging (XVI) system, where the source-to-isocenter and the source-to-detector distances were set as 100 and 153.8 cm, respectively. Further, the flat panel detector (FPD), measuring 512×512 pixels, with a 0.8-mm scale in the actual system, was set to the size of 400×400 pixels with 1-mm scale in the virtual system to reduce the data volume of the cone-beam projections. In this study, "no-shifted FPD" was considered [30]. Although the extended field of view (FOV) could be applicable by the shifting FPD technique, the CBCT imaging to the HN region have a sufficient FOV to view the entire HN with no-shifted FPD.

The cone-beam projection images were produced by considering 1) direct X-ray, 2) scattered X-ray, and 3) the statistical noise. The contribution from 1) was simulated with the employed X-ray energy spectrum and the elemental information of human phantom based on the material-based forward projection algorithm (MBFPA) [31,32]. Whereas, that from 2) was simulated by the DL model on the FPD using a database created with MC simulation. Finally, contribution of 3) on the projection image was modeled with the normal distribution by referring to a signal-to-noise ratio (SNR) on the reconstructed CBCT image domain. This projection model is much faster than the MC simulation. Typically, CBCT projection data (per projection) can be created within 12 s using our model, whereas the MC (6×10^9 events) simulation requires 183 s, using the computer employed in this study (Section 2.3).

2.1.2.1. Direct X-ray simulation.
Direct X-ray on the FPD was simulated by MBFPA based on the information of both the materials in an object and the X-ray spectrum. MBFPA calculates the intensity ratio of X-rays at the ith detector I_i, to the source I_0 according to Lambert-Beer's law with the line attenuation coefficient $\mu_j(E)$ in the jth voxel as,

$$\frac{I_i}{I_0} = \frac{\sum_E \alpha(E) n_0(E) e^{\sum_j -a_{ij}\mu_j(E)}}{\sum_E \alpha(E) n_0(E)}, \quad (3)$$

where E is the photon energy, $\alpha(E)$ is the fraction of the corresponding photon energy bin, n_0 is the photon number in the X-ray source, and a_{ij} is the photon pass length in voxel j of the object, representing an element known as *system matrix*. In this study, the X-ray energy spectrum was estimated using an artificial neural network model proposed in a previous study[33], based on the percent depth dose (PDD) measurement using the real Elekta XVI system (100 kVp, discretized with a bin width of 1 keV). The measurements were performed via the Elekta XVI system in University of Tokyo Hospital and the estimated spectrum is depicted
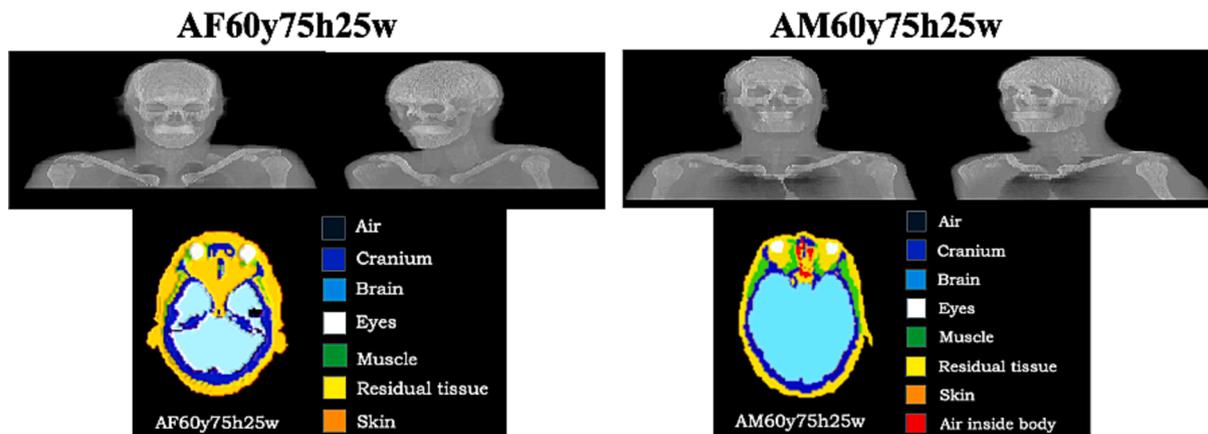

Fig. 2. Examples in human library. "AF60y75h25w" implies 60 years, 75% in height, and 25% in weight for female, whereas "AM60y75h25w" implies 60 years, 75% in height, and 25% in weight for male. The organs in phantoms assigned in ICRP 110 report [23] were considered. The densities in each anatomy are reassigned depending on the age, height, and weight without considering the difference of gender. Further descriptions are provided in the Appendix.
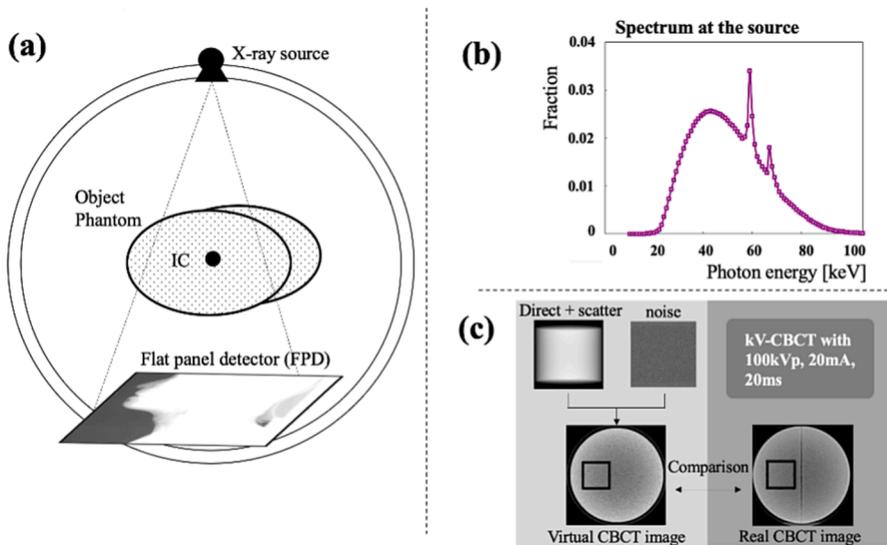
Fig. 3. Virtual cone-beam computed tomography (CBCT); (a) overview, (b) the energy spectrum employed in this study (100 kVp), and (c) the adjustment of the noise model. The energy spectrum in (b) was estimated from the percent depth dose measurement [33]. The standard deviation of Eq. (6) was determined by comparing the signal-to-noise ratio (SNR) of CBCT image for a 25-cm diameter cylindrical water phantom reconstructed in virtual system with that of real CBCT image within a box in (c). In the real CBCT scan, 100 kVp, 20 mA, and 20 ms parameters were employed.

in Fig. 3(b). The attenuation $\mu_j(E)$ in the jth voxel can be expressed as the sum of the attenuation coefficients for each element m included in an object,

$$\mu_j(E) = \sum_m w_m \mu_{m,j}(E, Z, \rho), \tag{4}$$

where w_m is the weight (fraction) of the mth element and for the human phantoms, $m \in$ H, C, N, O, P, and Ca were considered. For the energy range considered in this study, the linear attenuation coefficient $\mu_{m,j}(E, Z, \rho)$ can be written as the sum of the processes of the photoelectric effect, Compton scattering, and Rayleigh scattering as follows:

$$\mu_{m,j}(E, Z, \rho) = \rho Z \frac{N_A}{A} [\sigma_{pe}(E, Z) + \sigma_{Comp}(E, Z) + \sigma_{Rayl}(E, Z)], \tag{5}$$

where N_A and A are the Avogadro constant and atomic weight, respectively, and $\sigma_{pe}, \sigma_{Comp}$, and σ_{Rayl} are the cross sections owing to the photoelectric effect, Compton scattering, and Rayleigh scattering, respectively, which were obtained from the open X-ray database provided in a previous study[34].

2.1.2.2. Scattered X-ray simulation. Because the scattered X-ray is the primary cause of the image degradation of CBCT using FPD, its removal on the CBCT has been actively studied by many researchers. In this study, the accurate simulation of the scattered X-ray is crucial to establishing the virtual CBCT simulator. This study used a U-net model in DL to generate the scattered map on the FPD [12,35]; the network structure is shown in Fig. 4. The training database was composed of the scattered X-ray distribution on the FPD and the corresponding direct X-ray image, both of which were generated by the GPU accelerated MC
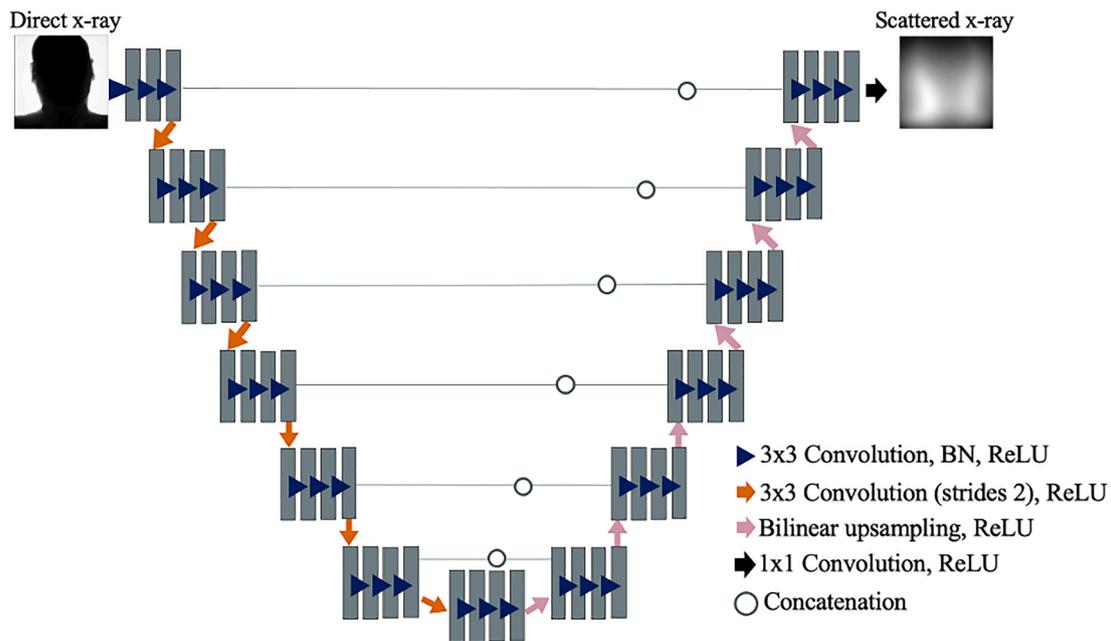

Fig. 4. U-net model for scattered X-ray distribution on the flat-panel detector (FPD), where the input is the direct X-ray image and the output is the scattered one. Encoding and decoding parts have 7 layers. Three consecutive 3×3 convolutions followed by a rectified linear unit (ReLU) are applied for each level of the encoding path. At the end of each part, the spatial resolution is reduced by two while the number of channel is doubled by applying a 3×3 convolution with 2×2 stride, also followed by a ReLU. In the decoding part, the last convolution of each layer is replaced by a bilinear upsampling.

simulation code, that is, MCGPU [18] with 6×10^9 photons. The cone-beam projection images for 30 human phantoms with 90 directions (4° interval) of gantry angle were used for the model development. To moderate the statistical noise due to the limited number of photons on the MC-based projections, a 5×5 median filter was applied to the simulation results. The training datasets were then augmented by flipping the projections vertically and horizontally, which resulted in 8100 ($=30 \times 90 \times 3$) projections. In addition, the size of the input and output images was 256×256, which was rescaled into the size of FPD, that is, 400×400 with 1-mm scale.

The training of the U-net model was performed by using Google's deep learning framework TensorFlow (version 1.12.0) and the Keras library (version 2.1.6), with 900 epochs and batch size of 4. Further, the Adam optimizer was used, where the learning rate was dynamically reduced to half after five epochs without improvement from the initial value of 0.0004. The mean absolute error (MAE) was employed as a loss function. In addition, the model weights were initialized using a Glorot uniform initialization [36]. In this study, 20% of the training data were used for validation to avoid the overfitting and tune the above parameters. The training via the computer used in this study was completed in 35 h.

2.1.2.3. Noise.
Noise is inevitable in actual CT acquisition. In this study, the noise in X-ray detectors was modeled using the Gaussian distributed random noise for the X-ray intensity as follows,

$$I_i^* = I_i + \sigma_i z_i, \tag{6}$$

where I_i^* is the X-ray intensity generated in the *i*-th detector of a virtual CT, and z_i is a standard normal random variable. Typically, the Poisson distribution is used in the noise model; however, it can be approximated via the Gaussian distribution with the standard deviation σ_i $I_0 \sqrt{I_i}$ provided the intensity is sufficiently high. Specifically, we set $10^5 < I_0 < 10^6$ to have the same order of signal-to-noise (SNR) ratio as that observed in a real CT system (SNR $\sim$ 10, Section 3.1).

2.1.3. Evaluation of the virtual CBCT simulator
The evaluation of the virtual CBCT simulator was performed in two stages; one was to evaluate the projection image by comparing with the MC simulation and the other was to evaluate the reconstructed CBCT image through comparisons with the real CBCT acquired in the phantoms. The former evaluates by dividing into the direct X-ray and the scattered X-ray. In the MC simulation, MCGPU was employed, where the scattered X-ray includes the photons that travel into the FPD with the interaction histories of the Compton and Rayleigh scatterings, whereas the direct X-ray includes the residual photons from the photoelectric process in addition to those scatterings. The phantom resolution was $1 \times 1 \times 2$ mm^3 and 22-classified anatomical information was considered both in the MC and virtual CBCT simulations. In particular, the quantitative evaluation was performed on the scattered X-ray maps with 540 projections (90 gantry angles $\times$ 6 human phantoms in our human library) using the mean absolute percentage error (MAPE).

$$MAPE = 100 * \frac{1}{N} \sum_i^N \frac{|s_i^{pred} - s_i^{MC}|}{s_i^{MC}} \tag{7}$$

where s_i^{pred} and s_i^{MC} are the model-predicted and MC scattered X-ray counts in the ith detector, respectively, and N is the number of the pixels in FPD.

The latter evaluates the SNR on a virtual reconstructed CBCT image through comparisons with that in the Rando phantom (Radiology Support Devices, Inc. USA) acquired in the real CBCT system (Elekta XVI system at The University of Tokyo Hospital). The virtual CBCT image included the artifact via the noise generated on the projections (Eq. (6)), whose magnitude was determined by the water phantom measurement with a nominal tube voltage of 100 kVp, 20-cm filter (S20 filter), no

bowtie filter, 20 mA, 20 ms, and the full-angle rotation. This magnitude was validated on the additional measurement using the Rando phantom with the same protocol as that in the water phantom measurement. Thus, the SNR of the CBCT image for a brain area of the Rando phantom was compared with that in one of the virtual human phantoms of the test data.

2.2. Application

2.2.1. Material decomposition model
As an application of the virtual CBCT simulator, the DL-based EMD model was developed and its performance was investigated. For the EMD model, a large training dataset including the elemental information and the corresponding CBCT image is required. As the collection of such dataset clinically is impractical, the EMD model based on the virtual CBCT simulator demonstrates one of the advantages of using this simulator.

2.2.1.1. Datasets.
The human library described in Section 2.1.1 includes the information of the elemental distribution. Each phantom in the human library was projected over 360° with 1° increment using the virtual CBCT simulator described in Section 2.1.2. Based on this cone-beam projection data, CBCT images were reconstructed using the Feldkamp-Davis-Kress (FDK) method [37,38], where the reconstructed sizes were $270 \times 270 \times 120$ and $270 \times 270 \times 140$ pixels for female and male phantoms, respectively, with $1 \times 1 \times 2$ mm^3 resolution. Almost all CBCT systems in clinical applications employ the FDK algorithm; therefore, the CBCT image should be prepared with this algorithm considering the clinical application of the material decomposition model. Thus, the CBCT database labeled with the elemental information was created and then used in the training of the following DL model.

2.2.1.2. Deep learning model.
The U-net based EMD model was developed as shown in Fig. 5, where, the input is a 2D image of each CBCT slice cropped as 256×256 pixels [39]. Through downsampling, 4×4 convolution with 2×2 stride was applied for each level of the encoding path. Further, each convolutional layer was batch-normalized and was activated by ReLU, except for the downsampling layers, where LeakyReLU was used after batch normalization. Upsampling and skip connection were subsequently performed to output six images corresponding to the six elemental distributions. In the expansive path, 2×2 nearest neighbor upsampling was designed to restore the spatial resolution, similar to the strided convolution in the contracting path. In total, 2340 images generated from 9 phantoms $\times$ 120 slices for females and 9 phantoms $\times$ 140 slices for males were used to train the U-net based MD model, with 468 (20%) being used for the model validation. The mean squared error (MSE) was selected as a loss function of the training with the Adam optimizer, and the learning rate was dynamically reduced by a half after five epochs without improvement (a minimum learning rate of 0.00001 was set). The training required 8 h.

2.2.2. Evaluation of MD model and the application to the real CBCT images

2.2.2.1. Human phantom study.
The developed U-net based EMD model was evaluated by 6 human phantoms belonging to our human library but excluded in the training data. In addition, human phantoms created from real 3D CT images were also used in the EMD model evaluation. To observe the effect of scattered X-ray and the noise, cases with and without scatter/noise were evaluated. The evaluation was performed using the root mean squared error (RMSE):

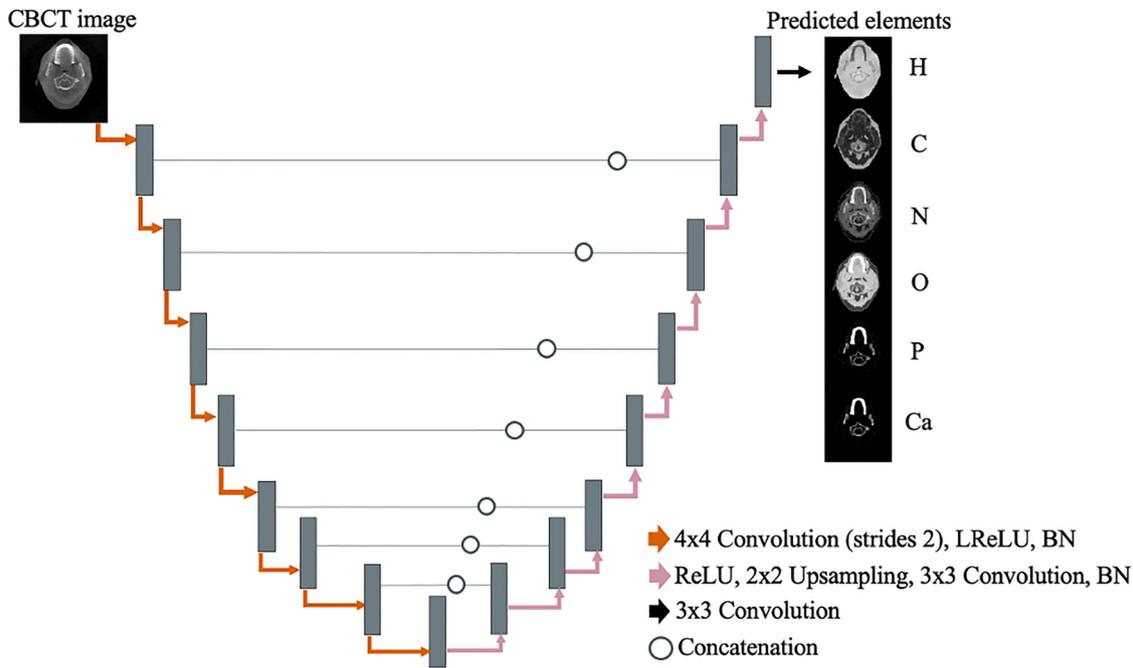$$RMSE_m = \sqrt{\frac{\sum_i^N \left(\rho_{m,i}^{pred} - \rho_{m,i}^{GT} \right)^2}{N}} \tag{8}$$

Fig. 5. U-net model for elemental material decomposition prediction. A 4 × 4 convolution (strides with 2) followed by a leaky rectified linear unit (LReLU) is applied for each level of the encoding path. In the decoding part, a 2 × 2 upsampling is followed by the 3 × 3 convolution in each layer.

where $m \in$ (H, C, N, O, P, Ca), $\rho_{m,i}^{pred}$ ($\rho_{m,i}^{GT}$) is the predicted (ground truth) density for m element in the ith pixel, and N is the number of the pixels in the reconstructed EMD image. In addition to RMSE, the structural similarity index measure (SSIM) with the ground truth in each six macro-element was also evaluated as follows

$$SSIM_m = \frac{(2\mu^{pred}\mu^{GT} + c_1)(\sigma^{pred,GT} + c_2)}{\left(\left(\mu^{pred}\right)^2 + \left(\mu^{GT}\right)^2 + c_1\right)\left(\left(\sigma^{pred}\right)^2 + \left(\sigma^{GT}\right)^2 + c_2\right)}, \quad (9)$$

where μ^{pred} (μ^{GT}) and σ^{pred} (σ^{GT}) are the average value and standard deviation of the predicted (true) density inside the region-of-interest, respectively, $\sigma^{pred,GT}$ is the covariance between the predicted and the ground truth density, and the constant values, c_1 and c_2, are included to avoid instability [40]. To perform the RMSE and SSIM evaluation, masking was performed to restrict the evaluation area as the bounding box of the anatomical region.

2.2.2.2. Rando phantom and patient study. The U-net based EMD model was applied to the real Rando phantom and patient data obtained from The University of Tokyo Hospital. The CBCT images of Rando phantom and an HN cancer patient were used to demonstrate the feasibility of the proposed EMD method. These were acquired using the Elekta XVI system with a nominal tube voltage of 100 kVp, S20 filter, no bowtie filter, 20 mA, 20 ms, and the full rotation for Rando phantom, and 10 mA, 10 ms, and the half-angle rotation for the patient, respectively. The size of the reconstructed images was 270 × 270 pixels with 1 × 1 mm² resolution, which was cropped to 256 × 256 pixels for use as the input of the U-net based EMD model. In the experiment using the real system, no ground truth for the elemental distribution was provided. Therefore, in this study, we defined the ground truth using the *planning* CT (Canon Aquilion LB) in the Rando phantom experiment and compared it with the elemental distribution predicted using the CBCT image to check the validity. Details regarding the EMD model using the planning CT can be found elsewhere [39]. Because of the different timings of the scans between the planning CT and CBCT, a rigorous quantitative comparison is still difficult. However, based on comparisons in terms of the nearest slices, an assessment of whether the EMD obtained from the CBCT

images is consistent with that obtained from the planning CT images is possible.

2.3. Environment of computation

Data-preparation process and model training and testing were implemented using C++ and Python programing codes, respectively. A Linux server equipped with an NVIDIA TITAN RTX GPU with 24 GB of memory was used in this study. Our developed CBCT simulator and the high-resolution version of ICRP 110 human reference phantoms are available in the Github repository.

https://github.com/hagaakihiro/VirtualCBCT

3. Results

3.1. Evaluation of virtual CBCT image

Fig. 6 shows the images produced by the virtual CBCT system developed in this study, where the projection image for (a) the direct X-ray and (b) for the scattered X-ray, and (c) the reconstructed CBCT images are depicted. As evident in (a), the direct X-ray from the virtual system is comparable with the MC result. The small difference presumably originates from the difference in the observation of the X-ray photon. The MBFPA focuses on the X-ray path starting from the X-ray source to the center of the detector bin. In contrast, the MC method observes the photons somewhere in the detector bin; thus, it can indicate the difference in the projection, particularly, at the edge of the high density such as the bone.

As evident in (b), the generation of the scattered X-ray image in the virtual system also works well when using the U-net model shown in Fig. 4. As described in Section 2.1.2, the present model was developed with the training data obtained by adopting the median filter on the scattered map calculated by MC simulation; thus, the scattered X-ray image generated by the virtual system was smoother than that by MC simulation. Quantitatively, the MAPE between the present model and the MC simulation was 1.99% with 540 projections (6 test phantoms) on average, and in any case, the MAPE was smaller than 10%. Here, <30 s was required to generate of one projection image. Based on the direct
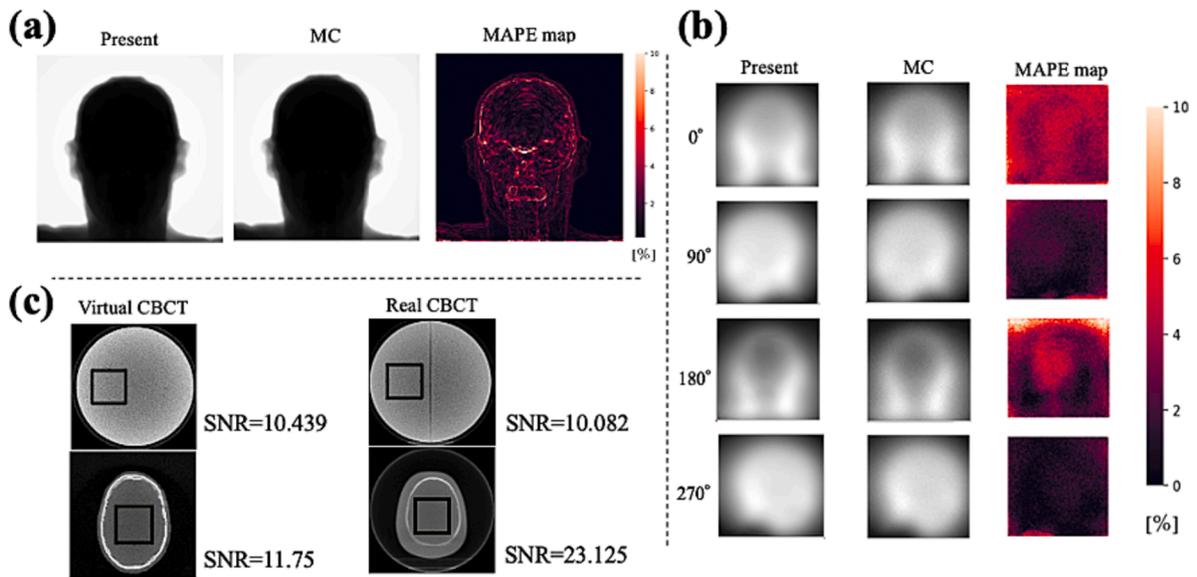
Fig. 6. Projection image and the reconstructed image with virtual cone-beam computed tomography; (a) direct and (b) scattered X-ray images on flat panel detector (FPD) are depicted, where for comparison, mean absolute percentage error (MAPE) maps between our method and Monte Carlo method are shown on the right side. (c) Reconstructed CBCT images after adding the Gaussian distributed noise on FPD are depicted, where in the real CBCT system, 100 kVp, 20 mA, and 20 ms parameters were employed. The signal-to-noise ratio (SNR) is evaluated inside the box depicted.

and scattered X-ray results, it can be concluded that the developed virtual CBCT system produced the accurate projection image very efficiently.

Through the addition of noise in the above projection image, the CBCT images were generated as depicted in Fig. 6(c), where the upper two images are the reconstructed images of the water phantoms, whereas the lower two images are that of the human phantoms. As described in Section 2.1.2, because the amount of the noise was modeled by the SNR in the reconstructed image of the water phantom, the SNR on

the virtual CBCT was comparable with that on the real CBCT for the water phantom. However, in case of the human phantoms, a difference in the SNR between the virtual and the real human phantoms (Rando phantom) was observed. This difference may be attributed to the difference in the sizes of the human and water phantoms used in the noise model development; A relatively large size of the water phantom (25 cm diameter) generated a large noise model. This result implies that this noise model is excessively large to simulate the CBCT image for HN region. Nevertheless, this model can still be employed because the
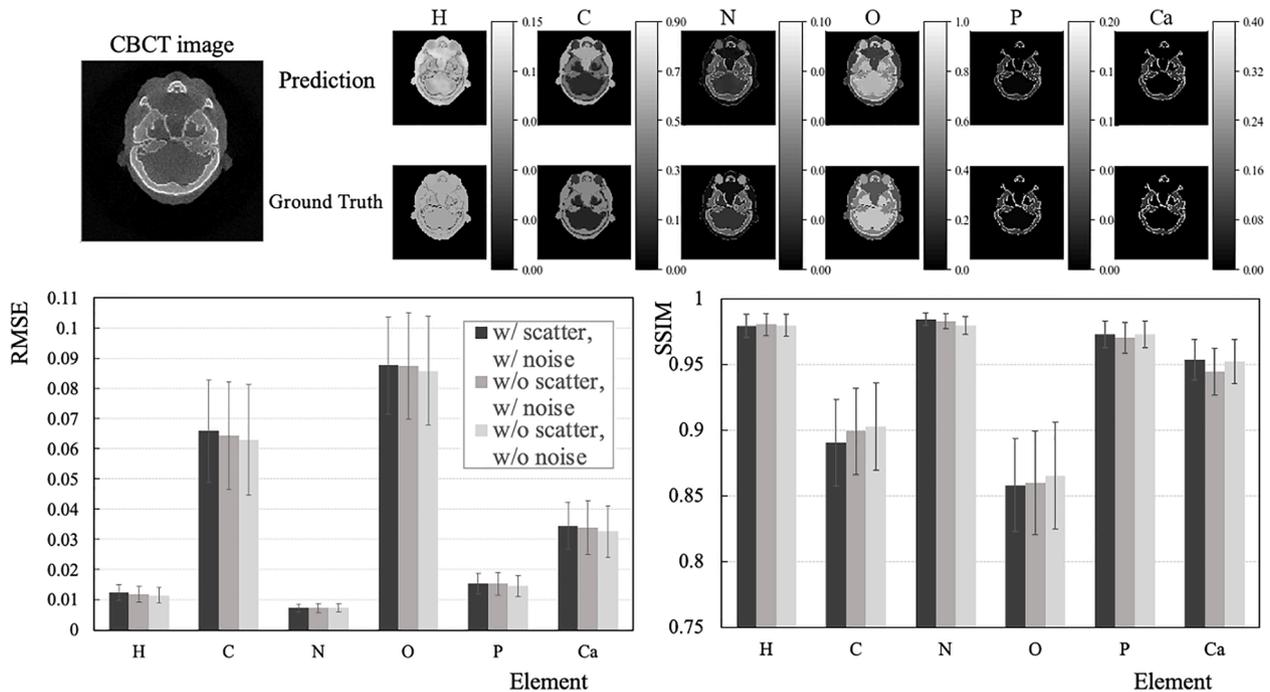
Fig. 7. (Top) Example of predicted distribution for 6 major elements in human body from a virtual cone-beam computed tomography (CBCT) image reconstructed with the scatter and noise. (Bottom) Root mean squared error (RMSE, left) and structural similarity index measure (SSIM, right) of the elemental material decomposition predicted for each element with the test data (6 human phantoms). "w/ scatter" and "w/ noise" imply the conditions with scatter and noise, respectively, whereas "w/o scatter" and "w/o noise" imply the scatter- and the noise-free conditions, respectively.
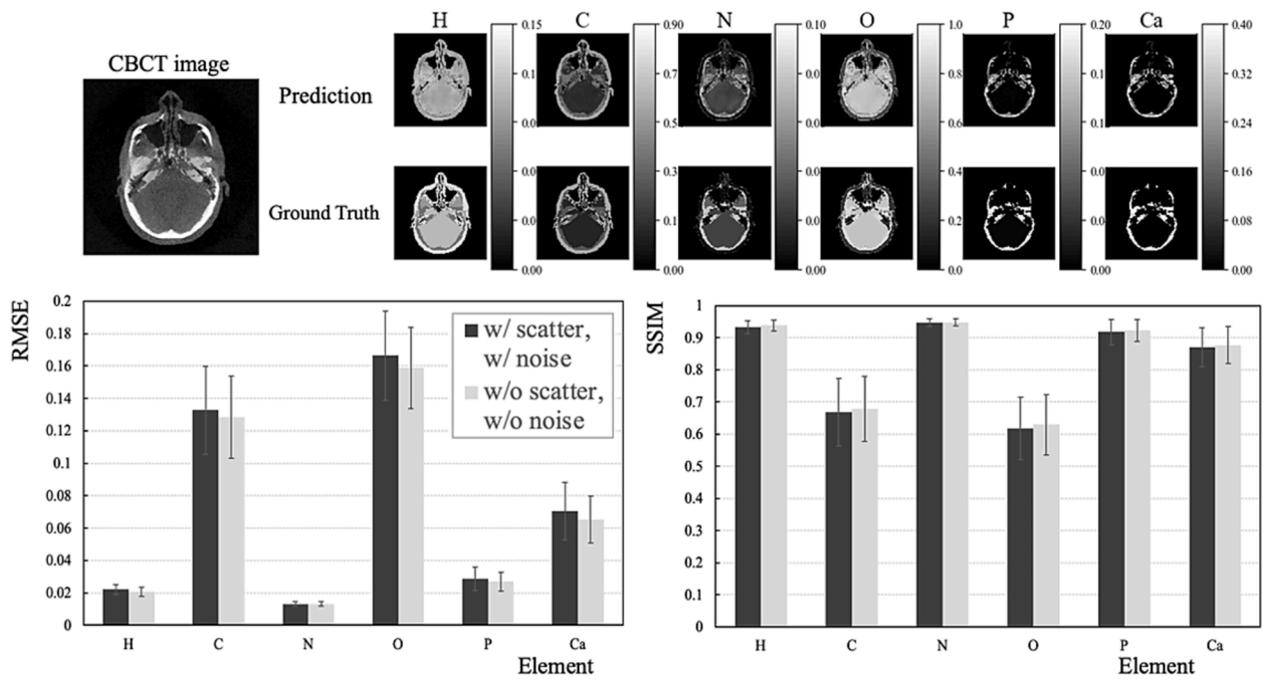
Fig. 8. Same as in Fig. 7, except for the testing data where the human phantom deformed from the ICRP phantom into the real 3D CT image, which is independent of our human library.

amount of the noise varies based on the imaging protocol such as the tube current and exposure time. Moreover, the CBCT images with a large noise can be useful as the training data to develop a robust model in the DL model.

3.2. Evaluation of the material decomposition using CBCT

The RMSE and SSIM of the EMD predicted for each element with the test data (6 human phantoms) are shown in Fig. 7, where the "w/o

scatter" and "w/o noise" imply the scatter-free and the noise-free conditions, respectively, were used as an input image generation in the development of the EMD models. This result shows that the predicted distribution in H, P, N, and Ca is consistent with the ground truth compared to that in C and O. The difficulty in predicting the C and O distributions may be attributed to the large deviation of C and O in biological tissues. Thus, one of the difficult tasks in EMD for the human body is to distinguish between anatomies with similar attenuation coefficients (electron densities or CT values) but different material
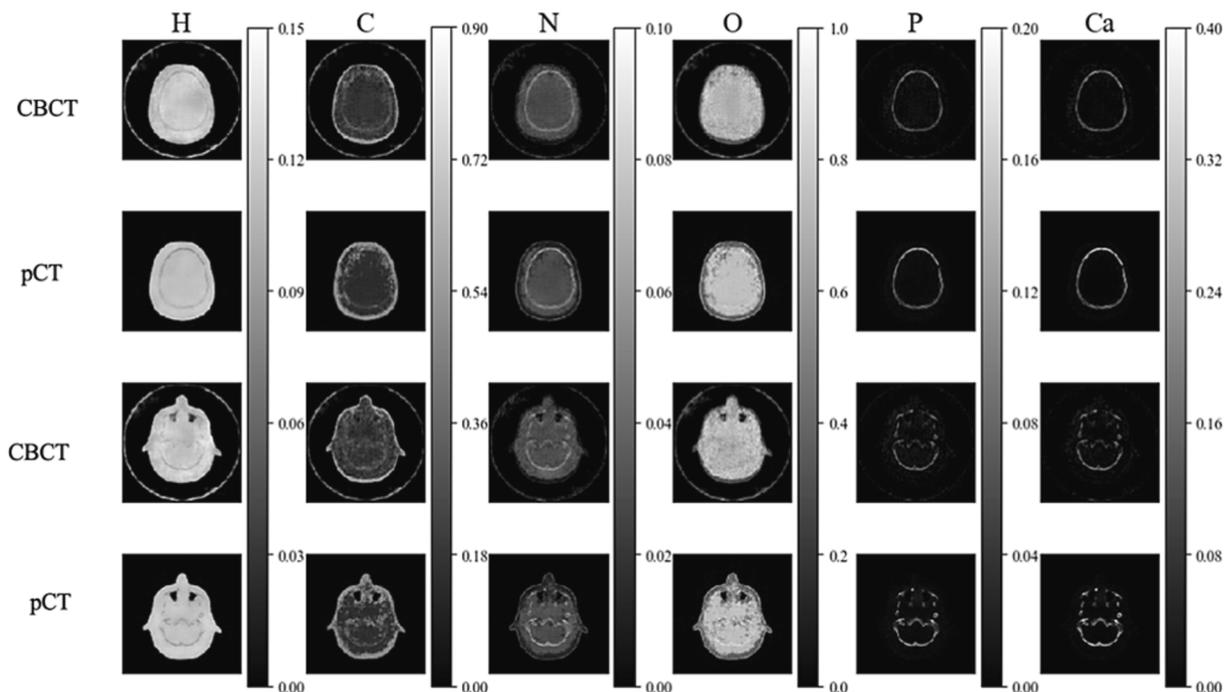

Fig. 9. Example of predicted distribution for 6 major elements for the cone-beam computed tomography (CBCT) images and the planning CT (pCT) images of Rando phantom acquired by real CBCT machine (Elekta Synergy) and real planning CT machine (Canon Aquilion LB).

weights. Nevertheless, it can be concluded that the EMD model can comprehensively predict the elemental distributions. Fig. 8 shows the RMSE and SSIM obtained for the human phantom deformed from the ICRP phantom into the real 3D CT image, which is independent of our human library. The results indicate that the EMD model developed with our human library can predict the elemental distribution well and exhibits a generalization ability. However, the quantitative accuracy was decreased compared with the test results presented in Fig. 7, implying the need for variety in the human library to avoid overlearning.

As evident from Figs. 7 and 8, the prediction using CBCT images without the scatter and noise condition is slightly better than that with the scatter and noise condition. Thus, the CBCT image degradation owing to the scattered X-ray and the statistical noise has effect on the prediction accuracy in the DL-based EMD model. Further, the present result showed that the SSIM in scatter- and noise-free condition is significantly different from that in scatter and noise condition for C and O distribution (paired t-test, p < 0.03). In addition, the noise reduction as well as the scattered X-ray removal in projection images measured by FPD is significant, although the impact of this effect is minimal.

3.3. Material decomposition using real CBCT system

To investigate the feasibility of this study, the proposed EMD approach was applied to the experimental kV CBCT (Elekta XVI system). Fig. 9 shows the elemental distribution with the Rando phantom, where the EMD results are compared with those predicted using the planning CT image. As evident, the EMD from the CBCT images produced

comparable results with that of the planning CT images. Fig. 10 shows the elemental distribution for an HN cancer patient. Even under different CT scan protocols, results produced the similar distributions. In addition, they revealed that P and Ca were extracted clearly in the high-density materials. Moreover, the enhanced images indicate that the bone and the soft tissue are distinguishable mainly from the P and Ca distributions. As shown in the previous subsection, it is considered that the present EMD model reconstructed the elemental distribution mainly using the structural information. Because of this property, the reconstructed result appears to be reasonable. In contrast, there is no reliability on the EMD in the couch as well as the fixture used for the HN patient as no training data included these materials.

4. Discussion

The combination of computational human phantoms and a virtual CBCT simulator can provide a breakthrough solution for accurate data collection in medical image informatics. In addition, the proposed simulator can provide a rigorous association with information that has never been obtained in clinical data such as a material distribution. In this study, we demonstrated this utility by developing the DL-based EMD model with the proposed simulator, where the model was trained and validated via reconstructed CBCT images from our phantom library. It was also shown that the EMD model was insensitive to the scattered X-ray and noise, and produced reasonable material distributions for a real patient CBCT. In addition, predictions of the elemental material distributions were performed in 30 s using GPU, implying that the proposed
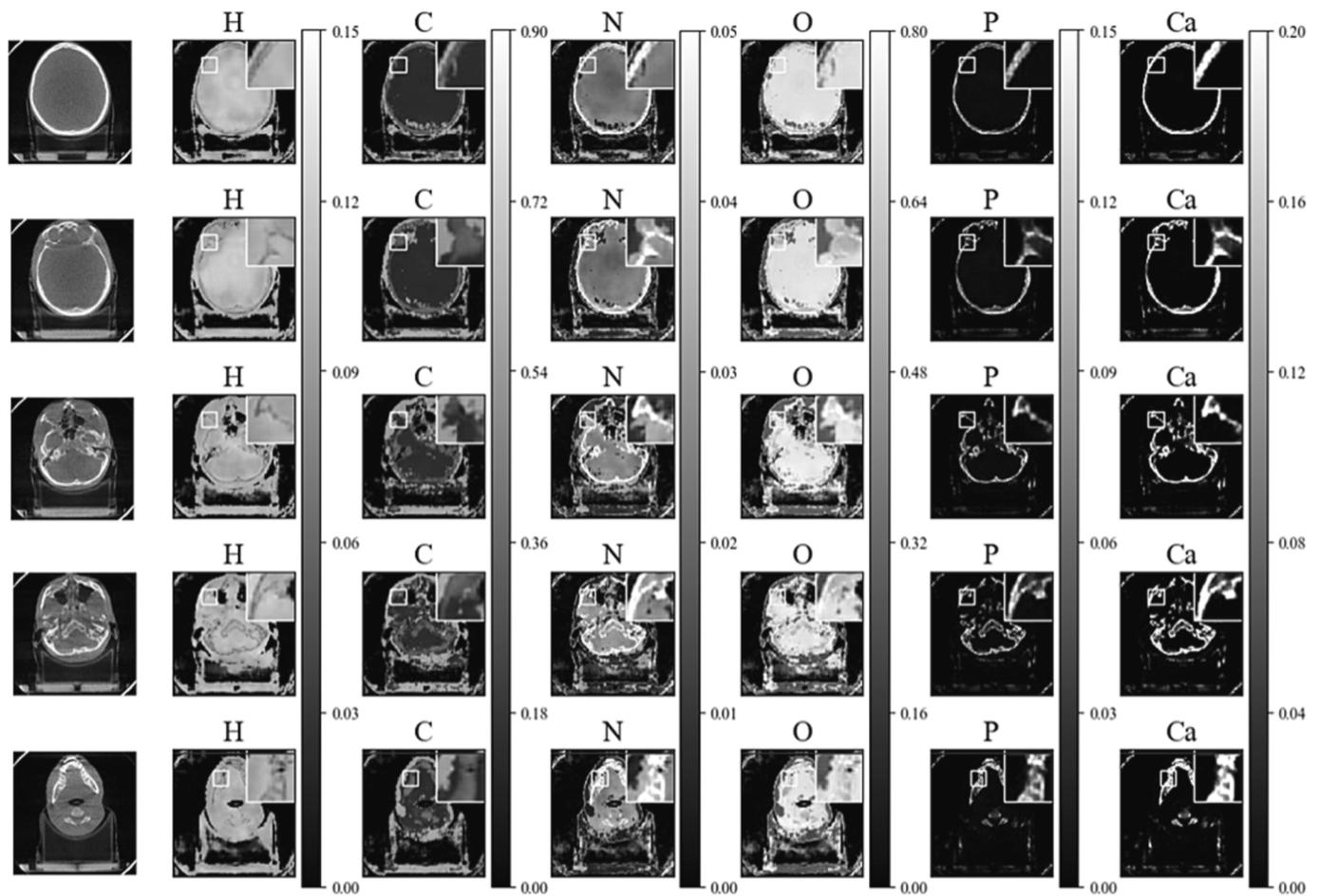

Fig. 10. Example of predicted distribution for 6 major elements for the cone-beam computed tomography (CBCT) images of a patient acquired by real CBCT machine (Elekta Synergy), where 100 kVp, 10 mA, and 10 ms parameters were employed. The enhanced images show the boundary between the bone and soft tissue.

model is applicable to the ART with CBCT acquired before treatment. The importance of elemental information in ART is particularly emphasized in hadron therapy. The impact of fragmentation on dose calculation is indeed significant, as exemplified in carbon ion therapy, where approximately 15% of the physical dose at the Bragg peak is attributed to fragments [41–44]. Accurate elemental information is essential for such dose calculations, highlighting the crucial role it plays in evaluating fragmentation and assessing the subsequent distribution of dose (biological dose) distribution.

In this study, statistics of the human population was followed in the creation of the human phantom library. The construction of realistic human phantoms models is highly desired in the fields of radiological diagnosis, radiation therapy, and radiation protection, which include the investigation of the novel CT device and/or reconstruction algorithm, quality assurance and quality control in the delivered dose, impact of the cosmic rays in the space, etc. Our human library was extended from the ICRP reference phantoms by referring to the methods in a previous study, wherein the diversity of anthropometric characteristics was presented [25]. In addition, anatomical diversities were developed by considering the deformation of the brain and the deviation in the anatomical density and the weight fraction of elemental components. Consequently, we generated more than 36 phantoms with parameters such as gender, age, height, and weight. In addition to the ICRP human phantom, four-dimensional (4D) extended cardiac-torso based on National Library of Medicine data [45] is among the broadly utilized phantom libraries. The first human 4D phantom library was developed by Segars et al. [46]. These libraries were extracted using real CT data of normal patients having different BMIs. Another difference between these human libraries and that proposed in this study was the introduction of the deviation in the density and elemental composition fraction at each medium to develop the quantitative EMD imaging.

This study successfully demonstrated the feasibility of material decomposition using the virtual CBCT system. We specifically investigated the impact of scattered X-rays and statistical noise on the prediction accuracy of the DL-based EMD model. While the predictions using CBCT images with scatter and noise removal exhibited statistically superior performance compared to those including scatter and noise, the observed difference was not substantial (as shown in Figs. 7 and 8). It is widely recognized that scattered X-rays can degrade the image quality in CBCT imaging. However, the utilization of the DL model, which was trained on anatomical shapes, may have mitigated the adverse effects, resulting in accurate predictions.

The virtual CBCT system can be prospective in various applications other than that in the MD in the medical imaging research. For example, an anatomical segmentation model can be developed by a small extension of the proposed model. In the current segmentation approach, the contours (boundaries) of each anatomy are prepared from the data in radiotherapy patients, where the segmentation is performed by hand. Such contours include an inter-observer error and it is inevitable in the development of the segmentation model. However, we can develop a model without such the error by using the virtual CBCT system, because the anatomical boundaries are rigorous owing to the use of the human library. Similarly, metal-artifact reduction may be applied in the reconstructed domain by including the metal database in the human library. Moreover, the novel reconstruction algorithm and the device development can be investigated using the virtual system. In this study, the impact of the scattered X-ray and the electrical noise in the MD model development was investigated. Thus, the other artifacts observed

in the CBCT system such as the ghosting artifact [47] can be investigated by extending our virtual system.

Through this study, we demonstrated the usefulness of the generative image model by the virtual CBCT system in DL for quantitative EMD imaging. However, several limitations still exist. Firstly, the variety of body shapes in our library for the HN region is insufficient. This is due to the limited number of mediums that can be well-deformed, such as fat, at different ages. Additionally, the facing direction was unchanged from the ICRP phantom basis. We developed 36 phantoms in the library, which may not provide a comprehensive representation of the HN region. To address this issue, a combination with the DIR approach may help generate different types of phantoms. Secondly, the virtual CBCT system used in this study still possesses certain simplified features [48–53], such as disregarding the energy dependence on FPD, bowtie filter, and panel shift. Moreover, phantom motion during the acquisition of projection images was also not considered. On the other hand, the virtual CBCT system can be adapted to incorporate real situations, and more realistic features are expected to be included in the near future. Thirdly, apart from data generation, the DL model used in this study has limitations, especially in terms of interpretability. It is typically considered a black box, making it difficult to extract meaningful explanations for prediction results. Therefore, careful validation is required in medical applications. These limitations should be addressed in future studies.

5. Conclusions

We developed a virtual CBCT system with a head-and-neck human phantom library, allowing for the inclusion of an arbitral X-ray spectrum and generating the cone-beam projection image with the scattered X-ray through the use of a DL model and statistical noise. The reconstructed CBCT images are associated with the material information, and in this study, the material decomposition was performed to demonstrate the utility of the developed virtual CBCT system.

The results of our investigation demonstrate the potential of accurately estimating elemental compositions using the U-net model with virtual CBCT images input labeled with material information, as it is robust against scattered X-rays and noise, though further validation with various CBCT data is necessary to ensure reliability. Our method offers advantages in estimating elemental distributions that are typically challenging to obtain in a clinical setting. It holds promising applications in ART, particularly in the context of hadron therapy.

Declaration of Competing Interest

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

Appendix

In this Appendix, we summarize the development of our human library. Table A1 presents the human phantoms utilized in this study. The procedure for creating the human library is outlined below:

Table A1
Specifics of all human phantoms used in this study.

Base	ICRP 110	TCIA
Element	H, C, N, O, P, Ca	H, C, N, O, P, Ca
Resolution (mm^3)	$1 \times 1 \times 2$	$1 \times 1 \times 2.4996$
Slice(s)	120 (female) or 140 (male)	80
No. of phantoms	36	1

1. Generating high-resolution versions of the ICRP 110 human phantoms for both female and male subjects.

2. Scaling the phantoms using Equations (1) and (2), where the height and weight values are obtained from the 25%, 50%, and 75% percentiles at ages 60, 70, and 80 using the NHANES database [27]. During this process, a non-rigid deformation was applied to the brain anatomy, where the DVF was artificially generated; First, the binding box of the brain was defined and the grids for the DVF were placed three-dimensionally at an equal spacing of 8 mm. Then, the DVF in the nearest grid to the surface of the binding box with $\vec{r}$ was produced as

$$\vec{T} = \frac{\vec{t}}{\left| \vec{r} - \vec{r}_0 \right|^2},$$

where $\vec{r}_0$ is the centroid of the brain. Further, the magnitude of the deformation was controlled by $\vec{t}$, which was sampled from the normal distribution. Furthermore, the deformed brain was scaled by the value of the power of one third of the ratio between the weight of ICRP 110 and the weight presented in a previous study [25] and rotated three-dimensionally within a range of $\pm2°$. However, a part of the generated brain can overlap with the other anatomies and in this case, the priority of the brain was considered as lower than the others. In this procedure, 18 female and 18 male phantoms were generated.

3. The density and elemental fractions in the anatomy were subsequently sampled from a Gaussian distribution with standard deviations of 5% and 10% respectively, based on the ICRP reference data. These values were then assigned to the phantoms mentioned above.

4. In addition to the phantoms derived from the ICRP 110 human library, additional human phantoms were created using 3D DIR. This process involved registering the ICRP 110 human phantom with real head and neck CT images obtained from The Cancer Imaging Archive (TCIA). An example of this process is illustrated in Fig. A1.
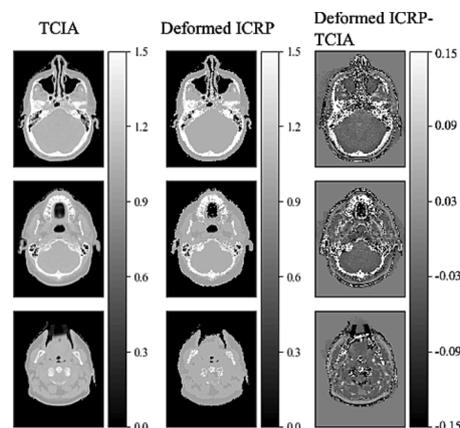
Fig. A1. Examples of targeted TCIA images (first column), deformed ICRP 110 reference phantom (second column), and difference between Deformed ICRP and TCIA images (third column).

References

[1] Unkelbach J, Bortfeld T, Martin BC, Soukup M. Reducing the sensitivity of IMPT treatment plans to setup errors and range uncertainties via probabilistic treatment planning. Med Phys 2009;36:149–63. https://doi.org/10.1118/1.3021139.

[2] Barker Jr JL, Garden AS, Ang KK, O'Daniel JC, Wang H, Court LE, et al. Quantification of volumetric and geometric changes occuring during fractionated radiotherapy for head-and-neck cancer using an integrated CT/linear accelerator system. Int J Radiat Oncol Biol Phys 2004;59:960–70. https://doi.org/10.1016/j.ijrobp.2003.12.024.

[3] Shinde P, Jadhav A, Shankar V, Gupta KK, Dhoble NS, Dhoble SJ. Evaluation of kV-CBCT based 3D dose calculation accuracy and its validation using delivery fluence derived dose metrics in head and neck cancer. Phys Med 2002;96:32–45. https://doi.org/10.1016/j.ejmp.2022.02.014.

[4] Haga A, Nakagawa K, Shiraishi K, Itoh S, Terahara A, Yamashita H, et al. Quality assurance of volumetric modulated arc therapy using Elekta Synergy. Acta Oncol 2009;48(8):1193–7.

[5] Posiewnik M, Piotrowski T. A review of cone-beam CT applications for adaptive radiotherapy of prostate cancer. Phys Med 2019;59:13–21. https://doi.org/10.1016/j.ejmp.2019.02.014.

[6] Cho S, Lim S, Kim C, Wi S, Kwon T, Youn WS, et al. Enhancement of soft-tissue contrast in cone-beam CT using an anti-scatter grid with a sparse sampling approach. Phys Med 2020;70:1–9.

[7] Wang J, Mao W, Solberg T. Scatter correction for cone-beam computed tomography using moving blocker strips: a preliminary study. Med Phys 2010;37:5792–800. https://doi.org/10.1118/1.3495819.

[8] Yao W, Leszczynski KW. An analytical approach to estimating the first order X-ray scatter in heterogeneous medium. Med Phys 2009;36:3145–56. https://doi.org/10.1118/1.3152114.

[9] Bootsma GJ, Verhaegen F, Jaffray DA. Efficient scatter distribution estimation and correction in CBCT using concurrent Monte Carlo fitting. Med Phys 2015;42:54–68. https://doi.org/10.1118/1.4903260.

[10] Kida S, Nakamoto T, Nakano M, Nawa K, Haga A, Kotoku J, et al. Cone beam computed tomography image quality improvement using a deep convolutional neural network. Cureus 2018;10:e2548.

[11] Chen L, Liang X, Shen C, Jiang S, Wang J. Synthetic CT generation from CBCT images via deep learning. Med Phys 2020;47:1115–25. https://doi.org/10.1002/mp.13978.

[12] Lalonde A, Winey B, Verburg J, Paganetti H, Sharp GC. Evaluation of CBCT scatter correction using deep convolutional neural networks for head and neck adaptive proton therapy. Phys Med Biol 2020;65(24):245022.

[13] Rong Y, Smilowitz J, Tewatia D, Tomé WA, Paliwal B. Dose calculation on kV cone beam CT images: an investigation of the HU-density conversion stability and dose

accuracy using the site-specific calibration. Med Dosim 2010;35:195–207. https://doi.org/10.1016/j.meddos.2009.06.001.

[14] Kaplan LP, Elstrøm UV, Møller DS, Hoffmann L. Cone beam CT based dose calculation in the thorax region. Phys Imaging Radiat Oncol 2018;7:45–50. https://doi.org/10.1016/j.phro.2018.09.001.

[15] Sato Y-H, Sakata D, Bolst D, Simpson EC, Guatelli S, Haga A. Development of a more accurate Geant4 quantum molecular dynamics model for hadron therapy. Phys Med Biol 2022;67(22):225001.

[16] Durante M, Paganetti H. Nuclear physics in particle therapy: a review. Rep Prog Phys 2016;79(9):096702.

[17] Wohlfahrt P, Richter C. Status and innovations in pre-treatment CT imaging for proton therapy. Br J Radiol 2020;93(1107):20190590.

[18] Badal A, Badano A. Accelerating Monte Carlo simulations of photon transport in a voxelized geometry using a massively parallel graphics processing unit. Med Phys 2009;36:4878–80. https://doi.org/10.1118/1.3231824.

[19] Shen C, Li B, Lou Y, Yang M, Zhou L, Jia X. Multi-energy element-resolved cone beam CT (MEER-CBCT) realized on a conventional CBCT platform. Med Phys 2018; 45:4461–70. https://doi.org/10.1002/mp.13169.

[20] Zhu J, Su T, Zhang X, Yang J, Mi D, Zhang Y, et al. Feasibility study of three-material decomposition in dual-energy cone-beam CT imaging with deep learning. Phys Med Biol 2022;67(14):145012.

[21] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A, et al. GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;2018:68. https://doi.org/10.3322/caac.21492.

[22] Ratko TA, Douglas GW, de Souza JA, Belinson SE, Aronson N. Radiotherapy Treatments for Head and Neck Cancer Update. *Rockville (MD): Agency for Healthcare Research and Quality (US)*. 2014;15-EHC001-EF.

[23] Menzel HG, Clement C, DeLuca P. ICRP Publication 110. Realistic reference phantoms: an ICRP/ICRU joint effort, A report of adult reference computational phantoms. Ann ICRP 2009;39:1–164. https://doi.org/10.1016/j.icrp.2009.09.001.

[24] Klein S, Staring M, Murphy K, Viergever MA, Pluim JP. elastix: a toolbox for intensity-based medical image registration. IEEE Trans Med Imaging 2010;29: 196–205. https://doi.org/10.1109/tmi.2009.2035616.

[25] Akhavanallaf A, Xie T, Zaidi H. Development of a library of adult computational phantoms based on anthropometric indexes. IEEE Trans Radiat PlasmaMed Sci 2019;3:65–75. https://doi.org/10.1109/TRPMS.2018.2816072.

[26] Bushby KM, Cole T, Mathews JN, Goodship JA. Centiles for adult head circumference. Arch Dis Childh 1992;67:1286–7. https://doi.org/10.1136/adc.67.10.1286.

[27] Fryar CD, Carroll MD, Gu Q, Afful J, Ogden CL. Anthropometric reference data for children and adults: United States, 2015–2018. National Center for Health Statistics. *Vital Health Stat*. 2021;3:1-44.

[28] Looker AC, Borrud LG, Hughes JP, Fan B, Shepherd JA, Sherman M. Total body bone area, bone mineral content, and bone mineral density for individuals aged 8 years and over: United States, 1999–2006. Vital Health Stat 2013;11:1–78.

[29] Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, et al. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. J Digit Imaging 2013;26(6):1045–57.

[30] Spezi E, Downes P, Radu E, Jarvis R. Monte Carlo simulation of an X-ray volume imaging cone beam CT unit. Med Phys 2009;36:127–36. https://doi.org/10.1118/1.3031113.

[31] Li KW, Fujiwara D, Haga A, Liu H, Geng LS. Physical density estimations of single- and dual-energy CT using material-based forward projection algorithm: A simulation study. Br J Radiol 2021;94:20201236. https://doi.org/10.1259/bjr.20201236.

[32] Li K-W, Fujiwara D, Haga A, Liu H, Geng L-S. kV–kV and kV–MV DECT based estimation of proton stopping power ratio - A simulation study. Physcia Medica 2021;89:182–92. https://doi.org/10.1016/j.ejmp.2021.07.038.

[33] Hasegawa Yu, Haga A, Sakata D, Kanazawa Y, Tominaga M, Sasaki M, et al. Estimation of X-ray energy spectrum of cone-beam computed tomography scanner using percentage depth dose measurements and machine learning approach. J Phys Soc Jpn 2021;90(7):074801.

[34] Newville M. XrayDB ver.4.47, 2021. Available from: https://xraypy.github.io/XrayDB/ or https://github.com/xraypy/XrayDB.

[35] Maier J, Eulig E, Vöth T, Knaup M, Kuntz J, Sawall S, et al. Real-time scatter estimation for medical CT using the deep scatter estimation (DSE) method and robustness analysis with respect to different anatomies dose levels tube voltages and data truncation. Med Phys 2019;46:238–49. https://doi.org/10.1002/mp.13274.

[36] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks, *Proc. 13th AISTATS.* 2010;9:249-256.

[37] Feldkamp LA, Davis LC, Kress JW. Practical cone-beam algorithm. J Opt Soc Am A 1984;1:612–9. https://doi.org/10.1364/JOSAA.1.000612.

[38] Webb S. A modified convolution reconstruction technique for divergent beams. Phys Med Biol 1982;37:419–23. https://doi.org/10.1088/0031-9155/27/3/009.

[39] Fujiwara D, Shimomura T, Zhao W, Li K-W, Haga A, Geng L-S. Virtual computed-tomography system for deep-learning-based material decomposition. Phys Med Biol 2022;67(15):155008.

[40] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 2004;13:600–12. https://doi.org/10.1109/TIP.2003.819861.

[41] Schardt D, Schall I, Geissel H, Irnich H, Kraft G, Magel A, et al. Nuclear fragmentation of high-energy heavy-ion beams in water. Adv Space Res 1996;17 (2):87–94.

[42] Matsufuji N, Fukumura A, Komori M, Kanai T, Kohno T. Influence of fragment reaction of relativistic heavy charged particles on heavy-ion radiotherapy. Phys Med Biol 2003;48:1605–23. https://doi.org/10.1088/0031-9155/48/11/309.

[43] Haettner E, Iwase H, Schardt D. Experimental fragmentation studies with 12C therapy beams. Radiat Prot Dosimetry 2006;122(1–4):485–7. https://doi.org/10.1093/rpd/ncl402.

[44] Bolst D, Cirrone GAP, Cuttone G, Folger G, Incerti S, Ivanchenko V, et al. Validation of Geant4 fragmentation for heavy ion therapy. Nucl Instrum Methods Phys Res A 2017;869:68–75.

[45] Segars WP, Mendonca S, Grimes J, Sturgeon G, Tsui BMW. 4D XCAT phantom for multimodality imaging research. Med Phys 2010;37:4902–15. https://doi.org/10.1118/1.3480985.

[46] Segars WP, Bond J, Frush J, Hon S, Eckersley C, Williams CH, et al. Population of anatomically variable 4D XCAT adult phantoms for imaging research and optimization. Med Phys 2013;40(4):043701.

[47] Schulze R, Heil U, Groβ D, Bruellmann DD, Dranischnikow E, Schwanecke U, et al. Artefacts in CBCT: a review. Dentomaxillofac Radiol 2011;40(5):265–73.

[48] van der Heyden B, Schyns LEJR, Podesta M, Vaniqui A, Almeida IP, Landry G, et al. VOXSI: A voxelized single- and dual-energy CT scenario generator for quantitative imaging. Phys Imaging Radiat Oncol 2018;6:47–52.

[49] O'Connell J, Lindsay C, Bazalova-Carter M. Experimental validation of Fastcat kV and MV cone beam CT (CBCT) simulator. Med Phys 2021;48:6869–80. https://doi.org/10.1002/mp.15243.

[50] Abadi E, Harrawood B, Sharma S, Kapadia A, Segars WP, Samei E. DukeSim: A realistic, rapid, and scanner-specific simulation framework in computed tomography. IEEE Trans Med Imaging 2019;38:1457–65. https://doi.org/10.1109/tmi.2018.2886530.

[51] Wu M, FitzGerald P, Zhang J, Segars WP, Yu H, Xu Y, et al. XCIST-an open access X-ray/CT simulation toolkit. Phys Med Biol 2022;67(19):194002.

[52] van Aarle W, Palenstijn WJ, Cant J, Janssens E, Bleichrodt F, Dabravolski A, et al. Fast and flexible X-ray tomography using the ASTRA toolbox. Opt Express 2016;24 (22):25129.

[53] Jia X, Yan H, Cervino L, Folkerts M, Jiang SB. A GPU tool for efficient, accurate, and realistic simulation of cone beam CT projections. Med Phys 2012;39:7368–78. https://doi.org/10.1118/1.4766436.