



PAPER

OPEN ACCESS

RECEIVED
10 September 2019REVISED
5 February 2020ACCEPTED FOR PUBLICATION
13 February 2020PUBLISHED
8 April 2020

Original Content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



DeepDose: Towards a fast dose calculation engine for radiation therapy using deep learning

C Kontaxis, G H Bol, J J W Lagendijk and B W Raaymakers

Department of Radiotherapy, University Medical Center Utrecht, Heidelberglaan 100, Utrecht 3584 CX, The Netherlands

E-mail: c.kontaxis@umcutrecht.nl**Keywords:** dose engine, IMRT, deep learning, AI, treatment planning, plan adaptation, MR-linac

Abstract

We present DeepDose, a deep learning framework for fast dose calculations in radiation therapy. Given a patient anatomy and linear-accelerator IMRT multi-leaf-collimator shape or segment, a novel set of physics-based inputs is calculated that encode the linac machine parameters into the underlying anatomy. These inputs are then used to train a deep convolutional network to derive the dose distribution of individual MLC shapes on a given patient anatomy.

In this work we demonstrate the proof-of-concept application of DeepDose on 101 prostate patients treated in our clinic with fixed-beam IMRT. The ground-truth data used for training, validation and testing of the prediction were calculated with a state-of-the-art Monte Carlo dose engine at 1% statistical uncertainty per segment. A deep convolution network was trained using the data of 80 patients at the clinically used 3 mm³ grid spacing while 10 patients were used for validation.

For another 11 independent test patients, the network was able to accurately estimate the segment doses from the clinical plans of each patient passing the clinical QA when compared with the Monte Carlo calculations, yielding on average 99.9%±0.3% for the forward calculated patient plans at 3%/3 mm gamma tests. Dose prediction using the trained network was very fast at approximately 0.9 seconds for the input generation and 0.6 seconds for single GPU inference per segment and 1 minute per patient in total.

The overall performance of this dose calculation framework in terms of both accuracy and inference speed, makes it compelling for online adaptive workflows where fast segment dose calculations are needed.

1. Introduction

The accurate modelling of the delivered dose to the patient is an essential part of the treatment planning process in photon radiation therapy. Several dose calculation engines have been used in 3D treatment planning systems (TPS) including pencil beam (PB) (Mohan *et al* 1986), collapsed cone (CC) (Ahnesjö 1989) and Monte Carlo (MC) (Rogers 2006) based ones. MC simulations have been proven to yield the best dose accuracy compared to other algorithms by modelling the full particle transport especially around tissue interfaces and air cavities where varying types of materials are present (Arnfield *et al* 2000, Krieger and Sauer 2005).

One downside of the early MC implementations was the long computation times when compared to the PB/CC algorithms. In the recent years this has been addressed partly due to the speed increase in newer hardware and by GPU based MC implementations (Hissoiny *et al* 2011, Tian *et al* 2015). Fast dose calculations are especially crucial for MRI-guided applications that include online dose calculations/treatment replanning where a new plan has to be generated and approved in a few minutes while the patient lies on the treatment couch (Henke *et al* 2018, Werensteijn-Honingh *et al* 2019). Moreover, as we progress to more advanced MRI-guided interventions, intrafraction replanning will require almost real-time and accurate dose calculations for use during optimization (Kontaxis *et al* 2017).

In Intensity Modulated Radiation Therapy (IMRT) dose calculations are performed in multiple timepoints during treatment planning depending on the optimization process used. In a typical two-phase optimization (Bortfeld *et al* 1994), the dose engine is initially used to calculate the ideal dose distributions in the patient anatomy for all included beams split into several thousand beam elements. Then, following an optimization sequencing step that yields deliverable multi-leaf-collimator (MLC) shapes or segments, the dose engine calculates the dose of these shapes while taking into account the machine's MLC, jaws and patient anatomy. For Direct Aperture Optimization (Shepard *et al* 2002) methods the MLC shapes and their weights are simultaneously being updated during the optimization requiring multiple iterations of dose calculations.

Machine learning and more recently deep learning (DL) with convolutional networks has been applied in the field of treatment planning. The ability to accurately predict the treatment dose has been demonstrated for several treatment sites including prostate, head and neck and lung (Nguyen *et al* 2019, Mahmood *et al* 2018, Barragán-Montero *et al* 2019) using different types of network architectures. DL based dose prediction can be utilized to assist the treatment planning specialist during the plan optimization process by providing an accurate estimate of the optimal dose distribution.

The prediction in these methods relies on the correlation of patient anatomy and delineated structures to the total treatment dose and does not take into account the actual machine parameters of the linac including MLC leaf/jaw positions and gantry/MLC angles of the individual segments. The direct correlation of machine parameters to patient anatomy would enable a flexible dose prediction framework equivalent to a dose engine.

In this work we explore the use of deep convolutional networks to predict the dose of individual MLC shapes, bypassing the use of a conventional dose engine. We introduce the DeepDose framework, a DL pipeline that correlates the MLC shapes of the linear-accelerator (linac) to the patient anatomy by encoding meaningful physical information into the learning process. For this proof-of-concept experiments we train and apply this pipeline on MLC shapes of the clinical plans from prostate patients previously treated in our clinic with conventional linear-accelerators. This initial work aims to demonstrate the network's ability to predict the dose for arbitrary shapes in new patient anatomies before proceeding to more specific/targetted applications.

2. Materials and methods

2.1. Patient data

We used the data from 101 prostate patients previously treated in our clinic with 5-beam IMRT on Elekta linacs with 10 MV flattened beam and an MLCi collimator with a $40 \times 40 \text{ cm}^2$ field size, 80 leaf pairs and 1 cm leaf width. The prescribed gantry and collimator angles were set to 40° , 100° , 180° , 260° , 320° and 0° , 3° , 6° , 354° , 357° respectively. For these patients the prostate boost volume was treated with 77 Gy in 35 fractions.

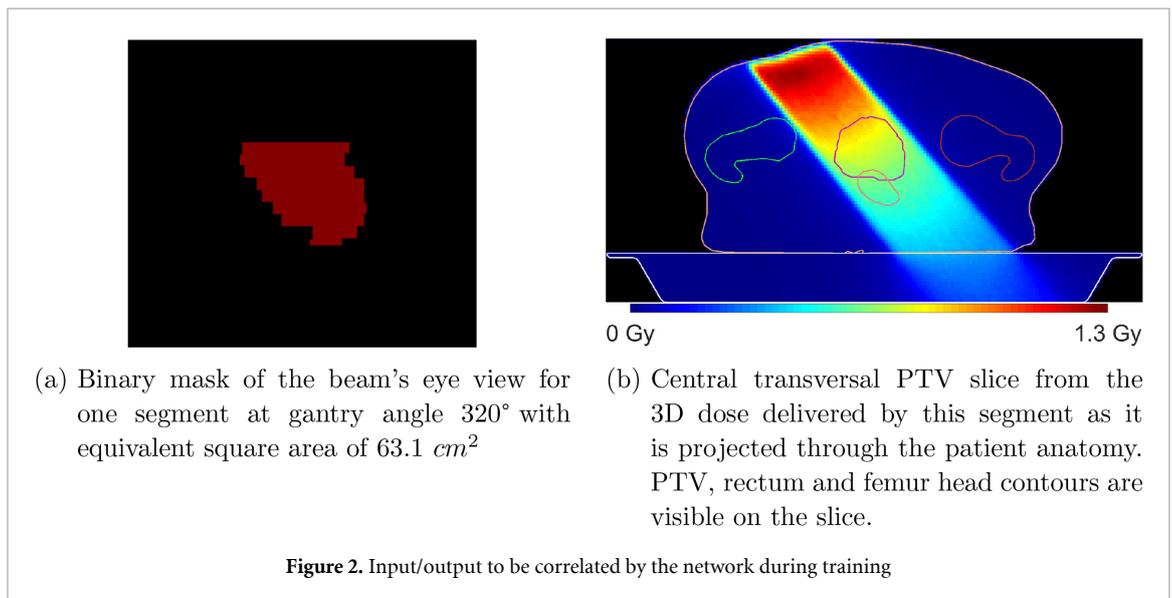
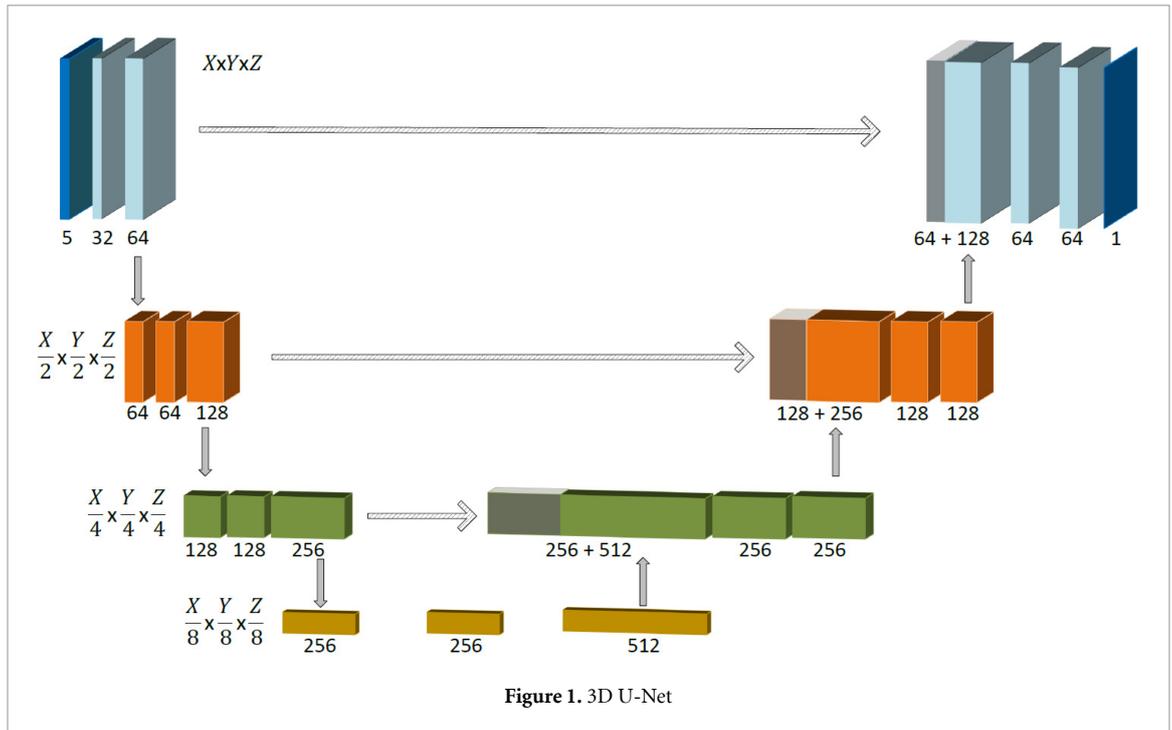
For each patient the clinical plan was parsed to extract the MLC shapes used during treatment, in total 4176 segments among all patients. Each segment is matched to the anatomical information of its corresponding patient leading to a unique patient-segment combination used as a training/validation/testing dataset for our network. The experiments, including all input and output data, were performed at a 3 mm^3 grid spacing which is the clinical resolution used for treatment planning and dosimetric evaluation. All patient grids included the treatment couch and were cropped to the same size of $192 \times 128 \times 80$, while maintaining their respective isocenters at the center of the volume and verifying that the grid safely includes the total dose distribution of the patient. The data were split in 80 patients for training, 10 for validation and 11 for testing, and thus 3290 training, 437 validation and 449 testing segments respectively. The average patient age was 71 ± 5.9 years, 69 ± 6.6 years and 70 ± 8.7 years for the training, validation and test sets.

The ground-truth dose distributions for each of the 4176 segments were calculated using the research version of the GPUMCD Monte Carlo dose engine (Hissoiny *et al* 2011) using a statistical uncertainty of 1%. Each segment was calculated with 100 Monitor Units (MU) enabling a standardized dose output which can then be scaled as needed. The ones corresponding to the training/validation datasets provide the target doses during training and validation, while the ones corresponding to the test cases are used to evaluate the prediction results of the network.

2.2. DeepDose

2.2.1. DL network

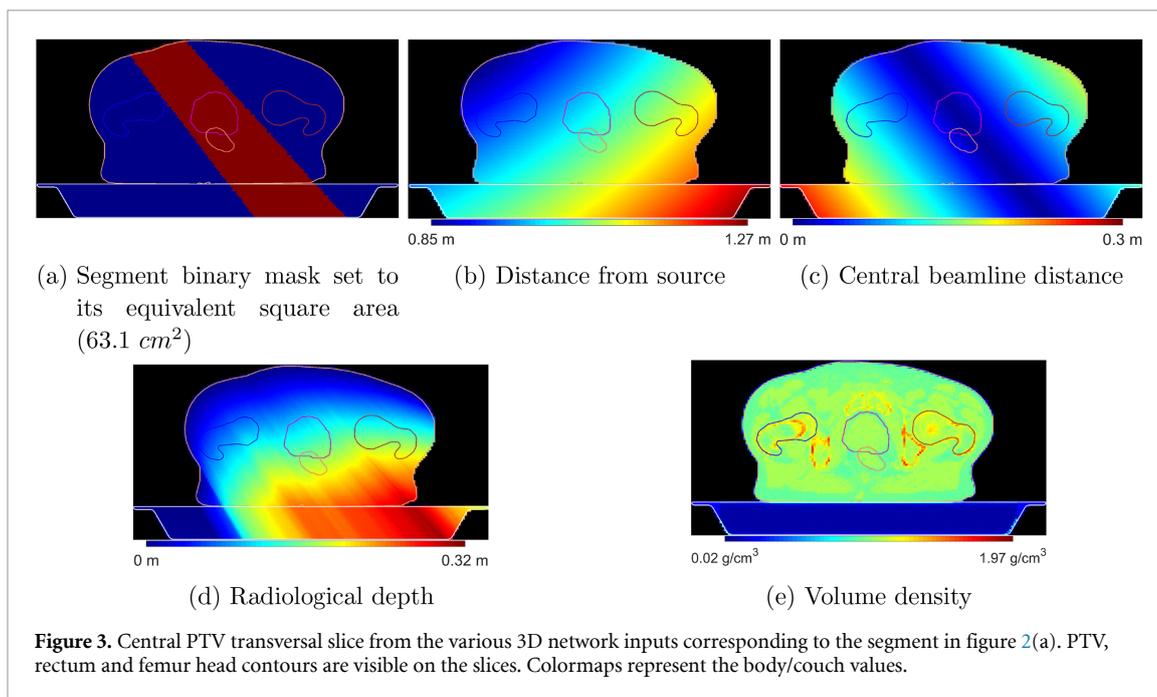
We used the 3D U-net network originally published by Çiçek *et al* (2016) for the purpose of volumetric organ segmentation. The U-net architecture includes a decoding and encoding path (figure 1). The encoding path consists of successive layers of two $3 \times 3 \times 3$ convolutions with batch normalization + rectified linear units (ReLU) and a $2 \times 2 \times 2$ max pooling with stride 2, while the decoding path consists of $2 \times 2 \times 2$



deconvolutions with stride 2 and two $3 \times 3 \times 3$ convolutions with batch normalization + ReLus. For each resolution of the decoding path there is a skip connection passing high-resolution features from the encoding path. The last layer also includes a $1 \times 1 \times 1$ convolution with a single output channel corresponding to the predicted dose. The network expects a set of 3D volume inputs of dimensions $X \times Y \times Z$ and yields a dose prediction output of the same size. The experiments were performed by using NiftyNet, a medical physics oriented framework, in which all different network parameters can be specified including the input/output volume $X \times Y \times Z$ size (Gibson *et al* 2018). The included NiftyNet implementation of the U-net was used, modified to yield a single channel output.

2.2.2. Input

Given an MLC shape the goal is to have a network that predicts its 3D dose in any given patient anatomy (figure 2). For that purpose we defined a network that expects 5 different 3D volumes as input, each modelling a physical quantity/feature implicitly used also by conventional dose engines (figure 3). The network's output is the target 3D dose distribution of that segment. The treatment couch is included in all inputs as well as the dose calculation, matching the clinical protocol for dose calculations in prostate treatment at our clinic.



The first input (figure 3(a)) is the 3D mask of the segment ray-traced from the source through the grid encoding the MLC shape into the patient anatomy. The value of all voxels inside the segment mask is set to the area of the equivalent square shape corresponding to each segment, in order to implicitly encode the contribution due to the different field size in the scattering component for the different non-square fields.

The next three inputs pass various distance information to the network enabling a deeper correlation of the anatomy to the machine parameters. Input 2 (figure 3(b)) provides the distance of each voxel from the linac source, modelling the inverse square law of the photon fluence. Input 3 (figure 3(c)) enables the network to learn the varying dose deposition effect for an off-axis MLC shape by providing the shortest distance of each voxel from the central beam line passing through the center of the linac MLC. Input 4 (figure 3(d)) provides the radiological depth allowing the network to take into account the tissue inhomogeneity. More specifically, following the segment as it goes through the anatomy, each voxel is assigned to its distance from the body surface weighted by the average density along this path up to that point. Finally, the electron density per voxel is calculated from the treatment CT of the patient and provided as input, essential for the attenuation and dose deposition in the patient anatomy (figure 3(e)).

The average equivalent square area of all used segments among the training, validation and testing patients was $25.9 \pm 16.7 \text{ cm}^2$, $25.9 \pm 16.8 \text{ cm}^2$ and $25.7 \pm 16.0 \text{ cm}^2$ while the average source to body surface distance was $83.5 \pm 3.0 \text{ cm}$, $83.4 \pm 3.1 \text{ cm}$ and $82.9 \pm 3.2 \text{ cm}$ respectively.

2.2.3. Training

We trained the deep convolutional network using a $3 \times 3 \times 3 \text{ mm}^3$ grid spacing also used in the clinical setting. For more robust training and data augmentation we utilized patch-based training using only a sub-volume of the 3D inputs for each iteration during the learning process. The patch size was set to $32 \times 32 \times 32$ leading to a $9.6 \times 9.6 \times 9.6 \text{ cm}^3$ volume. A batch size set of 16 patches was selected and the root mean squared error (RMSE) was used as the loss function. The ADAM optimizer was used with learning rate, beta1, beta2 and epsilon parameters set to 10^{-4} , 0.9, 0.999 and 10^{-8} respectively. During training validation was performed every ten iterations on the validation dataset to establish the performance of the network.

2.3. Analysis

The experiments were performed on a system with a dual Intel® Xeon® E5-2670 v3, 64 GB RAM. The training/validation/testing of the deep learning network as well as the GPUMCD dose calculations used a single NVIDIA® GTX Titan card.

The predicted dose distributions were evaluated by extracting the central beam and lateral profiles at the isocenter against the ground-truth target doses. Moreover, gamma analysis tests of 3%/3 mm, 2%/2 mm and 1%/1 mm were performed for the predicted/target dose distributions of each individual segment while taking into account the voxels within 10%–100% of the dose maximum.

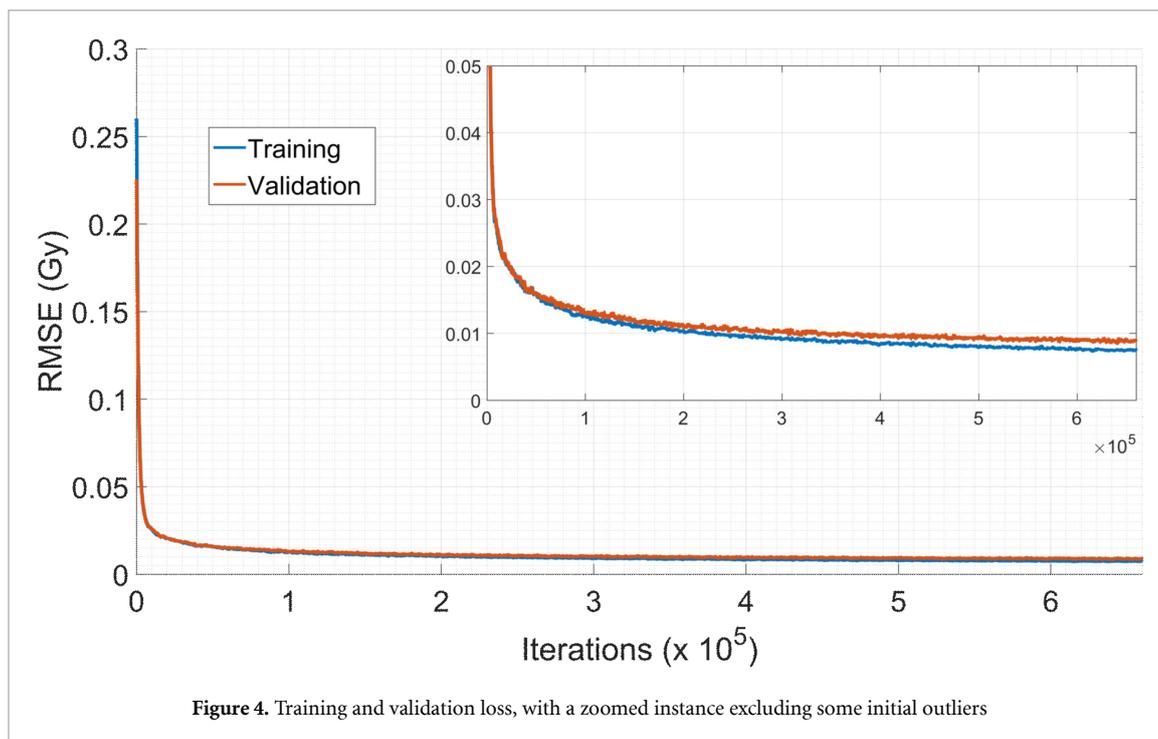


Figure 4. Training and validation loss, with a zoomed instance excluding some initial outliers

In addition, the predicted dose distributions for each patient's segments were summed after being scaled according to the MU values in the clinical treatment plan in order to generate the total forward-calculated plan dose. This was then directly compared to the clinical dose distribution for that patient—also calculated with 1% statistical uncertainty—by comparing DVH parameters and calculating the same gamma analysis stats for voxels within 50%–100% of the dose maximum similarly to what is used clinically during QA tests.

Finally, the absolute and relative dose differences were calculated between predicted and target dose distributions for the voxels within 10%–100% of the dose maximum.

3. Results

The training of the network took approximately 5 days on a single Titan card. It was trained for 6.6×10^5 iterations with 16 patches per batch and thus having processed over 10^7 random patches in total. At these iteration points the validation losses were stable and thus training was stopped to avoid potential over-fitting (figure 4). The average validation and training RMSE was 1.2×10^{-2} and 1.1×10^{-2} Gy respectively.

The trained network was then used to infer the dose per segment for the independent test patients. Inference was very fast at 0.6 sec for the whole patient volume per segment. Subsequently the individual segment doses—all based on 100 MU calculations—were weighted using their MU values from the original plan and summed, leading to the total predicted plan dose.

Overall a very good agreement between the target and predicted dose distributions is observed. Some indicative profiles in the central and lateral directions to the segment dose are plotted in figure 5 for two segments. The network prediction follows the target profiles faithfully, capturing the details encoded in the target dose for both the patient and treatment couch.

The average difference and standard deviation (SD) among all individual segment doses was $-0.4\% \pm 2.6\%$ (0.01 ± 0.01 Gy in absolute values) for the voxels within 10%–100% of dose maximum. For the forward calculated plan with a target prescription of 77 Gy, the 11 test patients showed an average difference of $0.4\% \pm 0.5\%$ (0.4 ± 0.2 Gy in absolute values) for the same inclusion criteria.

Figure 6 shows the DVH and central PTV slice for the total predicted and target dose distributions for one of the test patients. The DVH lines are very close to each other in the low and high dose levels for all structures including femur heads, bladder, rectum and PTV/EBV. It is interesting to see that also the dose of the treatment couch is accurately modelled.

For each of the 449 segments among all patients the gamma analysis pass rate was calculated (figure 7(a)). The average pass rates were $89.9\% \pm 5.1\%$, $97.8\% \pm 3.0\%$ and $99.4\% \pm 2.5\%$ for the 1%/1 mm, 2%/2 mm and 3%/3 mm respectively, demonstrating very good agreement while taking into account low dose values at the cGy level.

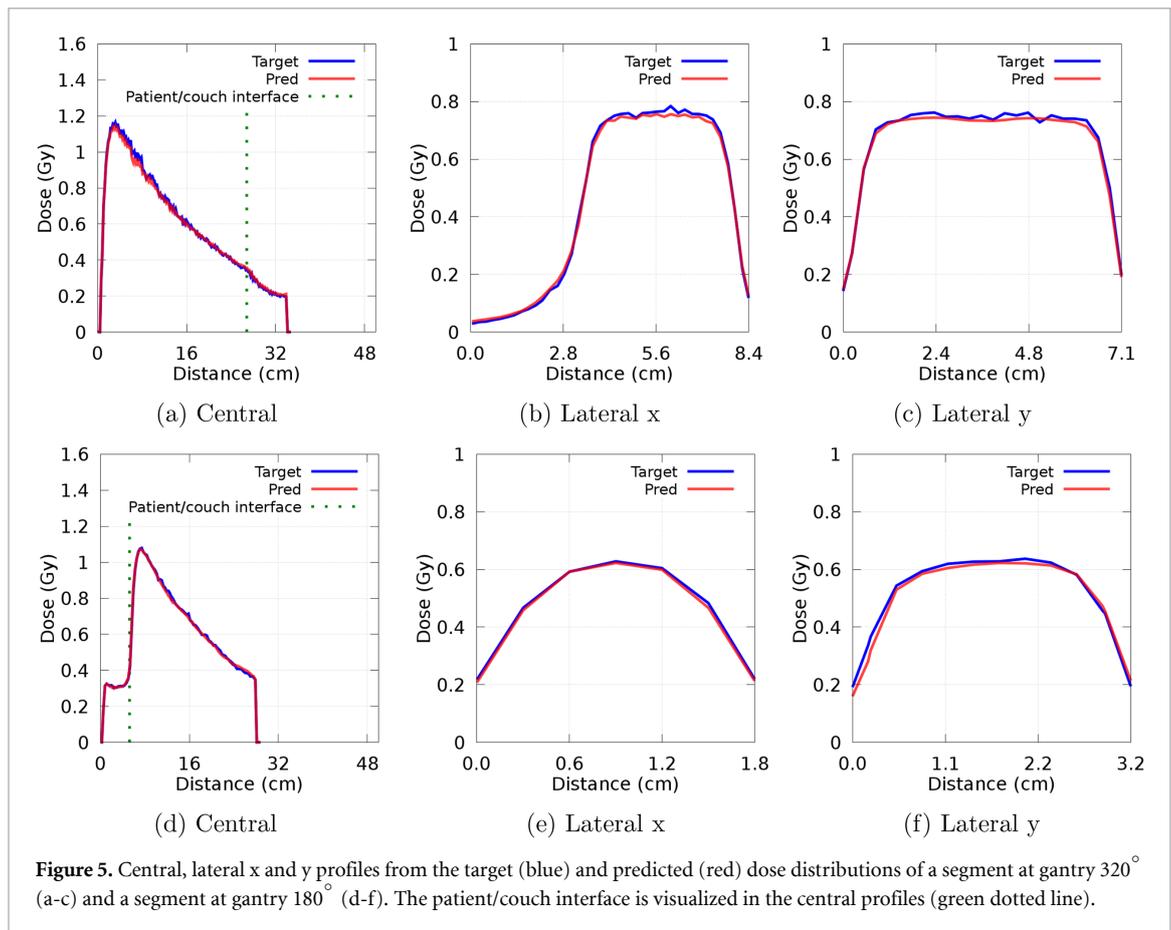


Figure 8 shows one of the mispredicted outlier segments from figure 7(a). In these cases the shape and magnitude of the dose is overall correctly predicted but the network fails to do so at the region where a single narrow slit is open and allows radiation to pass through. Such a shape is not frequently present in this type of prostate plans, which tend to be more uniform and round modelling the shape of the prostate.

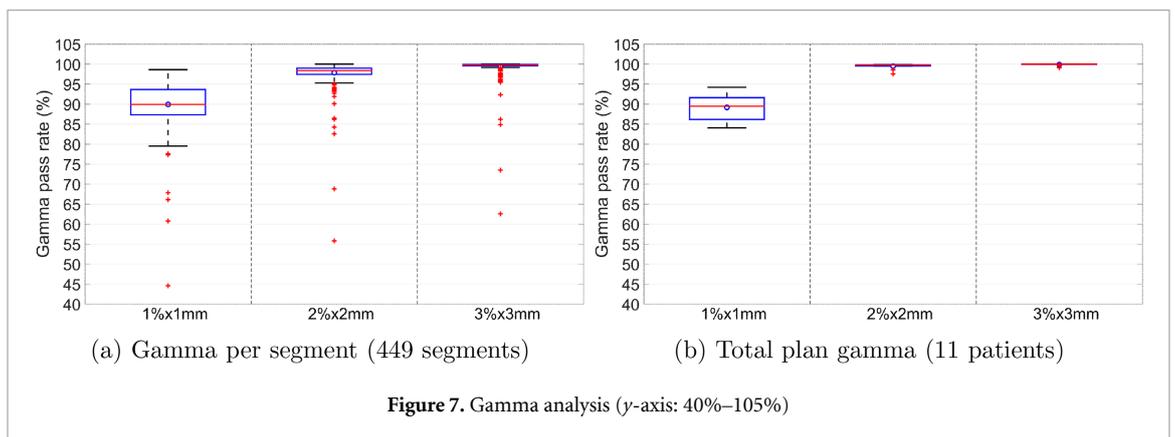
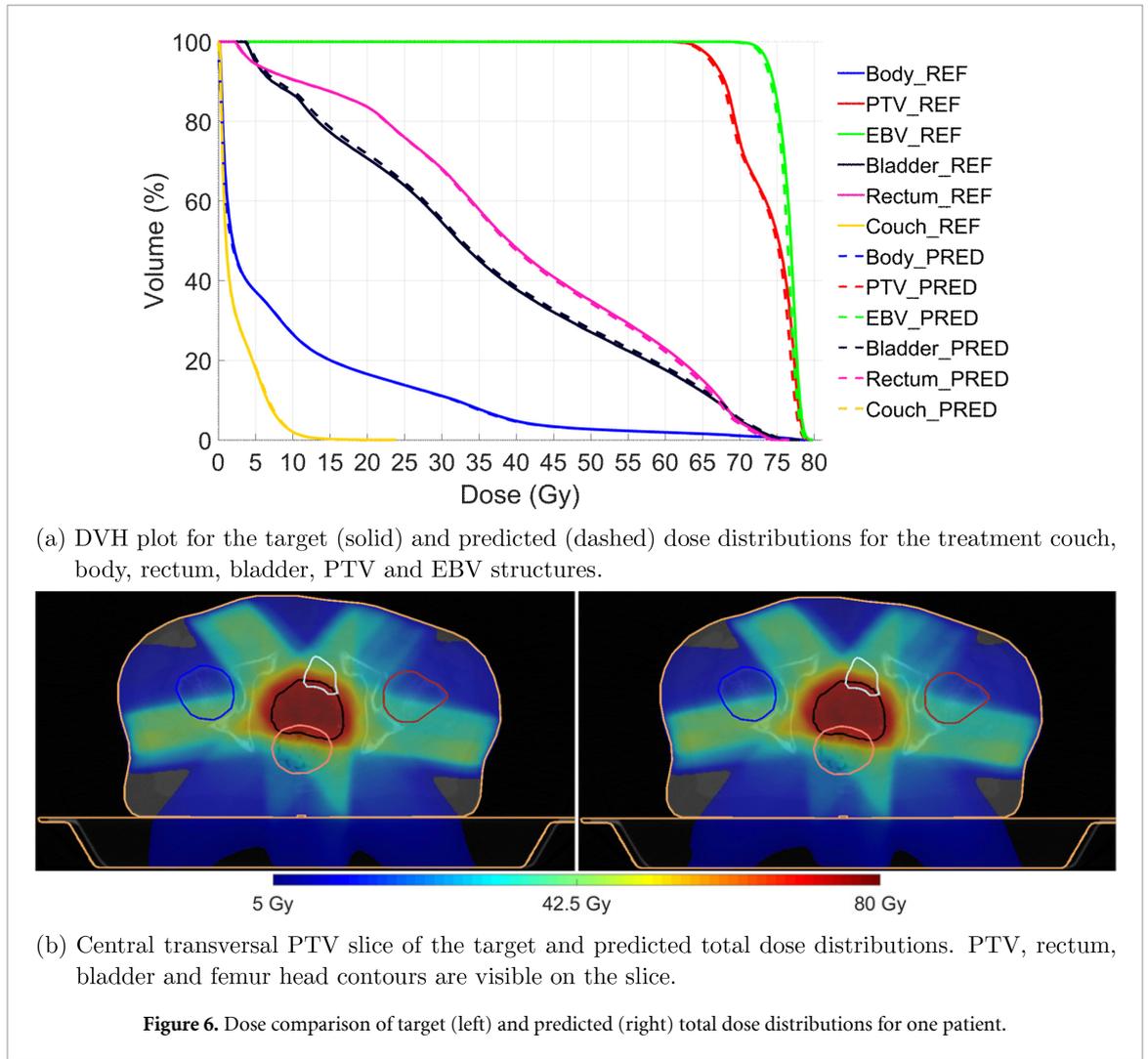
These outliers do not affect the total dose distributions for each patient, which achieve average pass rates of $89.2\% \pm 3.3\%$, $99.5\% \pm 0.7\%$ and $99.9\% \pm 0.3\%$ for the 1%/1 mm, 2%/2 mm and 3%/3 mm respectively were found (figure 7(b)). These values show excellent agreement to the clinical dose, yielding clinical grade dose distribution that pass current clinical QA tests which would allow minimum values of 95% at 3%/3 mm.

Figure 9 shows boxplots of critical DVH points included in the planning constraints for targets and OARs between the ground-truth and predicted total dose distributions for the test patients. The average relative D99% difference was $0.6\% \pm 0.5\%$ (0.4 ± 0.4 Gy in absolute values) for the PTV and $0.5\% \pm 0.4\%$ (0.4 ± 0.3 Gy in absolute values) for the CTV structures (figure 9(a)) respectively. Rectum and bladder V72Gy differences were on average $0.3\% \pm 0.6\%$ and $0.4\% \pm 0.6\%$ respectively (figure 9(b)).

4. Discussion

We have presented the DeepDose framework, a DL application for accurate dose calculation of IMRT MLC shapes in varying anatomy. We have formulated a novel set of physics-based 3D inputs which provide an accurate mapping from MLC shapes and patient anatomy to dose via a conventional 3D U-net network architecture. In these first results focusing on prostate radiotherapy for conventional linacs, we show that these inputs can model the problem successfully at the clinical isotropic 3 mm^3 grid spacing.

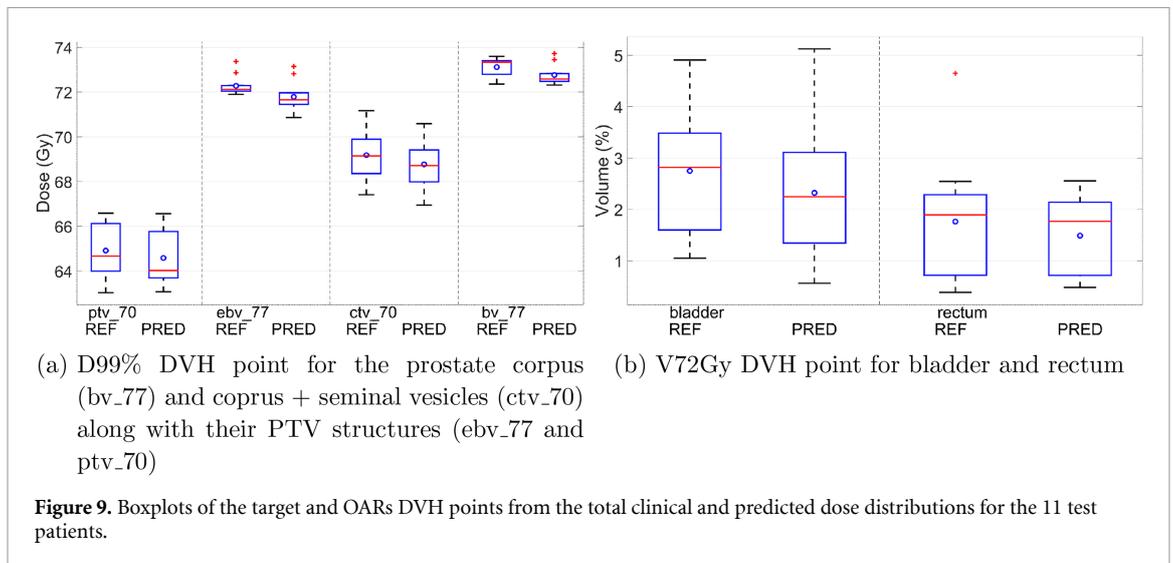
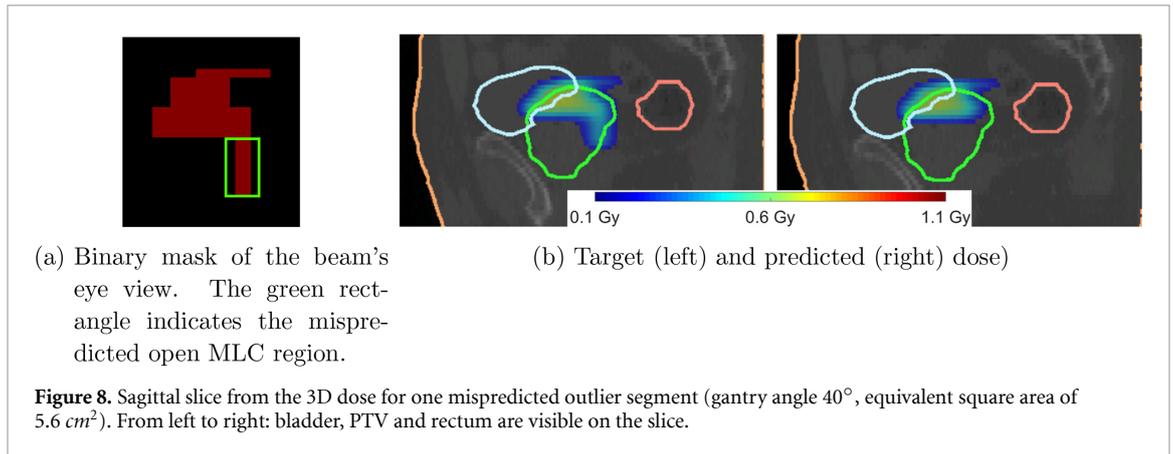
In contrast to previously published DL applications that focus on predicting the total plan dose for a new patient based on imaging data and anatomical structures, in this work we propose an elaborate framework that predicts the dose for each individual segment from a patient's plan based on the actual machine parameters of the linac. This enables a variety of applications for DL based dose prediction besides usage as decision support systems in the treatment planning phase. Initially this framework can be used as a secondary dose engine in the clinical QA workflows, to quickly assess the clinical generated plans prior to treatment. Moreover, in the future it could be extended to a new type of planning system utilizing deep learning in various sub-components able to calculate accurate treatment plans in an online setting.



As demonstrated in this work the framework yields accurate results passing the clinical QA tests at this standard grid spacing. The resulting accuracy is expected to increase by training the network at finer resolution grids—ideally down to 1 mm^3 —as then subtler features will be learned and no interpolation effects will be introduced during analysis at lower resolutions.

Very precise and time consuming dose calculations with low uncertainty can be used for the generation of the training data followed by sub-second dose prediction for new anatomies/segments in the same uncertainty range, otherwise infeasible in an online setting with a conventional dose engine. In this work we trained the network with 1% uncertainty per segment already lower than the clinically used value of 3%.

The forward application of a trained network even in 3D is inherently fast and expanding such a framework to parallel implementations and/or utilization of more cards, would linearly decrease the total



calculation time reported in this work. Furthermore, if the DL approach is integrated into the planning system itself, allowing for native inference without the need of external software, then the overall time will decrease significantly. For the results presented in this work and an average of 41 MLC shapes, the equivalent full plan calculation with GPUMCD takes 2.9 minutes on average per patient while the DeepDose framework is considerably faster at approximately 1 min (37 sec for input generation and 25 sec for inference). In case we are only interested in the dose of the complete plan (Nguyen *et al* 2019), a similar network based on these inputs could be trained to predict the total dose distribution of a new patient with a single inference run.

The inputs presented in section 2.2.2 collectively allow to map each point of the patient anatomy to the dose delivered by a particular MLC shape. This input set was designed to provide each randomly sampled sub-volume with the necessary global information to deduce its dose mapping, and—in these proof-of-principle results on prostate IMRT—led to excellent agreement to the clinical dose. Nonetheless, identifying a minimal set of inputs able to model the underlying problem will be a major part of our future work. As an initial investigation on a small subset of the available patient data, we trained four different networks starting with the 3D MLC mask and electron density of the patient and progressively added the next input as listed in figure 3. A random split of 11/2/12 patients for training, validation and testing was done while the hyperparameters of the main network were used. Table 1 shows the comparison of the predicted doses against the ground-truth data for each one of these networks. Networks 1–3 show a small improvement in individual gamma pass rates with decreasing variance as more inputs are used. This implies that the added inputs in 2 and 3 (distance from source and the center beamline respectively) do not add much more information that is not implicitly extracted from the mask and density. In contrast, the addition of the radiological depth (Network 4) boosts the predictive performance of the network leading to good agreement with the true doses.

These initial findings on a smaller dataset, suggest that depending on the application the proper set of inputs should be identified in order to avoid redundancy in the framework but also to increase calculation performance with less pre/post processing needed per input. Finally, during our experiments we observed

Table 1. Performance of the trained networks with various input combinations for a smaller training set. The inputs correspond to the order specified in figure 3. The relative difference between clinical and predicted dose distributions is shown in the third and fourth column for individual segments and total plan respectively. Similarly, the 3%/3 mm gamma pass rate between the distributions is shown in the fifth and sixth column. The details on the calculation of the statistics are described in section 2.3.

NetId	Inputs	Segment diff (%)	Plan diff (%)	Segment gamma (%)	Plan gamma (%)
1	a, e	-1.3 ± 5.8	0.8 ± 3.1	93.7 ± 8.7	91.6 ± 9.6
2	a, b, e	0.2 ± 5.3	2.9 ± 3.0	94.1 ± 7.8	91.4 ± 10.1
3	a-c, e	-0.5 ± 5.1	2.7 ± 2.8	94.4 ± 7.0	91.9 ± 9.7
4	a-e	-0.2 ± 3.5	2.8 ± 0.5	98.3 ± 2.5	99.7 ± 0.2

that training a similar network while using the whole patient volume could not yield accurate dose prediction irrespective of the input set. In theory though, the different factors including the radiological depth should be inferrable from an appropriately large patient volume while possibly demanding larger datasets.

For these initial proof-of-concept results we exclusively used patients that were treated with five beam IMRT plans. Given the nature of the inputs, the produced network is specialized on this particular treatment site and setup. The endpoint of the DeepDose framework is to be able to yield accurate predictions in multiple treatment sites, involving various anatomies and MLC shapes. Moreover, for coupling to a treatment planning system, different calculation modes ranging from MLC shapes to elementary beam elements might need to be supported. These features demand for more generalizable network architecture/input that will be able to support diverse anatomies and beams at arbitrary gantry/collimator angles. In practise this becomes computationally challenging in terms of possible anatomy/shape combinations that have to be encapsulated by the network. This is especially the case for the dynamic Volumetric Modulated Arc Therapy (VMAT) where—depending on the optimization approach—a treatment plan consists of at least 120 control points per arc and typically features a wide variety of MLC shapes. We are now investigating ways to approach this with more general-purpose networks.

In the context of MRI-guidance, we are looking into applying this framework on MR-linac treatments. New networks that incorporate the effects of the 1.5 T transverse magnetic field of the system into the dose calculations have to be developed, particularly in anatomical cases with varying tissue interfaces (Raaijmakers *et al* 2007). Finally, a longer term goal is to integrate DeepDose in our adaptive sequencing planning system specifically built for the purpose of adaptive MR-linac treatments (Kontaxis *et al* 2015). This coupling would allow for on-the-fly calculation of new segments within the open optimization loop during intrafraction adaptive workflows.

5. Conclusion

We present the DeepDose framework, a proof-of-concept deep learning pipeline able to infer the dose of an arbitrary MLC shape into a given patient anatomy by utilizing physics-based quantities. This is demonstrated in prostate radiotherapy patients previously treated in our clinic split into independent training, validation and testing datasets. We train a deep convolutional network at the clinical 3 mm³ grid spacing. This yields excellent agreement between the predicted segment doses and the ground-truth doses generated with a state-of-the-art Monte Carlo dose engine, passing the clinical QA gamma tests. The framework is able to calculate each 3D segment dose in sub-second timescales, matching the time-consuming Monte Carlo calculations at a low statistical uncertainty. We are now working on generalizing the framework towards a broader scope of applications and investigate its use in online adaptive workflows, in which fast dose calculations are needed.

Acknowledgments

The authors would like to thank Elekta AB, Stockholm, Sweden for providing some of their research software tools.

References

- Ahnesjö A 1989 Collapsed cone convolution of radiant energy for photon dose calculation in heterogeneous media *Med. Phys.* **16** 577–92
- Arnfield M R, Siantar C H, Siebers J, Garmon P, Cox L and Mohan R 2000 The impact of electron transport on the accuracy of computed dose *Med. Phys.* **27** 1266–74

- Barragán-Montero A M, Nguyen D, Lu W, Lin M-H, Norouzi-Kandalan R, Geets X, Sterpin E and Jiang S 2019 Three-dimensional dose prediction for lung IMRT patients with deep neural networks: robust learning from heterogeneous beam configurations *Med. Phys.* **46** 3679–91
- Bortfeld T R, Kahler D L, Waldron T J and Boyer A L 1994 X-ray field compensation with multileaf collimators *Int. J. Radiat. Oncol. Biol. Phys.* **28** 723–30
- Çiçek Ö, Abdulkadir A, Lienkamp S S, Brox T and Ronneberger O 2016 3D U-Net: Learning dense volumetric segmentation from sparse annotation ed Ourselin S, Joskowicz L, Sabuncu M R, Unal G and Wells W *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2016* (Berlin: Springer) pp 424–32
- Gibson E *et al* 2018 Niftynet: a deep-learning platform for medical imaging *Comput. Methods Programs Biomed.* **158** 113–22
- Henke L *et al* 2018 Magnetic resonance image-guided radiotherapy (MRIGRT): A 4.5-year clinical experience *Clin. Oncol.* **30** 720–727
- Hissoiny S, Raaijmakers A J, Ozell B, Despres P and Raaymakers B W 2011 Fast dose calculation in magnetic fields with GPUMCD *Phys. Med. Biol.* **56** 5119–29
- Kontaxis C, Bol G H, Lagendijk J J W and Raaymakers B W 2015 Towards adaptive IMRT sequencing for the MR-linac *Phys. Med. Biol.* **60** 2493–509
- Kontaxis C, Bol G H, Stemkens B, Glitzner M, Prins F M, Kerkmeijer L G W, Lagendijk J J W and Raaymakers B W 2017 Towards fast online intrafraction replanning for free-breathing stereotactic body radiation therapy with the MR-linac *Phys. Med. Biol.* **62** 7233–48
- Krieger T and Sauer O A 2005 Monte Carlo- versus pencil-beam-/collapsed-cone-dose calculation in a heterogeneous multi-layer phantom *Phys. Med. Biol.* **50** 859–68
- Mahmood R, Babier A, McNiven A, Diamant A and Chan T C Y 2018 Automated treatment planning in radiation therapy using generative adversarial networks *Proc. of the 3rd Machine Learning for Conf. Vol. 85 of Proc. of Machine Learning Research, PMLR (Palo Alto, California)* ed Doshi-Velez F, Fackler J, Jung K, Kale D, Ranganath R, Wallace B and Wiens J pp 484–99
- Mohan R, Chui C and Lidofsky L 1986 Differential pencil beam dose computation model for photons *Med. Phys.* **13** 64–73
- Nguyen D, Long T, Jia X, Lu W, Gu X, Iqbal Z and Jiang S 2019 A feasibility study for predicting optimal radiation therapy dose distributions of prostate cancer patients from patient anatomy using deep learning *Sci. Rep.* **9** 1076
- Raaijmakers A J E, Raaymakers B W, van der Meer S and Lagendijk J J W 2007 Integrating a MRI scanner with a 6 MV radiotherapy accelerator: impact of the surface orientation on the entrance and exit dose due to the transverse magnetic field *Phys. Med. Biol.* **52** 929–39
- Rogers D W 2006 Fifty years of Monte Carlo simulations for medical physics *Phys. Med. Biol.* **51** R287–301
- Shepard D M, Earl M A, Li X A, Naqvi S and Yu C 2002 Direct aperture optimization: a turnkey solution for step-and-shoot IMRT *Med. Phys.* **29** 1007–18
- Tian Z, Shi F, Folkerts M, Qin N, Jiang S B and Jia X 2015 A GPU OpenCL based cross-platform Monte Carlo dose calculation engine (goMC) *Phys. Med. Biol.* **60** 7419–35
- Werensteijn-Honingh A M *et al* 2019 Feasibility of stereotactic radiotherapy using a 1.5 T MR-linac: multi-fraction treatment of pelvic lymph node oligometastases *Radiother. Oncol.* **134** 50–4