

RESEARCH

Open Access



# Deep learning method for prediction of patient-specific dose distribution in breast cancer

Sang Hee Ahn<sup>1,2</sup>, EunSook Kim<sup>1</sup>, Chankyu Kim<sup>1</sup>, Wonjoong Cheon<sup>1</sup>, Myeongsoo Kim<sup>1</sup>, Se Byeong Lee<sup>1</sup>, Young Kyung Lim<sup>1</sup>, Haksoo Kim<sup>1</sup>, Dongho Shin<sup>1</sup>, Dae Yong Kim<sup>1</sup> and Jong Hwi Jeong<sup>1\*</sup>

## Abstract

**Background:** Patient-specific dose prediction improves the efficiency and quality of radiation treatment planning and reduces the time required to find the optimal plan. In this study, a patient-specific dose prediction model was developed for a left-sided breast clinical case using deep learning, and its performance was compared with that of conventional knowledge-based planning using RapidPlan™.

**Methods:** Patient-specific dose prediction was performed using a contour image of the planning target volume (PTV) and organs at risk (OARs) with a U-net-based modified dose prediction neural network. A database of 50 volumetric modulated arc therapy (VMAT) plans for left-sided breast cancer patients was utilized to produce training and validation datasets. The dose prediction deep neural network (DpNet) feature weights of the previously learned convolution layers were applied to the test on a cohort of 10 test sets. With the same patient data set, dose prediction was performed for the 10 test sets after training in RapidPlan. The 3D dose distribution, absolute dose difference error, dose-volume histogram, 2D gamma index, and iso-dose dice similarity coefficient were used for quantitative evaluation of the dose prediction.

**Results:** The mean absolute error (MAE) and one standard deviation (SD) between the clinical and deep learning dose prediction models were  $0.02 \pm 0.04\%$ ,  $0.01 \pm 0.83\%$ ,  $0.16 \pm 0.82\%$ ,  $0.52 \pm 0.97$ ,  $-0.88 \pm 1.83\%$ ,  $-1.16 \pm 2.58\%$ , and  $-0.97 \pm 1.73\%$  for  $D_{95\%}$ ,  $D_{mean}$  in the PTV, and the OARs of the body, left breast, heart, left lung, and right lung, respectively, and those measured between the clinical and RapidPlan dose prediction models were  $0.02 \pm 0.14\%$ ,  $0.87 \pm 0.63\%$ ,  $-0.29 \pm 0.98\%$ ,  $1.30 \pm 0.86\%$ ,  $-0.32 \pm 1.10\%$ ,  $0.12 \pm 2.13\%$ , and  $-1.74 \pm 1.79$ , respectively.

**Conclusions:** In this study, a deep learning method for dose prediction was developed and was demonstrated to accurately predict patient-specific doses for left-sided breast cancer. Using the deep learning framework, the efficiency and accuracy of the dose prediction were compared to those of RapidPlan. The doses predicted by deep learning were superior to the results of the RapidPlan-generated VMAT plan.

**Keywords:** Deep learning, Dose prediction, Volumetric modulated arc therapy (VMAT), Knowledge-based planning (KBP), RapidPlan™

\*Correspondence: [jonghwi@ncc.re.kr](mailto:jonghwi@ncc.re.kr)

<sup>1</sup> Department of Radiation Oncology, Proton Therapy Center, National Cancer Center, 323, Ilsan-ro, Ilsandong-gu Gyeonggi-do, Goyang-si 10408, South Korea

Full list of author information is available at the end of the article



© The Author(s) 2021. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

As radiation therapy treatment technology is advancing, the treatment outcomes of cancer patients are gradually improving. Recently, advanced treatment modalities, such as intensity modulation radiation therapy (IMRT) and volumetric arc therapy (VMAT), have been applied to deliver higher doses to tumor areas, while reducing the therapeutic doses to normal organs compared to conventional 3D conformal radiation therapy. However, the ability of these advanced treatment methods to produce an optimal plan varies according to the experience of the planner, and a planning time-consuming task must be repeated until the treatment goal is reached. Therefore, studies [1–6] have been conducted to improve treatment planning efficiency and quality, while reducing planning time and effort by using knowledge-based techniques for dose prediction in radiotherapy.

Currently, commercial software is available in the form of RapidPlan (version 13.6, Varian Oncology Systems, Palo Alto, CA, USA) [7–9]. A knowledge-based planning (KBP) model is generated using a previous, clinically approved treatment plan data-based regression analysis as the dose-volume histogram (DVH) estimation algorithm of RapidPlan. For a new patient, the most similar treatment plan is provided within the estimated DVH model.

However, dose prediction using the KBP model has two limitations. First, it is difficult to include all characteristics of the inherent organ structure depending on the patient. Second, the DVH does not reflect the spatial dose distribution. It is possible to derive an unacceptable plan according to the dose distribution around an important organ at risk (OAR). If patient-specific dose prediction were possible while compensating for the limitations of the KBP model, the workload in clinical practice would be reduced.

Deep learning methods have proven to be effective in various fields, such as automatic segmentation [10–12], image registration [13–15], respiratory motion prediction [16–18], and toxicity prediction [19–21]. Studies on dose prediction using deep learning have also been reported.

Deep learning methods based on dose prediction of IMRT plans have been utilized for head and neck [22, 23],

rectal [24], prostate [25, 26], and lung [27] cancer cases. In addition, dose prediction using VMAT plans has been performed for head and neck [28], rectal [29], and prostate [30] cancer. There have also been other treatment techniques such as 3D dose prediction for head and neck cancer treatment using helical tomotherapy [31].

In those studies, evaluation of the test cases was conducted after learning a deep neural network using 2D or 3D computed tomography (CT) images and patient anatomical information. However, when evaluating the accuracies of the deep learning models, comparisons were not made under the conditions used in the commercial models; rather, only the DVH, mean dose, maximum dose, and gamma index [32] were employed for quantitative evaluation of the clinically accepted plan.

In this study, the doses predicted by developed deep learning model using only the anatomy information of left-sided breast cancer patients treated with VMAT were compared with the results predicted using RapidPlan under the same clinically acceptable conditions.

## Methods

### Patients and treatment planning

Fifty-five patients with left-sided breast cancer diagnosed at the National Cancer Center in South Korea in 2018 and 2019 were included in this study. The characteristics of the patients are listed in Table 1. All patients were staged according to the American Joint Committee on Cancer (AJCC) staging system [33].

Manual planning and treatment were performed for all patients with two coplanar VMAT arcs with a prescription of 4320 cGy in 16 fractions, a photon energy of 6 MV beam, gantry angles of 165°–290°, and collimator angles of 30° and 330°. The entire breast was included, while there was no nodal involvement.

All treatments were planned using the Varian Eclipse treatment planning system (version 13.6, Varian Oncology Systems, Palo Alto, CA, USA) with the analytical anisotropic algorithm (AAA). Table 2 shows the planning goals for the PTV, heart, left lung, right lung, and right breast.

The clinical treatment plan was optimized to ensure the following OAR constraints: no more than 5% of the heart received > 20 Gy (Volume 5 [%] ≤ 20 Gy); no more than 30% of the heart received > 10 Gy (Volume 30

**Table 1** Patient characteristics in this study

Patients	Female	Average age	Average left breast volume [cm <sup>3</sup> ]	Stage				
				I	II	III	IV	N/X
Training set	45	51	438.15	35	6	-	-	4
Testing set	10	56	491.51	7	3	-	-	-

**Table 2** Left-sided breast cancer clinical treatment planning goals

Target and normal organ structure	Planning goals
Planning target volume (PTV)	Dose 95 [%] = 100 Dose [max] ≤ 49.68 Gy
Heart	Dose [mean] ≤ 4 Gy Volume 5 [%] ≤ 20 Gy Volume 30 [%] ≤ 10 Gy
Lung (left, right)	Volume 15 [%] ≤ 20 Gy Volume 35 [%] ≤ 10 Gy Volume 50 [%] ≤ 5 Gy
Right breast	Dose [max] ≤ 5 Gy

\*Volume [% of total volume]

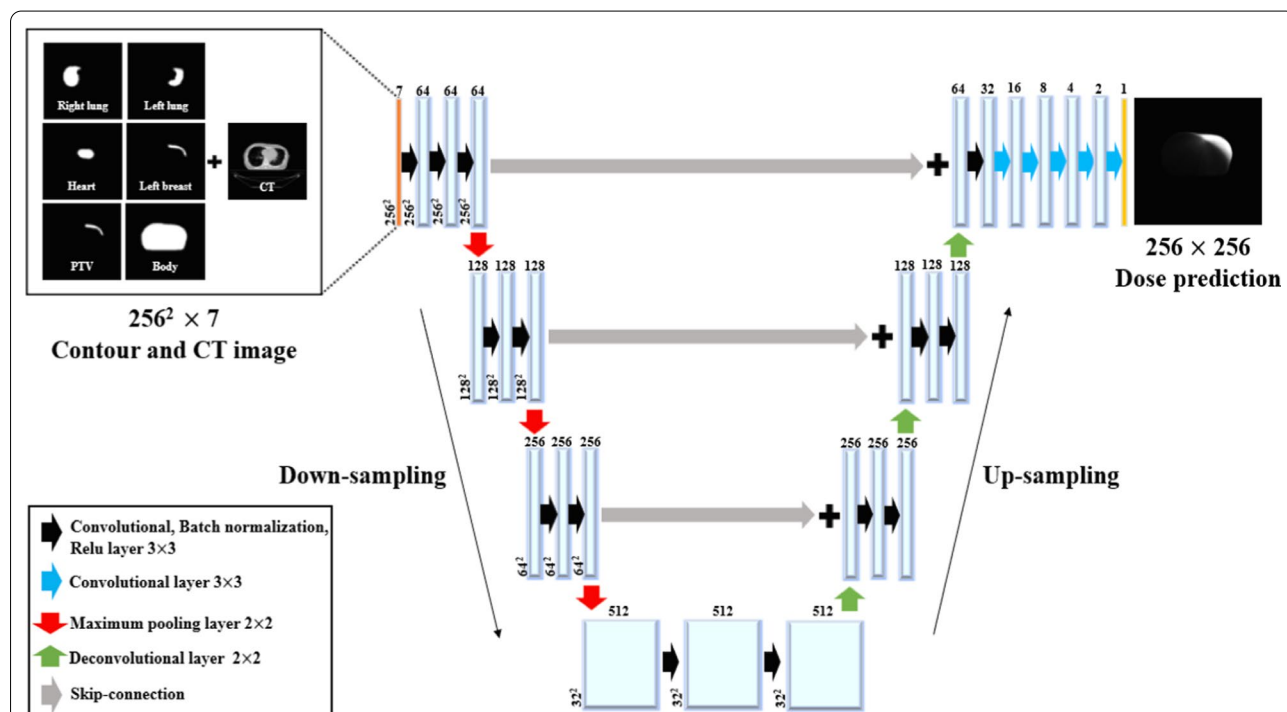
[%] ≤ 10 Gy); no more than 4 Gy of the heart mean dose received; no more than 15% of the lungs received > 20 Gy (Volume 15 [%] ≤ 20 Gy); no more than 35% of the lungs received > 10 Gy (Volume 35 [%] ≤ 10 Gy); no more than 50% of the lungs received > 5 Gy (Volume 50 [%] ≤ 5 Gy); no more than 5 Gy of the right breast max dose received. All the clinical confirmation plans were normalized such that 95% of the PTV received 100% of the prescription dose. For each plan, the contours of the PTV and OARs were determined by experienced physicians, and the dose distribution was confirmed by experienced physicians

and physicists. The same datasets were used in the Rapid-Plan method for comparison.

All CT images were acquired using General Electric Light Speed Radiotherapy System 4 (GE Medical Systems, Milwaukee, WI). CT images with the following dimensions were utilized for each axial slice: image matrix = 512 × 512, slice numbers = 76–106, pixel spacing = 0.98 mm, and slice thickness = 3.75 mm. The study protocol conformed to the ethical guidelines of the Declaration of Helsinki as revised in 1983 and was approved by the Institutional Review Board (IRB) of the National Cancer Center without an IRB number. All patient data were fully anonymized, and all methods were performed in accordance with the relevant guidelines and regulations outlined by our institution.

**Deep learning model for dose prediction**

The network used was based on the open-source library Keras (version 2.2.4) [34] and the reference implementation of U-Net [35]. Figure 1 shows the dose prediction deep neural network (DpNet) for dose prediction, which consists of a down-sampling (encoding) path and an up-sampling (decoding) path. For the encoding path, we used two 3 × 3 convolution layers, which had 64, 128, 256, and 512 filters. Each of these layers was followed by a rectified linear unit (ReLu) [36] and a maximum pooling layer. On the decoding path, we used a 2 × 2



**Fig. 1** Schematic diagram of the U-net based dose prediction deep neural network (DpNet) architecture used for volumetric arc radiation therapy (VMAT) dose distribution prediction

transposed convolution and two  $3 \times 3$  convolution layers followed by a ReLU activation function. Concatenation was performed with the corresponding feature map from the skip connection path and two convolution layers with  $3 \times 3$  filters. To avoid overfitting during training, batch normalization [37] and dropout (the dropout rate was set to 0.2) [38] were added to the layers. In the final layer, we used a  $1 \times 1$  convolution network with a sigmoid activation function. The mean squared error (MSE) loss function utilized in DpNet calculates the difference between the actual and predicted doses according to Eq. (1):

$$MSE = \frac{1}{n} \sum_{i=1}^n (D_p^i - D_c^i)^2, \quad (1)$$

where  $n$  is the total number of training samples and  $p$  and  $c$  are the predicted and clinical doses, respectively. We used Adam [39] as an optimizer with a learning rate of  $1.0E-04$  and mini-batch size of 15 images. The experiments were conducted on a computer workstation with an Intel i7 central processing unit with a 24 GB main memory and a computer unified device architecture library on a graphics processing unit (NVIDIA Titan-Xp with 12 GB of memory). Network training of the DpNet took approximately 48 h to run 5000 epochs on the training and validation datasets.

#### Dose prediction data preprocessing

The input of the dose prediction model was utilized for training and validation of the DpNet, including CT images as well as PTV, heart, left lung, right lung, and right breast contour images. Clinical plan datasets were obtained using Eclipse planning software (version 13.6, Varian Oncology Systems, Palo Alto, CA, USA). All CT images were converted into grayscale images, and the contouring points were converted into segmented contour images in binary format, as depicted in Fig. 1 [12]. All training images were resized from the conventional size of  $512 \times 512$  pixels to  $256 \times 256$  pixels owing to graph card memory resource limitations and to reduce the DpNet training time.

#### Deep-learning-based dose prediction process

The deep learning dose prediction (Dp) model is a DpNet training of contours, a CT image as input data, and the training and validation sets consisted of 35 and 10 patient datasets, respectively. The 10 test sets used for model validation were employed as independent, separate dataset images for DpNet, as shown in Fig. 2.

Five-fold cross-validation [40] was performed to improve the model accuracy because the training dataset was insufficient. The five-fold average loss  $\pm$  standard

deviation was  $0.09 \pm 0.01$  (training loss) and  $0.73 \pm 0.07$  (validation loss). The third fold performed the best, with the lowest validation loss of 0.65. This model was used to evaluate the dose predictions for the test set of patients. To compare the results with the clinical dose distribution, the predicted dose of the test set was normalized to 100% of the prescription dose in 95% of the PTV. For the dose prediction of the Dp model, the test process is carried out with the patient CT as an input, and dose prediction is performed not only on the PTV contour but also on the area that does not include the PTV contour as the entire area inside the body contour.

#### RapidPlan model for dose prediction

The RapidPlan dose prediction (Rp) model configuration consisted of two parts.

First, the geometric and dosimetric information of the patient was extracted from a group of selected available approved treatment plans (training and validation datasets), and an automated DVH estimation model was created based on the extracted features.

Second, the lower boundary of the DVH estimation model parameter of the new patient (test dataset) was generated by predicting the DVH using the trained, optimized model.

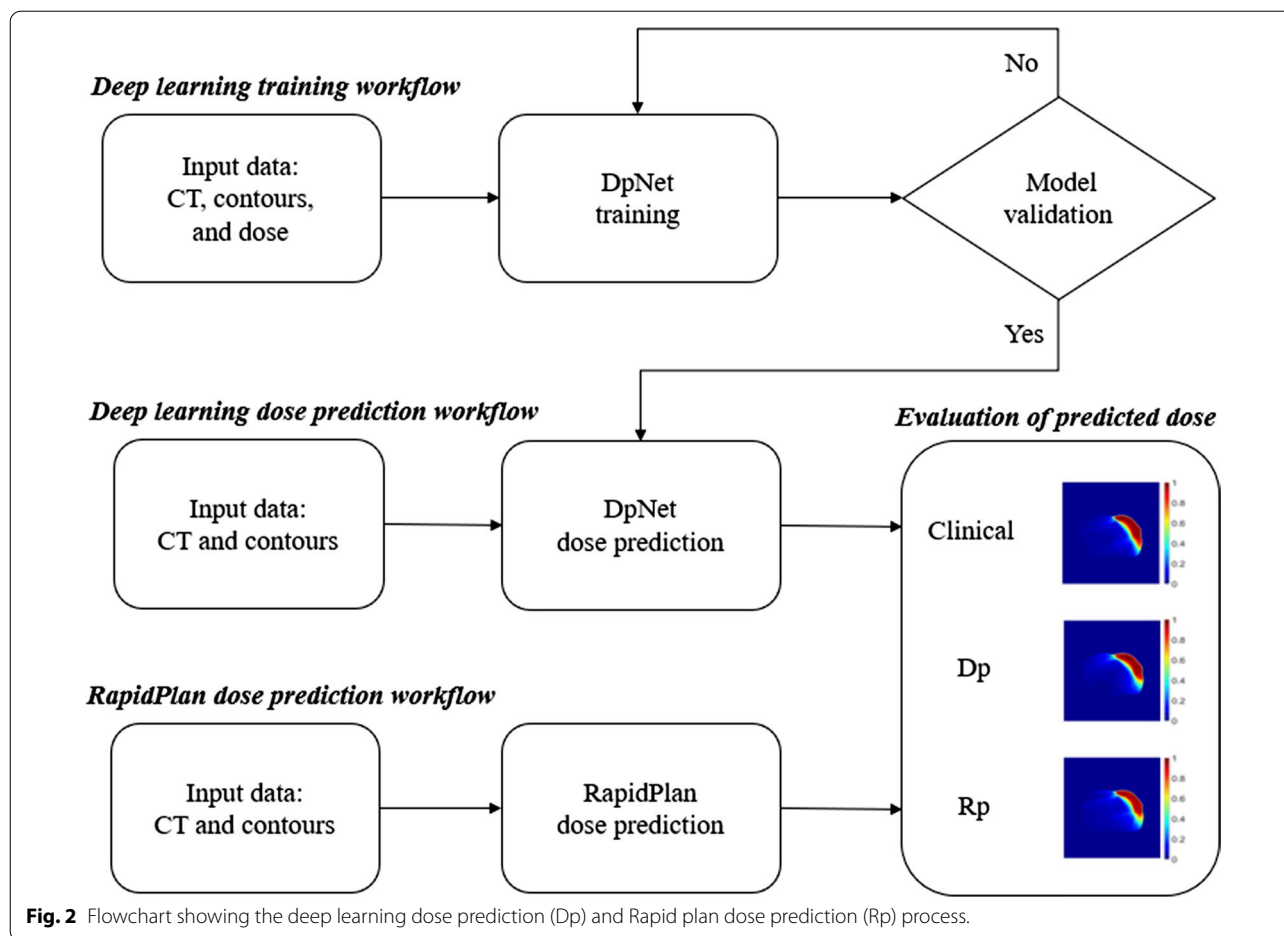
To compare and evaluate the performance of the Dp, the dose distribution was generated using RapidPlan with the same training, validation, and test datasets. AAA (version 13.6) was used as the dose calculation algorithm, and normal tissue objects were not used in the same way as in the clinical plan. All normalized dose distributions were the same as those in the clinical protocol of our institution, as mentioned above.

For PTV, the manual optimization objectives were applied in the same way as in the clinical plan; separately, the dose volume constraints were applied as OAR optimization objectives to the line objectives generated from the predicted DVH. The Rp plan is a result of the optimization process; thus, the similarity between the original plan and the Rp plan depends on the similarity of the objective setting in the original plan and in the Rp plan creation.

#### Quantitative dose prediction evaluation

To evaluate quantitatively the accuracy of Dp and Rp, the 3D dose distribution, the maximum and mean dose absolute differences between the clinical and predicted doses in the OARs and PTV, the DVH, 2D gamma analysis, and the isodose volume dice similarity coefficient (iDSC) [31] were used.

First, the clinical and predicted dose volumes of the Dp and Rp models were compared with the 3D dose distribution.



**Fig. 2** Flowchart showing the deep learning dose prediction (Dp) and Rapid plan dose prediction (Rp) process.

Second, in the indirect evaluation method, the absolute dose errors of the clinical and predicted doses were calculated using the following equation:

$$PTV, OARs \text{ percentage of absolute error} = \left| \frac{Clinical \text{ dose} - Predicted \text{ dose}(Dp, Rp)}{Clinical \text{ dose}} \right| \times 100\%. \tag{2}$$

Third, the clinical and predicted doses obtained using the DVH, the most commonly used treatment plan evaluation tool, were compared.

Fourth, the gamma analysis metric, which is utilized in evaluating complex modulated radiotherapy, was calculated by simultaneously considering the dose difference and distance to agreement. The clinical dose was compared with the 2D gamma index as a reference, and the 2D dose distribution corresponding to the transverse plane was calculated. The gamma index passing criteria were 3%/3 mm and 2%/2 mm, and the calculation for the whole body was performed without a dose

threshold. In addition, areas outside the body were not included in the gamma index calculation.

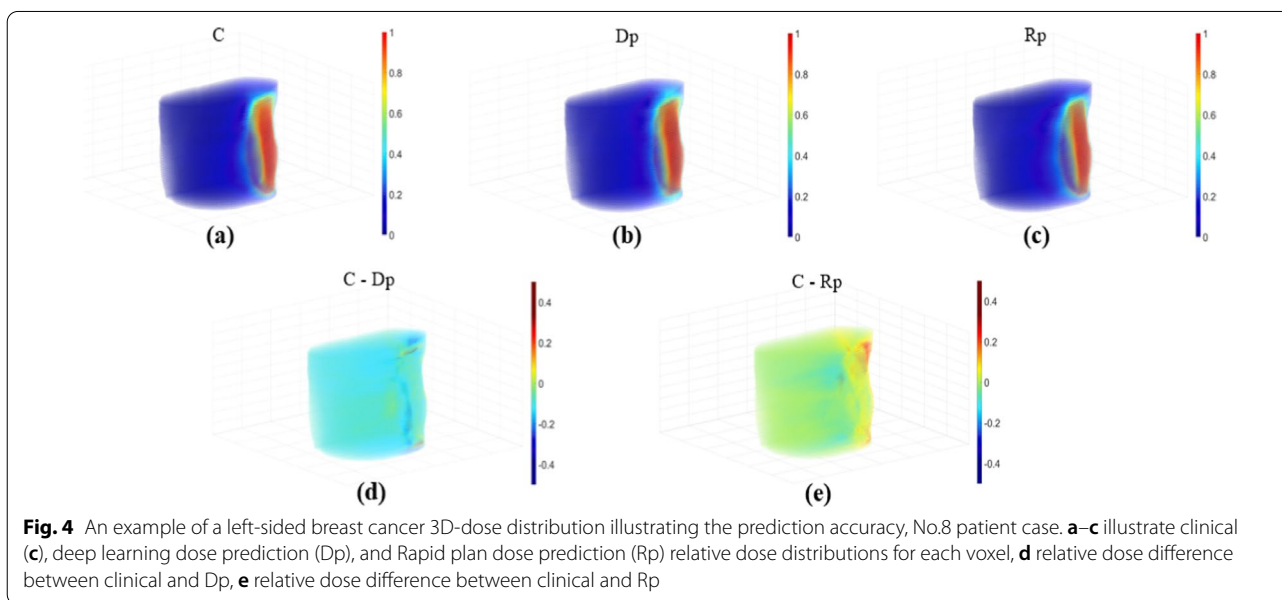
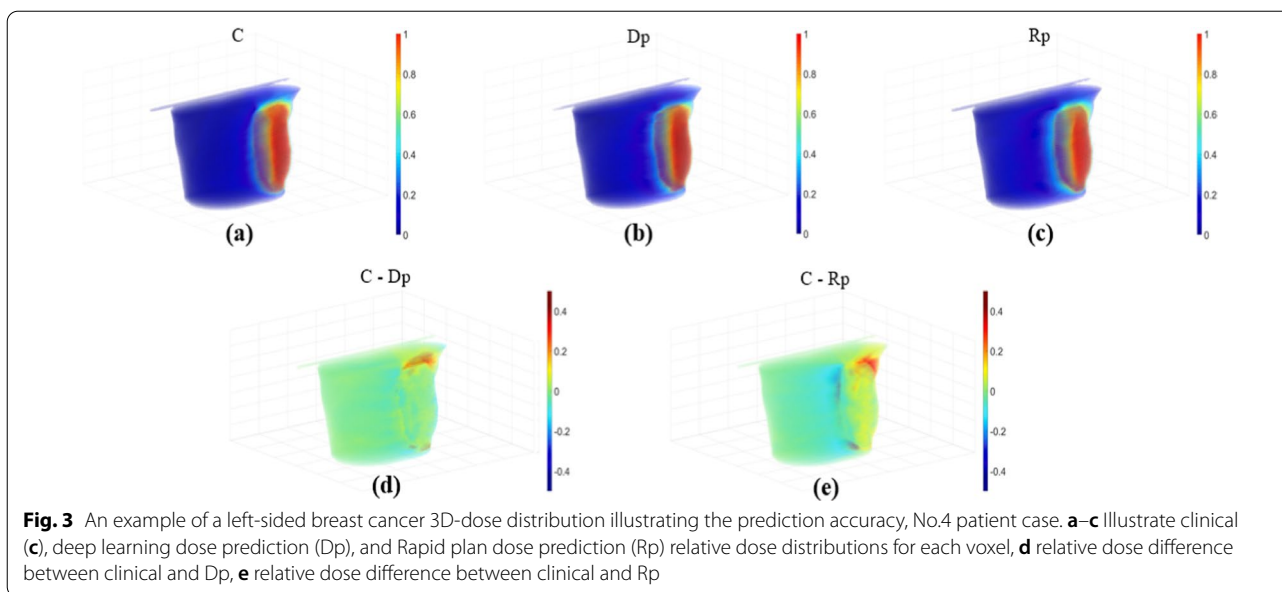
Finally, the dice similarity coefficient of the isodose

volume was evaluated in the 3D dose distribution. The iDSC method involves calculating the overlapping results of two different volumes according to the following equation:

$$iDSC = \frac{2|A \cap B|}{|A| + |B|}, \tag{3}$$

where A is the clinical isodose volume and B is the predicted isodose volume (Dp and Rp). iDSC takes values between zero and one. When iDSC approaches zero, the clinical and prediction results differ significantly.





However, as iDSC approaches one, the two volumes exhibit increased similarities.

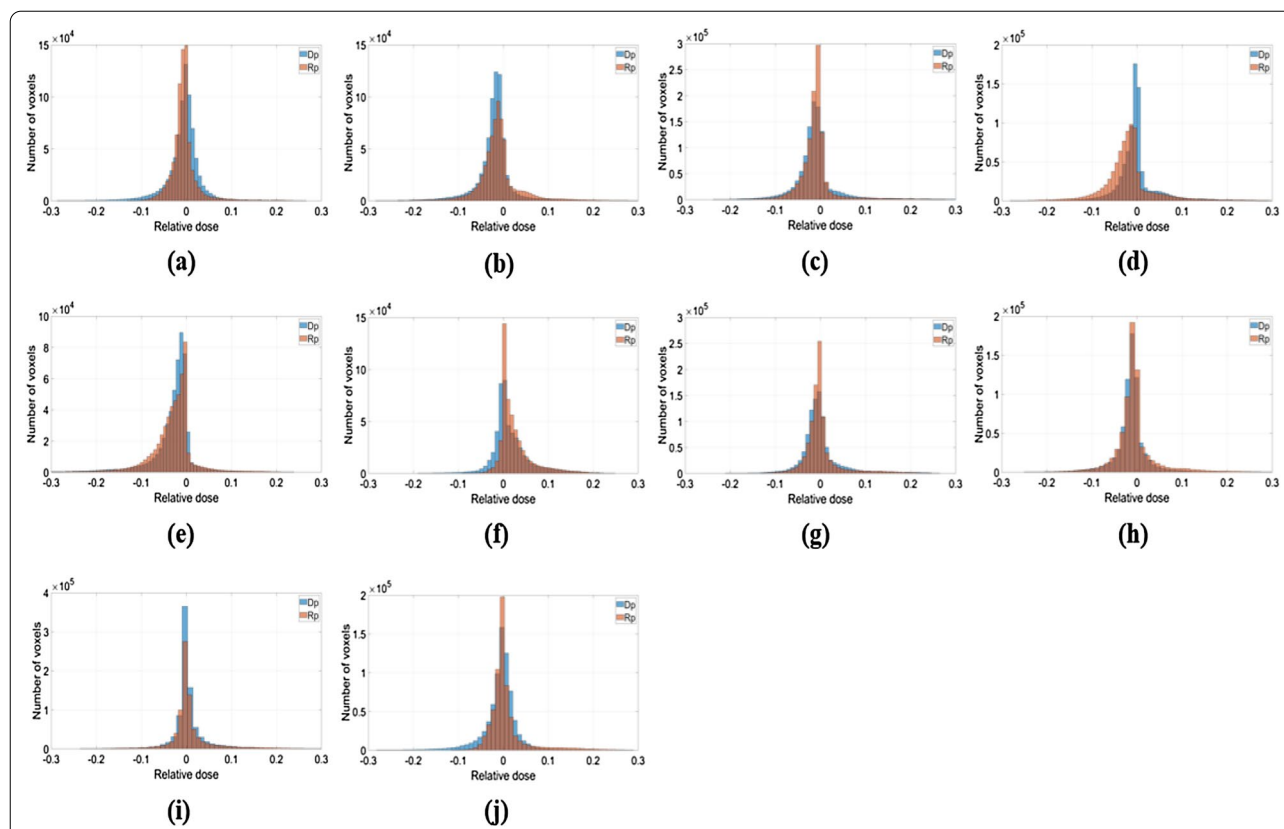
We used the Wilcoxon test to determine the statistical significance of the differences between the clinical Dp and Rp results.

**Results**

For quantitative evaluation, Figs. 3, 4 and 5 depict the relative dose differences for the test set between the clinical dose and the Dp and Rp 3D dose distributions and the differences from the clinical dose histogram. The

histogram distributions were compared; in cases 2, 4, and 5 of the test set, the Dp dose showed less difference from the clinical dose than the Rp dose. In cases 1, 6, and 10 of the test set, the Rp dose was more consistent with the clinical dose than the Dp dose, and there was no significant difference in cases 7–9, as demonstrated in Fig. 5.

Tables 2 and 3 present the absolute maximum and mean dose errors and SDs of the PTV and organ structures of the test set. For the absolute maximum dose error, the average differences between the errors of the Dp and Rp are 1.28%, 0.90%, -3.74%, -4.24%,



**Fig. 5** The relative 3D-dose distribution difference compared clinically with the Dp and Rp is shown by histogram for the ten-test data set. **a–j** are test set cases of No. 1–10. Zero value of relative dose was not included in the histogram plot

and  $-3.07\%$  for the body, left breast, heart, left lung, and right lung, respectively, as shown in Table 3. For the absolute mean dose error, the average differences between the

in Figs. 6 and 7. The Dp dose distribution results in the right lung are more consistent with the clinical results. Figures 8 and 9 present the 2D gamma analysis criteria

**Table 3** Comparative mean absolute error average ( $\pm$ SD) dosimetric results for OARs in 10 tested patients

Organ structure	$D_{mean}$				$D_{max}$			
	C-Dp (a) (%)	C-Rp (b) (%)	p value	$  a  -  b  (\%)$	C-Dp (a) (%)	C-Rp (b) (%)	p value	$  a  -  b  (\%)$
Body	$0.16 \pm 0.82$	$-0.29 \pm 0.98$	0.01	$-0.13$	$2.21 \pm 2.14$	$0.93 \pm 1.77$	0.01	1.28
Left breast	$0.52 \pm 0.97$	$1.30 \pm 0.86$	0.01	$-0.78$	$2.08 \pm 2.13$	$1.18 \pm 1.92$	0.13	0.9
Heart	$-0.88 \pm 1.83$	$-0.32 \pm 1.10$	0.20	0.56	$-1.79 \pm 8.43$	$5.53 \pm 4.97$	0.01	$-3.74$
Left lung	$-1.16 \pm 2.58$	$0.12 \pm 2.13$	0.05	1.04	$0.67 \pm 3.71$	$4.91 \pm 2.80$	0.01	$-4.24$
Right lung	$-0.97 \pm 1.73$	$-1.74 \pm 1.79$	0.25	$-0.77$	$-6.73 \pm 9.13$	$-9.8 \pm 10.03$	0.16	$-3.07$

Dp and Rp are  $-0.13\%$ ,  $-0.78\%$ ,  $0.56\%$ ,  $1.04\%$ , and  $-0.77\%$  for the body, left breast, heart, left lung, and right lung, respectively, as summarized in Table 3. In the PTV case, the differences in  $D_{95\%}$ ,  $D_{50\%}$ ,  $D_{2\%}$ , and  $D_{mean}$  between the Dp and Rp models are less than 1%, although the difference in  $D_{max}$  is larger, as observed in Table 4.

The DVHs of two patients (Nos. 4 and 8) were compared with the approved clinical results and are shown

of 2%/2 mm and 3%/3 mm with the Rp and Dp doses as references for the clinical dose. In Table 5, which summarizes the average gamma analysis passing rates calculated from all slices of the 2D dose with clinical dose distribution, there is no significant difference in the Dp model criteria of 2%/2 mm, and the passing rate of 0.03 is high at 3%/3 mm. In particular, the standard deviation of the

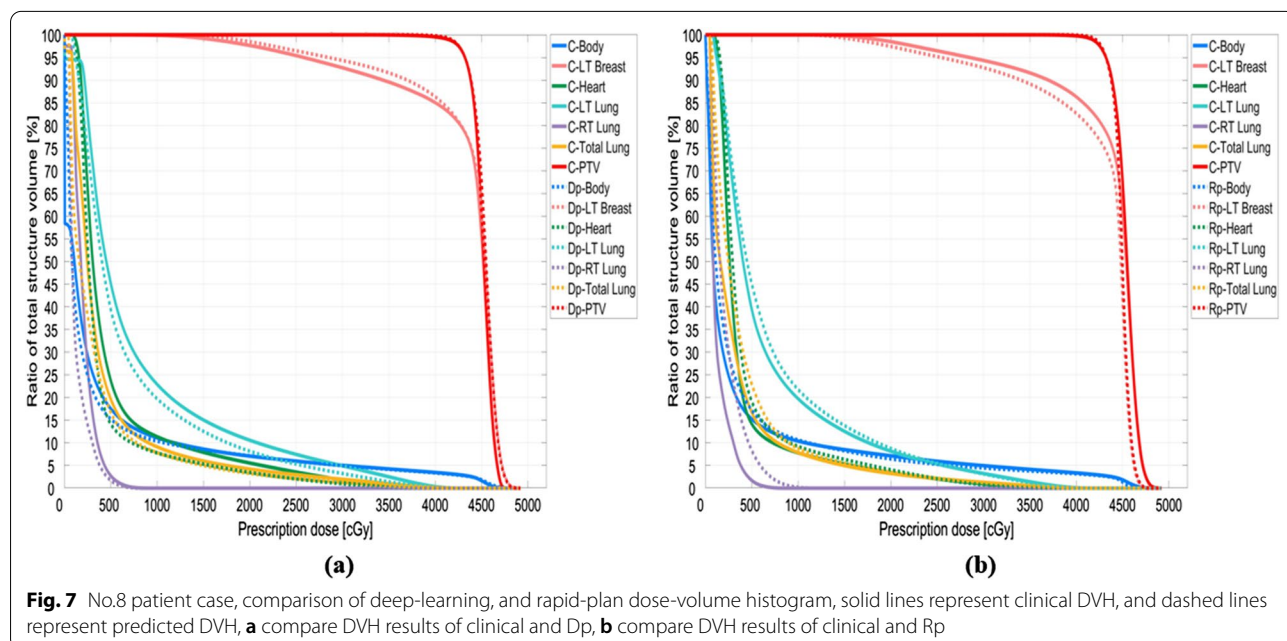
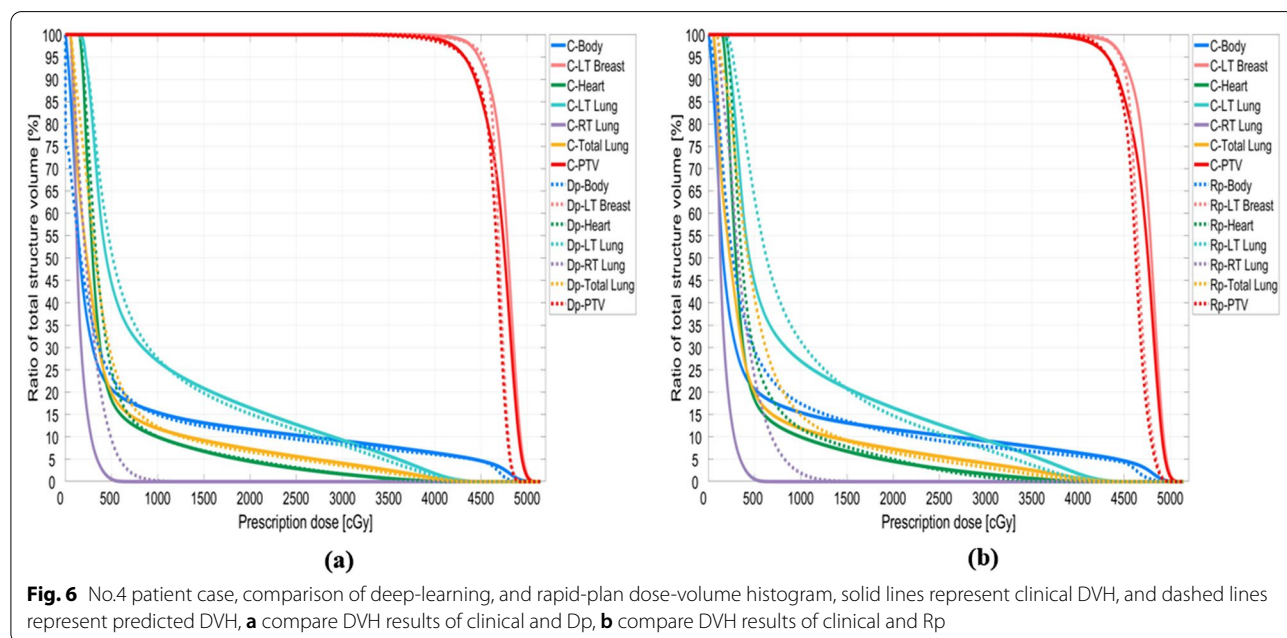
**Table 4** Comparative mean absolute error average ( $\pm$ SD) dosimetric results for PTV in 10 tested patients

Dosimetric index	PTV			
	C-Dp (a) (%)	C-Rp (b) (%)	p value	$  a - b  $ (%)
D <sub>95%</sub>	0.02 $\pm$ 0.04	0.02 $\pm$ 0.04	-	0
D <sub>50%</sub>	0.86 $\pm$ 0.53	0.90 $\pm$ 0.72	0.75	- 0.04
D <sub>2%</sub>	1.01 $\pm$ 0.67	0.91 $\pm$ 0.53	0.50	0.1
D <sub>max</sub>	1.51 $\pm$ 2.68	0.58 $\pm$ 1.82	0.08	0.93
D <sub>mean</sub>	0.01 $\pm$ 0.83	0.87 $\pm$ 0.63	0.01	- 0.86

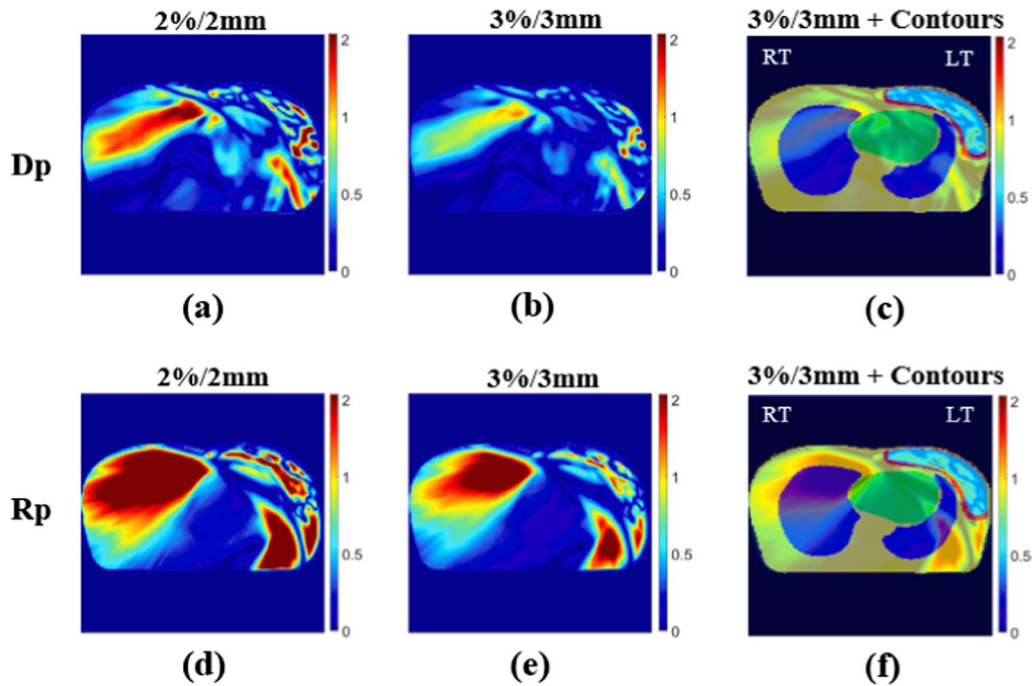
D<sub>n%</sub> is means the dose received by n % of the PTV

Rp model is approximately two times higher than that of the Dp model.

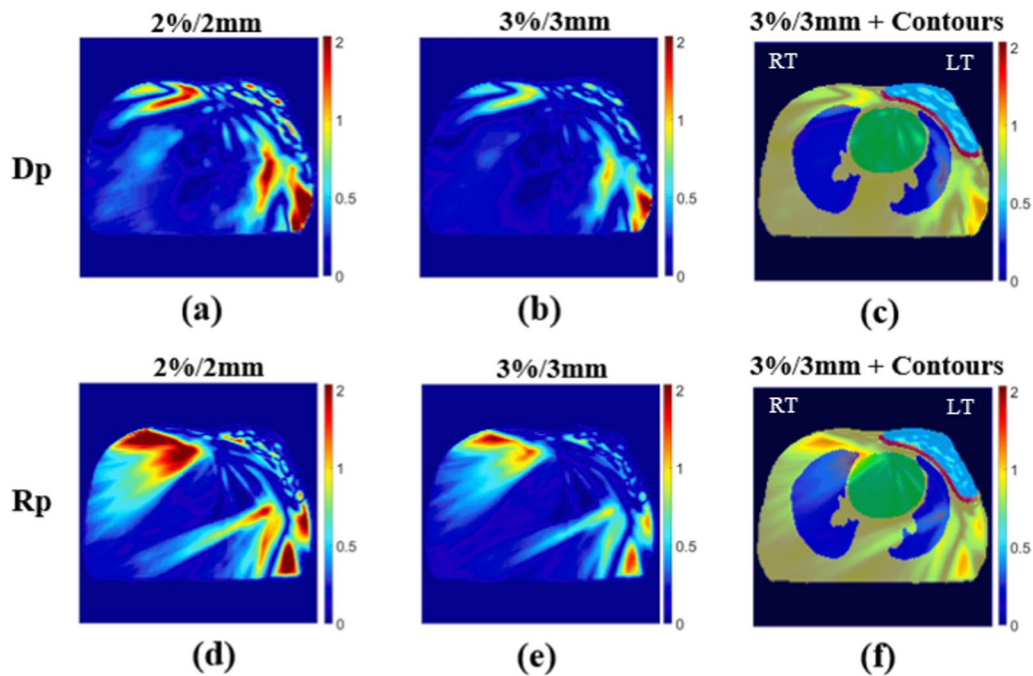
Figure 10 displays iDSC for 10 test datasets from 0 to 100% of the isodose volume. The solid red line represents the average iDSC, which usually ranges from zero to one, with one indicating an ideal match. The Rp and Dp low isodose volumes (ranging from 3 to 20%) show a tendency for iDSC to be less than 0.9. The Dp model always has iDSC > 0.9 at a high isodose volume (range from 90 to 100%). However, the Rp model has cases in which the iDSC of the test set (Nos. 2, 9, and 10) is less than 0.9.







**Fig. 8** 2D gamma analysis results of No.4 patient case at isocenter position, **a** dose difference and distance to agreement passing criteria of 2 %/2 mm Dp, **b** criteria of 3 %/3 mm Dp, **d** criteria of 2 %/2 mm Rp, **e** criteria of 3 %/3 mm Rp, **c, f** 3%/3mm gamma analysis results and contours overlapping images (the meaning of contour color, yellow: body, green: heart, blue: lungs, red: left breast, cyan: PTV)



**Fig. 9** 2D gamma analysis results of No.8 patient case at isocenter position, **a** dose difference and distance to agreement passing criteria of 2 %/2 mm Dp, **b** criteria of 3 %/3 mm Dp, **d** criteria of 2 %/2 mm Rp, **e** criteria of 3 %/3 mm Rp, **c** and **f** 3%/3mm gamma analysis results and contours overlapping images (the meaning of contour color, yellow: body, green: heart, blue: lungs, red: left breast, cyan: PTV)

**Table 5** All of the slice 2-D gamma analysis average pass rate with 3%/3 mm and 2%/2 mm criteria for 10 test sets

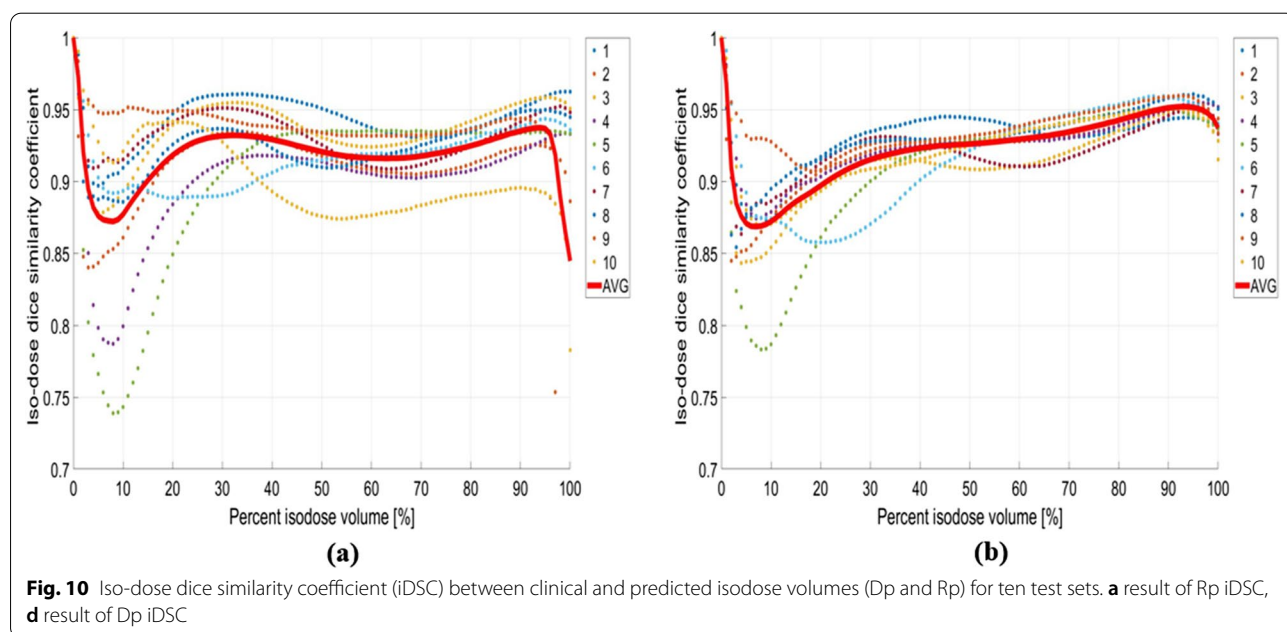
Dosimetric index	2%/2 mm		3%/3 mm	
	Dp	Rp	Dp	Rp
Ave	0.71	0.70	0.85	0.82
SD	0.06	0.12	0.04	0.10

**Discussion**

In this study, deep learning was utilized for VMAT dose distribution prediction using the anatomical features in the planning CT for left-sided breast cancer and the performance of this approach was compared with that of RapidPlan.

Our DpNet model consists of convolutional neural network layers to output the dose distribution according to the input target, OAR contours, and anatomical information (CT) (Fig. 1). Deep learning-based dose prediction

studies have been reported on tumors centrally located in the body, such as rectal [24] and prostate [25, 26] cancer tumors. In the case of breast cancer, the target anatomical position is close to the body outside the area and the left lung, and the dose conformity is lower than that in the prostate case [41, 42]. The approach using deep learning to predict the clinically accepted dose distribution in the case of inhomogeneity around the target is different from those utilized in previous studies [24–26]. In Table 6, Yoganathan and Zhang [43] predicted the dose distribution for left breast cancer and reported that the prediction  $D_{mean}$  error (over entire CT volume) was  $0.9 \pm 1.2$  Gy with an atlas-based method. Bai et al. [44] obtained  $0.48 \pm 2.27$  Gy and  $0.42 \pm 1.82$  Gy from dose prediction using the similarity selection method (SIM) based on the most similar atlas image and the weighted method (WEI\_F) applying weighted dose distribution with database images.



**Table 6** Summary of MADs presented in previous literature on left breast cancer

	Yoganathan et al. [40]	Bai et al. [41]		Our study	
Training data set	19	20		50	
Method	A	A (SIM)	A (WEI_F)	D (Dp)	A (Rp)
MAD (Gy) over entire CT volume	$0.9 \pm 1.1$	$0.48 \pm 2.27$	$0.42 \pm 1.82$	$0.32 \pm 0.07$	$0.35 \pm 0.08$

A: atlas-based dose prediction, D: deep learning-based dose prediction

The mean absolute differences (MADs) [43, 44] in the test set over the entire CT image according to the Dp and Rp models show small differences from the clinical plan dose, and the SD is also small, indicating consistent dose prediction.

As the patient datasets and breast cancer treatment protocols differ by institution, the ability to perform direct comparison is limited; however, evidence has shown that dose prediction using deep learning methods is possible. Tables 2–4 demonstrate that the Dp model obtained results superior to those of the Rp model because it differed less from the clinical plan. In particular, in Fig. 10, the Dp model shows  $iDSC < 0.9$  only in the region of low iso-dose volumes (from 2 to 20%), but for the Rp model in cases Nos. 2, 9, and 10, the results are less than 0.9 at high iso-dose volumes (from 95 to 100%) as well as low iso-dose volumes.

The calculated  $iDSC$  was less than 0.9 at high iso-dose volumes because the process of obtaining the optimal treatment plan using the Rp model optimization process and the manual method are different [45].

If the plan created using the Rp model does not satisfy the clinical goal in terms of the DVH and dose distribution, re-optimization must be performed, which is time consuming.

The time efficiency was determined based on the average times required by the Dp and Rp models for dose prediction, which were  $9.82 \pm 0.37$  s and  $676.08 \pm 81.23$  s, respectively (i.e., the times differ with statistical significance because the p-values were smaller than 0.05 when a ranked Wilcoxon test was performed).

The reason for why the time difference in the process in the dose prediction of the two models was large is that Rp spends a large amount of time on MLC (multi leaf collimator) optimization and final dose calculation. On the other hand, in the case of Dp, dose prediction is possible in seconds due to the fact that only image information is required as an input, while model training takes a long time.

In this study, since the goals of the models of Dp and Rp are different, it is difficult to accurately compare the performance. Rp can create a treatment plan, and it is also possible for the planner to directly intervene to create treatment goal or better plan. However, in the case of Dp, only 3D dose prediction in our model it is possible, and for a better plan, the process of training and validation of the model is required again. However, other studies only compared the clinical results [22–30] The Rp model was used to compare the dose prediction accuracy of the Dp model for the same patient test case.

The reason for the deviation in the prediction result of the Rp model may be due to the difference between the

quality of the input plan and the dosimetric strategy for Rp model training, as suggested by Fogliata et al. [46].

As shown in Figs. 8 and 9, the potential cause of the difference from the clinical result in the gamma analysis result in the right breast region is also a result of the imbalance in each plan objective.

This study had several limitations.

First, the current deep learning model provides only patient anatomical information (CT, contour) as inputs without dosimetric information. Second, only the learned one-type dose prediction is possible, and the deep neural network must be retrained for other IMRT or 3D-CRT treatment techniques and other sites.

Third, it is impossible to perform conversion into an executable treatment plan using the predicted dose distribution results.

Fourth, when selecting the plan data used to train the Rp plan, it is necessary to evaluate and confirm each plan by considering the difference between each primary object.

Nevertheless, dose predictability using deep learning was demonstrated in this study by quantitatively comparing the patient dose predictions obtained by automated radiation treatment planning [47] using deep learning with those determined existing commercial programs. The dose distributions predicted by deep learning will help reduce the iterative optimization process because planners can identify the areas in which to deliver increased or decreased doses in advance.

In a future study, to overcome the limitations of the currently developed Dp model, the dosimetric feature [24, 30] should be included in the input data to reflect the physical characteristics to increase the dose prediction accuracy. Based on the learned model, transfer learning [21, 48] will be applied to enable dose prediction for various treatment sites.

We will be able to evaluate the accuracy of the plan and the time it takes for the planning time after conducting a test on the new patient to see if the proposed model can provide effective guidance in order to satisfy the goal of the treatment plan.

We will develop a program using the Eclipse Scripting API to generate an optimal plan automatically based on the predicted dose.

## Conclusion

In this study, VMAT dose distribution predictions obtained by deep learning were compared with RapidPlan results for left-sided breast cancer patients using contour and CT images only. Our deep learning model produced superior dose predictions compared to RapidPlan and showed that dose prediction using deep learning is possible. In addition, radiation treatment planning

based on the dose predicted using deep learning will improve the radiation treatment process by reducing the time required for planning, while maintaining plan quality.

#### Abbreviations

AAA: Analytical anisotropic algorithm; CT: Computed tomography; CNN: Convolutional neural network; DpNet: Dose prediction deep neural network; DVH: Dose volume histogram; Dp: Deep learning dose prediction; ESAPI: Eclipse scripting API; iDSC: Isodose volume dice similarity coefficient; IMRT: Intensity modulation radiation therapy; KBP: Knowledge-based planning; MAE: Mean absolute error; MSE: Mean squared error; MAD: Mean absolute difference; OAR: Organs at risk; PTV: Planning target volume; ReLU: Rectified linear unit; RT: Radiotherapy; Rp: RapidPlan dose prediction; SD: Standard deviation; VMAT: Volumetric arc radiation therapy.

#### Acknowledgements

Not applicable.

#### Authors' contributions

SHA, ESK, and JHJ conceived the study, participated in its design and coordination, and helped draft the manuscript. ESK, WJC, DYK, and MSK developed RapidPlan treatment plans. SHA, ESK, CK, WJC, MSK, SBL, YKL, HK, DS, and DYK analyzed part of the data and interpreted the data. The SHA developed and designed the software and wrote the technical parts. All authors read and approved the final manuscript.

#### Funding

This study was supported by a National Cancer Center Grant (2110390-1).

#### Availability of data and material

The data are not publicly available because of patient privacy concerns but are available from the corresponding author upon reasonable request.

#### Declarations

##### Ethics approval and consent to participate

This study was approved by national cancer center review board and was conducted in accordance with the ethical standards of the Declaration of Helsinki.

##### Consent for publication

Not applicable.

##### Competing interests

The authors declare that they have no competing interest.

##### Author details

<sup>1</sup>Department of Radiation Oncology, Proton Therapy Center, National Cancer Center, 323, Ilsan-ro, Ilsandong-gu Gyeonggi-do, Goyang-si 10408, South Korea. <sup>2</sup>Department of Radiation Oncology, Yonsei Cancer Center, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, South Korea.

Received: 10 January 2021 Accepted: 19 July 2021

Published online: 17 August 2021

#### References

- Fogliata A, Nicolini G, Clivio A, Vanetti E, Laksar S, Tozzi A, Scorsetti M, Cozzi L. A broad scope knowledge-based model for optimization of VMAT in esophageal cancer: validation and assessment of plan quality among different treatment centers. *Radiat Oncol*. 2015;10(1):220.
- Wu H, Jiang F, Yue H, Li S, Zhang Y. A dosimetric evaluation of knowledge-based VMAT planning with simultaneous integrated boosting for rectal cancer patients. *J Appl Clin Med Phys*. 2016;17(6):78–85.
- Tol JP, Delaney AR, Dahele M, Slotman BJ, Verbakel WF. Evaluation of a knowledge-based planning solution for head and neck cancer. *Int J Radiat Oncol Biol Phys Biol Phys*. 2015;91(3):612–20.
- Schubert C, Waletzko O, Weiss C, Voelzke D, Toperim S, Roeser A, Puccini S, Piroth M, Mehrens C, Kueter JD, Hierholz K, Gerull K, Fogliata A, Block A, Cozzi L. Intercenter validation of a knowledge-based model for automated planning of volumetric modulated arc therapy for prostate cancer. The experience of the German RapidPlan Consortium. *PLOS ONE*. 2017;12(5):e0178034.
- Cagni E, Botti A, Micera R, Galeandro M, Sghedoni R, Orlandi M, Iotti C, Cozzi L, Iori M. Knowledge-based treatment planning: an inter-technique and inter-system feasibility study for prostate cancer. *Phys Med PM Int J Devot Appl Phys Med Biol Off J Ital Assoc Biomed Phys*. 2017;36:38–45.
- Scaggion A, Fusella M, Roggio A, Bacco S, Pivato N, Rossato MA, Peña LMA, Paiusco M. Reducing inter- and intra-planner variability in radiotherapy plan output with a commercial knowledge-based planning solution. *Phys Med*. 2018;53:86–93.
- Ma C, Huang F. Assessment of a knowledge-based RapidPlan model for patients with postoperative cervical cancer. *Precis Radiat Oncol*. 2017;1(3):102–7.
- Kubo K, Monzen H, Ishii K, Tamura M, Kawamorita R, Sumida I, Mizuno H, Nishimura Y. Dosimetric comparison of RapidPlan and manually optimized plans in volumetric modulated arc therapy for prostate cancer. *Phys Med PM Int J Devot Appl Phys Med Biol Off J Ital Assoc Biomed Phys*. 2017;44:199–204.
- Fogliata A, Reggiori G, Stravato A, Lobefalo F, Franzese C, Franceschini D, Tomatis S, Mancosu P, Scorsetti M, Cozzi L. RapidPlan head and neck model: the objectives and possible clinical benefit. *Radiat Oncol*. 2017;12(1):73.
- Hu P, Wu F, Peng J, Bao Y, Chen F, Kong D. Automatic abdominal multi-organ segmentation using deep convolutional neural network and time-implicit level sets. *Int J Comput Assist Radiol Surg*. 2017;12(3):399–411.
- Gibson E, Giganti F, Hu Y, Bonmati E, Bandula S, Gurusamy K, Davidson B, Pereira SP, Clarkson MJ, Barratt DC. Automatic multi-organ segmentation on abdominal CT with dense v-networks. *IEEE Trans Med Imaging*. 2018;37(8):1822–34.
- Ahn SH, Yeo AU, Kim KH, Kim C, Goh Y, Cho S, Lee SB, Lim YK, Kim H, Shin D, Kim T, Kim TH, Youn SH, Oh ES, Jeong JH. Comparative clinical evaluation of atlas and deep-learning-based auto-segmentation of organ structures in liver cancer. *Radiat Oncol*. 2019;14(1):213.
- Lei Y, Fu Y, Wang T, Liu Y, Patel P, Curran WJ, Liu T, Yang X. 4D-CT deformable image registration using multiscale unsupervised deep learning. *Phys Med Biol*. 2020;65(8):085003.
- de Vos BD, Berendsen FF, Viergever MA, Sokootti H, Staring M, Išgum I. A deep learning framework for unsupervised affine and deformable image registration. *Med Image Anal*. 2019;52:128–43.
- de Vos BD, Berendsen FF, Viergever MA, Staring M, Išgum I (2017) End-to-end unsupervised deformable image registration with a convolutional neural network. In: Lecture notes in computer science. Springer, Cham, pp 204–212.
- Wang Y, Yu Z, Sivanagaraja T, Veluvolu KC (2020) Fast and accurate online sequential learning of respiratory motion with random convolution nodes for radiotherapy applications. *Applied Soft Computing*, 95, <http://www.ncbi.nlm.nih.gov/pubmed/106528>.
- Lin H, Shi C, Wang B, Chan MF, Tang X, Ji W. Towards real-time respiratory motion prediction based on long short-term memory neural networks. *Phys Med Biol*. 2019;64(8):085010.
- Meyer A, Zverinski D, Pfahringer B, Kempfert J, Kuehne T, Sündermann SH, Stamm C, Hofmann T, Falk V, Eickhoff C. Machine learning for real-time prediction of complications in critical care: a retrospective study. *Lancet Respir Med*. 2018;6(12):905–14.
- Ibragimov B, Toesca D, Chang D, Yuan Y, Koong A, Xing L. Development of deep neural network for individualized hepatobiliary toxicity prediction after liver SBRT. *Med Phys*. 2018;45(10):4763–74.
- Ibragimov B, Toesca DAS, Chang DT, Yuan Y, Koong AC, Xing L. Automated hepatobiliary toxicity prediction after liver stereotactic body radiation therapy with deep learning-based portal vein segmentation. *Neurocomputing*. 2020;392:181–8.

21. Zhen X, Chen J, Zhong Z, Hrycushko B, Zhou L, Jiang S, Albuquerque K, Gu X. Deep convolutional neural network with transfer learning for rectum toxicity prediction in cervical cancer radiotherapy: a feasibility study. *Phys Med Biol*. 2017;62(21):8246–63.
22. Chen X, Men K, Li Y, Yi J, Dai J. A feasibility study on an automated method to generate patient-specific dose distributions for radiotherapy using deep learning. *Med Phys*. 2019;46(1):56–64.
23. Fan J, Wang J, Chen Z, Hu C, Zhang Z, Hu W. Automatic treatment planning based on three-dimensional dose distribution predicted from deep learning technique. *Med Phys*. 2019;46(1):370–81.
24. Zhou J, Peng Z, Song Y, Chang Y, Pei X, Sheng L, Xu XG. A method of using deep learning to predict three-dimensional dose distributions for intensity-modulated radiotherapy of rectal cancer. *J Appl Clin Med Phys*. 2020;21(5):26–37.
25. Nguyen D, Long T, Jia X, Lu W, Gu X, Iqbal Z, Jiang S. A feasibility study for predicting optimal radiation therapy dose distributions of prostate cancer patients from patient anatomy using deep learning. *Sci Rep*. 2019;9(1):1076.
26. Kajikawa T, Kadoya N, Ito K, Takayama Y, Chiba T, Tomori S, Nemoto H, Dobashi S, Takeda K, Jingu K. A convolutional neural network approach for IMRT dose distribution prediction in prostate cancer patients. *J Radiat Res*. 2019;60(5):685–93.
27. Barragán-Montero AM, Nguyen D, Lu W, Lin MH, Norouzi-Kandalan R, Geets X, Sterpin E, Jiang S. Three-dimensional dose prediction for lung IMRT patients with deep neural networks: robust learning from heterogeneous beam configurations. *Med Phys*. 2019;46(8):3679–91.
28. Nguyen, D., Jia, X., Sher, D., Lin, M. H., Iqbal, Z., Liu, H., & Jiang, S. Three-dimensional radiotherapy dose prediction on head and neck cancer patients with a hierarchically densely connected U-net deep learning architecture. 2018. <https://arxiv.org/abs/1805.10397>.
29. Song Y, Hu J, Liu Y, Hu H, Huang Y, Bai S, Yi Z. Dose prediction using a deep neural network for accelerated planning of rectal cancer radiotherapy. *Radiother Oncol J Eur Soc Ther Radiol Oncol*. 2020;149:111–6.
30. Ma M, Kovalchuk N, Buyyounouski MK, Xing L, Yang Y. Incorporating dosimetric features into the prediction of 3D VMAT dose distributions using deep convolutional neural network. *Phys Med Biol*. 2019;64(12):125017.
31. Liu Z, Fan J, Li M, Yan H, Hu Z, Huang P, Tian Y, Miao J, Dai J. A deep learning method for prediction of three-dimensional dose distribution of helical tomotherapy. *Med Phys*. 2019;46(5):1972–83.
32. Low DA, Harms WB, Mutic S, Purdy JA. A technique for the quantitative evaluation of dose distributions. *Med Phys*. 1998;25(5):656–61.
33. Lombardi A, Vitale V, Nigri G, Olivieri C, Mastrangeli MR, Bizzaglia E, Amanti C. Prognostic impact of the 8th edition American Joint Committee on Cancer (AJCC) cancer staging system on clinically negative lymph nodes (cN0) breast cancer patients. *Breast J*. 2020;26(7):1265–9.
34. Keras CF. The Python deep learning library. In: *Astrophysics Source Code Library*; 2018.
35. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. In: *Lecture Notes in Computer Science international conference on medical image computing and computer-assisted intervention*. Cham: Springer; 2015; pp 234–241.
36. Nair V, Hinton GE. Rectified linear units improve restricted Boltzmann machines. In: *Proceedings of the 27th international conference on machine learning (ICML-10)*; 2010; pp 807–814.
37. Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. *Clin Orthop Rel Res*. 2015.
38. Hinton GE, Srivastava N, Krizhevsky A, Sutskever I, Salakhutdinov RR. Improving neural networks by preventing co-adaptation of feature detectors. <https://arxiv.org/abs/1207.0580>. 2012.
39. Kingma DP, Adam BJ. A method for stochastic optimization. In: *Proceedings of the 3rd international conference for learning representations (ICLR 2015)*; 2015.
40. Arlot S, Celisse A. A survey of cross-validation procedures for model selection. *Stat Surv*. 2010;4:40–79.
41. Cao T, Dai Z, Ding Z, Li W, Quan H. Analysis of different evaluation indexes for prostate stereotactic body radiation therapy plans: conformity index, homogeneity index and gradient index. *Precis Radiat Oncol*. 2019;3(3):72–9.
42. Karpf D, Sakka M, Metzger M, Grabenbauer GG. Left breast irradiation with tangential intensity modulated radiotherapy (t-IMRT) versus tangential volumetric modulated arc therapy (t-VMAT): trade-offs between secondary cancer induction risk and optimal target coverage. *Radiat Oncol*. 2019;14(1):156.
43. Yoganathan SA, Zhang R. An atlas-based method to predict three-dimensional dose distributions for cancer patients who receive radiotherapy. *Phys Med Biol*. 2019;64(8):085016.
44. Bai X, Wang B, Wang S, Wu Z, Gou C, Hou Q. Radiotherapy dose distribution prediction for breast cancer using deformable image registration. *Biomed Eng Online*. 2020;19(1):39.
45. Lodwick WA, McCourt S, Newman F, Humphries S. Optimization methods for radiation therapy plans. In: *IMA volumes in mathematics & its applications*. Springer, New York, NY. 1999; pp 229–249.
46. Fogliata A, Cozzi L, Reggiori G, Stravato A, Lobefalo F, Franzese C, et al. RapidPlan knowledge based planning: iterative learning process and model ability to steer planning strategies. *Radiat Oncol*. 2019;14(1):1–12.
47. Mahmood R, Babier A, McNiven A, Diamant A, Chan TC. Automated treatment planning in radiation therapy using generative adversarial networks. <https://arxiv.org/abs/1807.06489>. 2018.
48. Bengio Y. Deep learning of representations for unsupervised and transfer learning. In: *Proceedings of the ICML workshop on unsupervised and transfer learning*; 2012; pp 17–36.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

