Radiotherapy and Oncology 170 (2022) 169–175

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Original Article

Multi-center analysis of machine-learning predicted dose parameters in brachytherapy for cervical cancer

Dominique Reijtenbagh [a,*], Jérémy Godart [a], Astrid de Leeuw [b], Yvette Seppenwoolde [a], Ina Jürgenliemk-Schulz [b], Jan-Willem Mens [a], Remi Nout [a], Mischa Hoogeman [a,c]

[a] Department of Radiation Oncology, Erasmus MC Cancer Institute, Rotterdam; [b] Department of Radiation Oncology, University Medical Centre Utrecht; and [c] Department of Medical Physics & Informatics, HollandPTC, Delft, The Netherlands

ARTICLE INFO

Article history:
Received 1 December 2021
Received in revised form 14 February 2022
Accepted 16 February 2022
Available online 24 February 2022

Keywords:
Cervical cancer
Image guided brachytherapy
Overlap volume histogram
DVH prediction

ABSTRACT

Background and purpose: Image-guided adaptive brachytherapy (IGABT) is a key component in the treatment of cervical cancer, but the nature of the clinical workflow makes it vulnerable to suboptimal plans, as the theoretical optimal plan depends heavily on organ configuration. Patient anatomy-based quality-assurance (QA) with overlap volume histograms (OVHs) is a promising tool to detect such suboptimal plans, and in this analysis its suitability as a multi-institutional clinical QA tool is investigated.
Materials and methods: A total of 223 plans of 145 patients treated in accordance with the current state-of-the-art IGABT protocols from UMC Utrecht (UMCU) and Erasmus MC (EMC) were included. Machine-learning models were trained to predict dose D_{2cm3} to bladder, rectum, sigmoid and small bowel with the help of OVHs. For this strategy, points are sampled on the organs-at-risk (OARs), and the distances of the sampled points to the target are computed and combined in a histogram. Machine-learning models can then be trained to predict dose-volume histograms (DVHs) for unseen data. Single-center model robustness to needle use and applicator type and multi-center model translatability were investigated. Performance of models was assessed by the difference between planned (clinical) and predicted D_{2cm3} values.
Results: Intra-validation of UMCU data demonstrated OVH model robustness to needle use and applicator type. The model trained on UMCU data was found to be robust within the same protocol on EMC data, for all investigated OARs. Mean squared error between planned and predicted D_{2cm3} values of OARs ranged between 0.13 and 0.40 Gy within the same protocol, indicating model translatability. For the former protocol cohort of Erasmus MC large deviations were found between the planned and predicted D_{2cm3} values, indicating plan deviation from protocol. Mean squared error for this cohort ranged from 0.84 to 4.71 Gy.
Conclusion: OVH-based models can provide a solid basis for multi-institutional QA when trained on a sufficiently strict protocol. Further research will quantify the model's impact as a QA tool.

Brachytherapy (BT) is an integral component of the treatment of locally-advanced cervical cancer (LACC) patients [1,2]. Over the past years the complexity of BT treatment has increased, and the standard has shifted towards image-guided adaptive brachytherapy (IGABT) [3,4]. In line with this development, the additional use of interstitial needles to intracavitary implants (IC + IS) instead of solely intracavitary implants (IC) has increased [5]. IC + IS techniques allow better shaping of the dose distribution, but can complicate the clinical workflow and the creation of optimal treatment plans [6].

Added to the increased complexity of BT treatments, the adaptive nature of the treatments makes IGABT workflows bound to time pressure, which can cause deterioration of treatment plan quality. Apart from adequate target coverage being the primary goal in cervical IGABT planning, the lowest dose to the maximally exposed 2 cm^3 of individual Organ At Risk (OAR) volumes (D_{2cm3}) is constrained and defined invariant to tumor size or to patient anatomy [7]. Depending on the patient and the use of interstitial needles, these constraints can be more easily met. To ensure a minimum quality of the treatment plans, the use of dummy-runs has been reported to improve familiarity with the protocol and to homogenize results among institutes [8–11]. Cohort analyses and comparisons of dose-volume parameters with other institutes can give insight in an institute's plan quality and possibly provide guidance for general improvement. However, this knowledge

* Corresponding author at: Erasmus MC Cancer Institute, Department of Radiation Oncology, Dr. Molewaterplein 40, 3015 GD Rotterdam, The Netherlands.
E-mail address: d.reijtenbach@erasmusmc.nl (D. Reijtenbagh).

https://doi.org/10.1016/j.radonc.2022.02.022
0167-8140/

cannot be utilized to recognize individual suboptimal plans and guide planning for individual patient cases.

As a solution, knowledge-based quality-assurance (QA) tools have the potential to assess the quality of a treatment plan. By utilizing prior high-quality plans, predictions for treatment plans of current patients can be made, while taking into account the specific anatomy of the individual patient [12]. Such QA tools can be trained on gold-standard datasets, and prospectively identify suboptimal plans that might benefit from further optimization. Theoretically, these tools could even be exchanged between institutes, provided these institutes employ a similar BT treatment protocol and equipment. Previous studies have reported the use of anatomy-based models to predict dose-volume parameters of external beam treatments [13–16]. As an example, Wang et al. proposed the use of Overlap Volume Histograms (OVHs) to prospectively detect outliers in terms of plan quality of external beam radiotherapy (EBRT) for prostate cancer patients [17]. With this methodology, points are sampled on an organ of interest, after which the distances of all points to the target are computed and combined in a histogram. The OVHs are then a suitable basis to predict Dose Volume Histograms (DVHs) with machine-learning strategies. For BT planning of LACC-patients, it has already been demonstrated that D_{2cm3} values of OARs depend heavily on the proximity of organs-at-risk (OARs) [14], but no report has been made to validate a knowledge-based QA tool in a multi-center setting.

The purpose of this study was to investigate the potential of OVH-based QA for cervical cancer patients in a clinical multi-institutional BT setting. Model robustness to needle use and applicator type was investigated based on the patient cohort of one institute, and multi-center translatability was validated with multi-center data.

Materials and methods

LACC-patients treated with HDR BT were included from UMC Utrecht (UMCU) and Erasmus MC Rotterdam (EMC). The study was approved by the medical ethics committee under number MEC-2021-0337. A total of 60 patients (120 plans) from UMCU and 14 patients (32 plans) from EMC (EMC-2) were included, treated in accordance with EMBRACE II protocol. Additionally, 71 MR-guided BT pre-EMBRACE II plans (71 patients) from Erasmus MC were included (EMC-1), treated with one or more fractions of MR-guided BT between 2015 and 2018 prior to the implementation of EMBRACE II protocol in the clinic (EMC-1). Details of the dosimetric aims for UMCU and both EMC cohorts can be found in Table 1. Treatment planning within the EMC-1 cohort had the same cumulative aims for the high-risk CTV (HRCTV) coverage and OAR D_{2cm3} constraints, with an exception for the small bowel (aim of 60 Gy instead of 70 Gy within EMBRACE II). This protocol did not include further planning aims from the EMBRACE II protocol, such as the use of an intermediate-risk CTV (IRCTV), and was therefore less strict. Further details on the planning for the EMC-1 cohort can be found in the work by Oud et al. [18].

Table 2
Treatment characteristics of included patients. Application and applicator types are grouped based on data from all included fractions per patient.

	UMCU	EMC-2	EMC-1
Number of plans	*n* = 120	*n* = 32	*n* = 71
Number of patients	60	14	71
Number of BT fractions given per patient			
4 fractions	60	4	3
3 fractions	0	10	68
Application type *per patient*			
IC	19	0	55
IC + IS	31	32	16
Mix fractions	10	0	0
Applicator type *per patient*			
Ovoid	24	32	71
Ring	28	0	0
Mix fractions	8	0	0

Applicators and application type (IC versus IC + IS) were registered. About 46% of the UMCU patients was treated with a ring-based applicator only, 40% with a ovoid-based applicator, and the remainder with a combination of both. Patients within the EMC-1 cohort were less often treated with additional needles (IC + IS) compared to patients within the EMC-2 cohort (23% versus 100%). Full treatment characteristics are shown in Table 2. Plots of distributions of planned D_{2cm3} of OARs (after scaling for coverage) and planning volumes of OARs and targets can be found in Appendix A.

Planning BT MR images, structure sets and the three-dimensional dose distributions were collected for all patients. OVHs between delineated OARs and HRCTV and between the delineated OARs and the IRCTV were used to quantify patient anatomy, as discussed above and analogous to previously described in literature [19]. Scaling of the full physical dose distributions was performed to account for differences in fractionation schemes. In clinical practice the BT fractions are given in addition to a 25-fraction external beam radiotherapy course of 1.8 Gy (UMCU and EMC-2), or 23-fraction course of 2 Gy (EMC-1). BT was administered in four fractions (UMCU) or generally three fractions (EMC). Full dose distributions were scaled to $D_{90\%}$ of the HRCTV to 7.5 Gy, for inter-patient and inter-institute comparison. This physical dose scaling was chosen to achieve a total biological dose of the HRCTV of 88 Gy (α/β = 10) given in 4 BT fractions. Dose-volume histograms (DVHs) were constructed for all OARs based on the scaled dose distributions. OVHs were constructed for all OARs separately with respect to HRCTV or IRCTV. Points were sampled on the OAR of interest, after which the closest distance to the target of interest was calculated for each sampled point. These distances were then combined into a histogram. The full dataset com-

Table 1
Overview of the dosimetric criteria for the full treatment (external beam radiotherapy and brachytherapy) for the EMBRACE II cohorts (UMCU and EMC-2) and pre-EMBRACE II cohort (EMC-1), α/β = 10 Gy is used for the HRCTV and α/β = 3 Gy for OARs.

Dosimetric criteria		UMCU and EMC-2		EMC-1	
Structure		*Limit EQD_{2Gy}*	*Aim EQD_{2Gy}*	*Limit EQD_{2Gy}*	*Aim EQD_{2Gy}*
HRCTV	$D_{90\%}$	85 Gy	>90 < 95 Gy	85 Gy	>90 < 95 Gy
IRCTV	$D_{98\%}$	–	>60 Gy	–	–
Bladder	D_{2cm3}	90 Gy	<80 Gy	90 Gy	<80 Gy
Rectum	D_{2cm3}	75 Gy	<65 Gy	75 Gy	<65 Gy
Sigmoid	D_{2cm3}	75 Gy	<70 Gy	75 Gy	<70 Gy
Small Bowel	D_{2cm3}	75 Gy	<70 Gy	75 Gy	<60 Gy

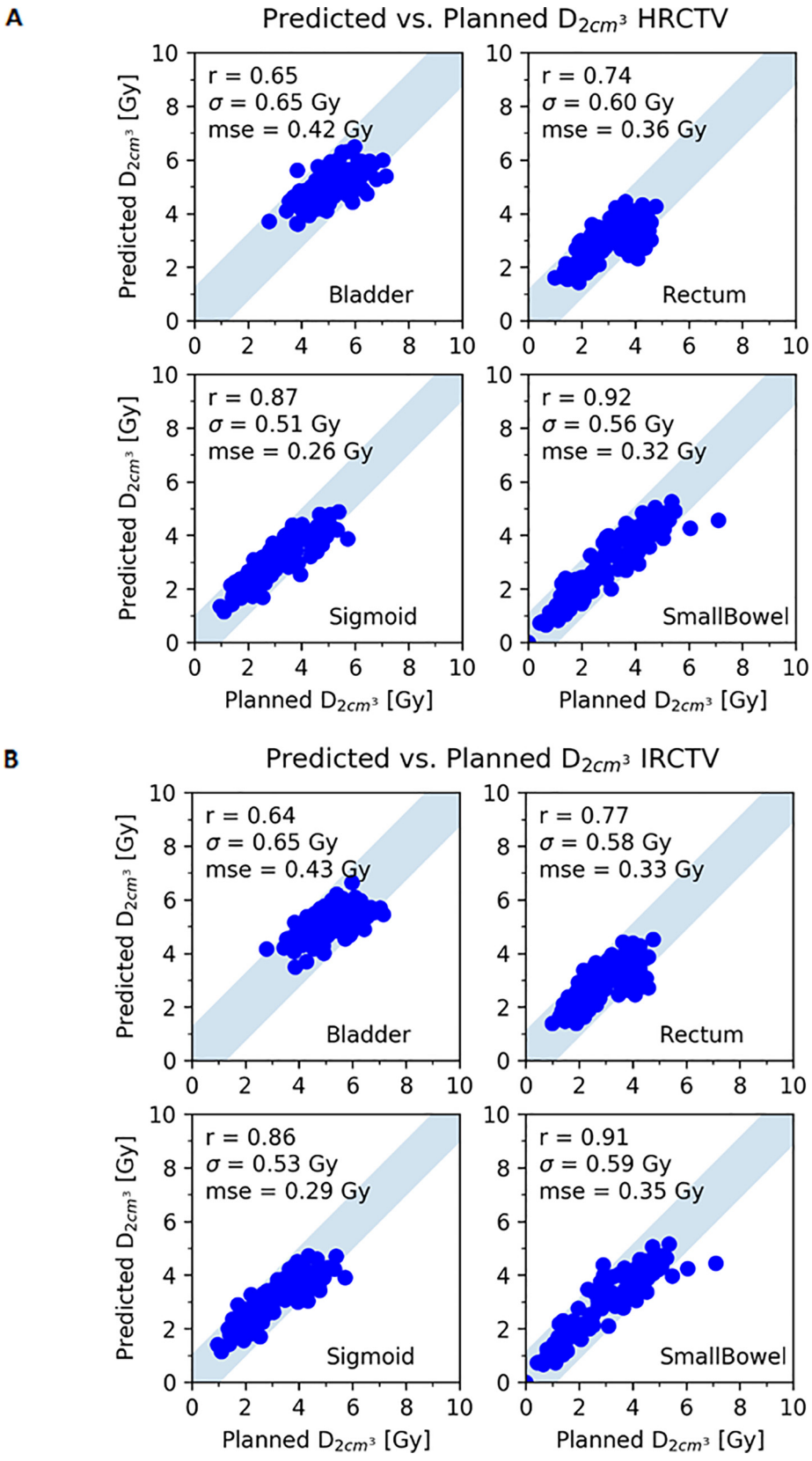

Fig. 1. Results of leave-one-out validation of UMCU data. OVHs and DVHs are constructed for all patients and all OARs with respect to the HRCTV (A) and IRCTV (B), and used to construct a random-forest model. The planned and predicted D_{2cm3} values are plotted here, together with the Pearson's correlation coefficient, the standard deviation of the mean and the mean squared error. The shaded area depicts the 95% Confidence Interval.

Table 3
Model performances for different data splits. Leave-one-out validation was used to determine the prediction interval for outlier detection. For all predicted values of the OARs the mean squared error is calculated, as well as the number of outliers outside of the prediction interval.

OVH with HRCTV		Bladder		Rectum		Sigmoid		Small Bowel	
Intra-validation	*Leave-one-out*	mse [Gy]	outliers (%)	mse [Gy]	outliers (%)	mse [Gy]	outliers (%)	mse [Gy]	outliers (%)
UMCU IC		0.55	2/38 (5)	0.43	2/38 (5)	0.28	2/38 (5)	0.61	2/38 (5)
UMCU Ovoid		0.56	2/48 (4)	0.41	4/48 (8)	0.34	1/48 (2)	0.35	3/48 (6)
UMCU		0.42	6/120 (5)	0.36	7/120 (6)	0.26	6/120 (5)	0.32	4/120 (3)
Inter-validation									
Training	*Testing*								
UMCU IC	UMCU IC + IS	0.45	3/62 (5)	0.38	3/62 (5)	0.69	11/62 (18)	0.53	3/62 (5)
UMCU Ovoid	UMCU Ring	0.55	2/56 (4)	0.58	6/56 (11)	0.37	4/56 (7)	0.45	3/56 (5)
UMCU	EMC-2	0.40	2/32 (6)	0.13	0/32 (0)	0.30	2/32 (6)	0.37	2/32 (6)
UMCU	EMC-1	4.69	30/71 (42)	3.46	33/71 (46)	1.35	25/70 (36)	0.84	10/56 (18)
UMCU	EMC-1 IC + IS	2.29	5/16 (31)	0.96	3/16 (19)	0.92	5/16 (31)	1.31	3/16 (19)
UMCU	EMC-1 IC	5.46	25/55 (45)	4.20	30/55 (55)	1.52	20/54 (37)	0.69	7/40 (18)

prised of an OVH and a DVH for each included plan, and for each combination of OAR and target structure.

After data pre-processing, random-forest models were constructed and tested for different sub-dataset combinations. Random forest networks were selected for their relative robustness to overfitting [20]. The training and testing was performed for all OARs (bladder, rectum, sigmoid and small bowel) and all target structures (HRCTV and IRCTV), if available. A trained model can then predict a full DVH for an unseen OVH of one type of OAR and target structure. Evaluations were done based on different data stratifications, first to investigate the robustness of the model within one center, then to investigate the applicability of the model to data from other center and other protocol. An overview of data collection, OVH construction and the modeling is displayed in Appendix B.

Data stratifications to investigate needle and applicator type robustness were:

- Training with IC UMCU plans (38), testing with UMCU IC + IS plans (62)
- Training with UMCU ovoid-based plans (48), testing with UMCU ring-based plans (56)

Data stratifications to investigate multi-center model translatability were:

- Training with all UMCU plans (120), testing with EMC-1 plans (71)
- Training with all UMCU plans (120), testing with EMC-2 plans (32)

For each stratification, each of the following steps were performed:

1. A leave-one-out analysis was performed on the training dataset to set the prediction interval (intra-validation phase)
2. A random-forest model was trained on the full training set and applied to the testing data (inter-validation phase)

Within the intra-validation phase for each iteration of the leave-one-out analysis, OVH data was reduced with Principal Component Analysis (PCA), after which a random-forest model was fitted to the remaining leave-one-out data. The model was then applied to the left-out OVH to predict the D_{2cm3} value, and the planned (clinically delivered) D_{2cm3} and predicted D_{2cm3} values were registered. The mean squared error (mse), the standard deviation of the mean (σ) and the Pearson correlation coefficient r were calculated for all planned and predicted D_{2cm3} values. To evaluate the use of OVHs in a QA setting, the prediction interval was defined as the 95% confidence interval of the difference between planned D_{2cm3} and predicted D_{2cm3} values. This interval reflects the natural distribution of the training dataset.

Within the inter-validation phase, the full training dataset was used to train a random-forest model after PCA application. The trained model was consequently applied to the OVH data of the testing set to find the predicted D_{2cm3} values. The difference between planned D_{2cm3} and predicted D_{2cm3} values was registered, after which the mse was calculated. Predictions outside of the prediction interval (outliers) were identified, as defined by the training set in the intra-validation phase. All code was implemented in Python 3.5.2 and DICOM handling was performed with in-house developed software Matterhorn. Data preprocessing and machine-learning was performed with Scikit-learn version 0.19.2 [21].

For all plans within the EMC-1 and EMC-2 cohort, protocol compliance was investigated. Clinical records about all plans regarding protocol compliance were collected. All plan records were categorized in three categories: protocol compliant, protocol non-compliant and unknown. The first category comprised plans where no problems were reported and the plans were mentioned to be adequate, or compliant. In the second category plans were placed that had reports of trade-offs that were considered during planning, such as OAR constraints that posed limitations to protocol compliance and clinical difficulties during treatment (e.g. problems with implantation). For some plans no available records or remarks could be retrieved. Plans having one or more OARs with outliers outside of the prediction interval were then investigated to find possible explanations for their deviations.

Results

Intra-validation results of the full UMCU dataset are displayed in Fig. 1A and 1B and Table 3 for all OARs with respect to HRCTV and IRCTV. The range of mse for all OARs varied between 0.26 and 0.43 Gy. Models based on HRCTV and IRCTV showed comparable σ, correlation r and mse for each target-organ pair. The least variation in planned dose was found for the bladder, as well as the lowest Pearson's correlation of all OARs. The light blue regions in the figures display the prediction interval. Similarly, results after stratification of the UMCU data based on application type (IC-based plans) and applicator type (ovoid-based) can also be found in Table 3 and in Appendix B and C. Stratification of the UMCU data based on applicator type (ovoid-based for training and ring-based for testing) showed comparable mse for bladder and sigmoid, but a slightly higher mse was observed for rectum and small bowel. Training on IC-based plans and validation with IC + IS-based plans resulted in similar mse for bladder and rectum for the valida-

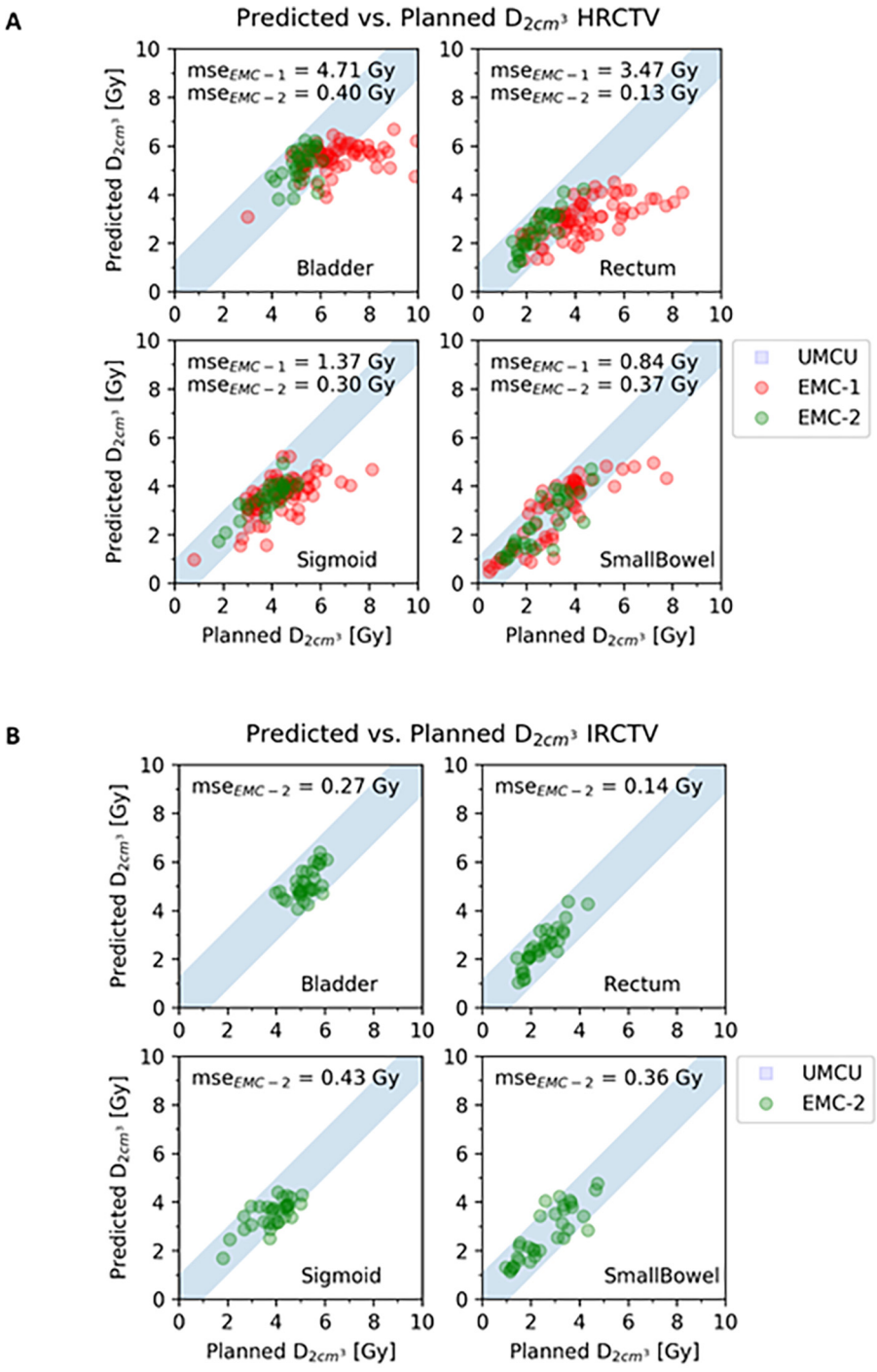

Fig. 2. Results of training with UMCU data, validation with EMC-2 and EMC-1 data. OVHs and DVHs are constructed for all patients and all OARs with respect to the HRCTV (A) and IRCTV (B), and used to construct a random-forest model. The planned and predicted D_{2cm3} values are plotted here, together with the mean squared error. The shaded area depicts prediction interval based on the leave-one-out validation of the UMCU data. EMC-1 patients were not planned with an IRCTV, therefore no data is visible for them in B.

tion data, but higher mse was found for sigmoid and bowel. The number of outliers outside of the prediction interval was comparable for training and testing, and is also reported in Table 3.

Results based on institute stratification (UMCU for training versus EMC-2 and EMC-1 for testing) are shown in Fig. 2A and B. No EMC-1 data points are displayed in Fig. 2B, as no IRCTV was delineated and used for planning within this cohort. Clear differences can be observed for all OARs between cohorts in Fig. 2A – the mse for all OARs of the EMC-2 cohort varied between 0.13 and 0.40 Gy, whilst the values for the EMC-1 cohort varied between 0.84 and 4.69 Gy. When investigating the clinical records of plan compliance for the EMC-1 and EMC-2 plans in Table 4, the majority of the plans within the EMC-2 cohort (81%) did not have an outlier for one or more of the OARs. The outliers that were found did correspond in 80% of the cases with reported difficulties to comply to the protocol. Within the EMC-1 cohort a far smaller percentage of

Table 4
Classification of plans from EMC-2 and EMC-1 with OVH-based predicted outliers. Clinical records were investigated to find possible reporting of planning difficulties that could demonstrate awareness of sub-optimality of these plans. Plans were marked as protocol compliant if a record did not contain such information, and with unknown when no report could be found.

OVH-based predicted outliers for EMC patients		All EMC-1 (n = 71)	EMC-1 IC (n = 55)	EMC-1 IC + IS (n = 16)	All EMC-2 (n = 32)
No outliers for all OARs	Protocol compliant	11	7	4	13
	Protocol non-compliant	8	6	2	5
	Unknown	0	0	0	8
Outlier (1 of OARs)	Protocol compliant	10	6	4	1
	Protocol non-compliant	7	5	2	3
	Unknown	1	1	0	2
Outliers (>1 of OARs)	Protocol compliant	2	2	0	0
	Protocol non-compliant	29	26	3	0
	Unknown	3	2	1	0

the plans (27%) did not have any predicted outliers for the OARs, and 48% of the plans had more than one outlier for the OARs. IC-based plans were also more often classified as outlier (76%) than IC + IS-based plans (63%).

Discussion

In this paper we presented a machine-learning based method to predict D_{2cm3} values of OARs in MR-based BT planning for cervical cancer patients in a multi-center setting. The models were found to be robust to application and applicator details within a single-center setting, and have proven to be translatable to a multi-center setting. Overall, inter-validation results were comparable to results reported in a prior study utilizing geometric features to depict patient anatomy [14], and slightly worse than another small (n = 20) study [15]. Both these studies reported results for CT-based planning, where target delineation is shown to differ from a MR-based setting [22]. Additionally, both publications described single-center studies.

For this study, dose distributions were compared based on physical dose, not based on EQD_{2Gy} dose. Taking this into account, it is expected that a direct model translation to Pulsed Dose Rate (PDR) type plans is not possible. PDR models will therefore most likely have to be constructed and validated separately. Furthermore, D_{2cm3} dose constraints for OARs were considered for this analysis as it is most widely used for BT planning of cervical cancer patients. Other dosimetric parameters were not considered in this work, but could be of interest in future studies.

Models based on HRCTV showed comparable results to IRCTV-based models for the EMC-2 cohort. Although planning aims for IRCTV are different than for HRCTV this is not entirely surprising, as the IRCTV is largely an isotropic expansion of the HRCTV. The OAR would be closer to the IRCTV, resulting in a shifted but similarly shaped OVH, yet the organ's DVH would not change. With a consistent shift of OVH, this has no expected consequences for model performance.

The EMC-1 cohort contained more plans that were marked as outliers than the EMC-2 cohort. EMC-1 patients were treated with a protocol similar to EMC-2 patients, containing the same D_{2cm3} aims and goals for all OARs except small bowel. However, needle use was less frequent for EMC-1 patients. Therefore, it could be argued that the high number of outliers was caused by the majority of patients being treated with only IC-based plans. Interstitial needles allow better modulation of the dose, and can improve target coverage while maintaining OAR sparing [6]. Patients within the UMCU cohort were preselected to be treated with IC-only plans if the anatomy and tumor size were expected to be favorable, but for EMC-1 patients this was not the case.

However, even for IC + IS patients within the EMC-1 cohort an unexpectedly high percentage of plans was classified as outliers. Treatment remarks concerning protocol compliance for these plans were investigated to find explanations for this. As an example, it was found that two outliers within this cohort were identified to originate from plans for patients suffering from tumor infiltration in the bladder or rectum, explaining the unexpectedly high dose to these organs. For nine outliers clinical records reported a higher accepted dose to specific OARs to obtain sufficient target coverage, possibly explained by suboptimal needle or implant position. Yet, the remainder of plans were reported to be protocol-compliant and still classified as outlier. This is particularly interesting from a QA point-of-view, as these plans might not have been the most optimal plans for the patients' anatomies, and could therefore have benefited from further optimization. It appears that a protocol ought to be sufficiently strict to be trained on to predict D_{2cm3} values for any OAR of new patients, which was not the case for the EMC-1 patients. A less strict protocol leaves too many degrees of freedom and results in suboptimal dose distributions.

In this study clinical data was used for both model-building and validation, implicitly representing trade-offs and other decision-making present during BT planning. Fully-automated planning could improve consistency of planning, and it has been demonstrated for LACC patients that fully-automated planning can provide high-quality plans [18,23–25]. Fully-automated planning could generate high-quality datasets to improve the prediction accuracy of OVH models, but a necessity for QA will remain to validate the quality and optimality of also fully-automated clinical plans [17].

To summarize, this work presents a method to predict D_{2cm3} values in a complex multicenter setting based on clinically-used treatment data. The D_{2cm3} predictions based on the individual patient geometry could encourage planners to improve planning beyond dosimetric constraints as dictated by protocol, decreasing normal tissue irradiation. The construction of the OVHs and the subsequent prediction of the achievable D_{2cm3} value is performed in mere seconds without the necessity of applicator reconstruction, which makes it a feasible tool for clinical practice. Models can be swiftly trained on different patient cohorts or treatment characteristics, implying they could be used to detect outliers in multi-center study settings without the need of sharing patient-sensitive data. Future studies will quantify the use of OVH-based models as a QA tool by re-planning patients outside of the prediction interval to improve plan quality, and with this quantify its potential to better personalize BT planning.

Conclusion

In this paper an OVH-based approach is presented to predict D_{2cm3} values in BT planning for LACC-patients for all relevant OARs in a multifactorial multi-center setting. The constructed model based on single-center data can identify plans from a different institute, for which trade-offs between OAR sparing and target coverage had to be made.

Conflicts of interest

This work was in part funded by a research grant of Elekta AB (Stockholm, Sweden). The funders had no role in study design, data collection and analysis, and decisions on preparation of the manuscript. Erasmus MC Cancer Institute also has research collaborations with Accuray Inc, Sunnyvale, USA and Varian Medical Systems Particle Therapy GmbH & Co. KG, Troisdorf, Germany. Professor Hoogeman reports a membership of the advisory board Accuray, Sunnyvale, USA.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.radonc.2022.02.022.

References

[1] Sturdza A, Pötter R, Fokdal LU, Haie-Meder C, Tan LT, Mazeron R, et al. Image guided brachytherapy in locally advanced cervical cancer: Improved pelvic control and survival in RetroEMBRACE, a multicenter cohort study. Radiother Oncol 2016;120:428–33.
[2] Pötter R, Tanderup K, Schmid MP, Jürgenliemk-Schulz I, Haie-Meder C, Fokdal LU, et al. MRI-guided adaptive brachytherapy in locally advanced cervical cancer (EMBRACE-I): a multicentre prospective cohort study. Lancet Oncol 2021;22:538–47.
[3] Haie-Meder C, Pötter R, Van Limbergen E, Briot E, De Brabandere M, Dimopoulos J, et al. Recommendations from Gynaecological (GYN) GEC-ESTRO Working Group☆ (I): concepts and terms in 3D image based 3D treatment planning in cervix cancer brachytherapy with emphasis on MRI assessment of GTV and CTV. Radiother Oncol 2005;74:235–45.
[4] Pötter R, Haie-Meder C, Limbergen EV, Barillot I, Brabandere MD, Dimopoulos J, et al. Recommendations from gynaecological (GYN) GEC ESTRO working group (II): concepts and terms in 3D image-based treatment planning in cervix cancer brachytherapy-3D dose volume parameters and aspects of 3D image-based anatomy, radiation physics, radiobiology. Radiother Oncol 2006;78:67–77.
[5] de Boer P, Jürgenliemk-Schulz IM, Westerveld H, de Leeuw AAC, Dávila-Fajardo R, Rasch CRN, et al. Patterns of care survey: radiotherapy for women with locally advanced cervical cancer. Radiother Oncol 2017;123:306–11.
[6] Fokdal L, Tanderup K, Hokland SB, Røhl L, Pedersen EM, Nielsen SK, et al. Clinical feasibility of combined intracavitary/interstitial brachytherapy in locally advanced cervical cancer employing MRI with a tandem/ring applicator in situ and virtual preplanning of the interstitial component. Radiother Oncol 2013;107:63–8.
[7] Pötter R, Tanderup K, Kirisits C, de Leeuw A, Kirchheiner K, Nout R, et al. The EMBRACE II study: The outcome and prospect of two decades of evolution within the GEC-ESTRO GYN working group and the EMBRACE studies. Clin Transl Radiat Oncol 2018;9:48–60.
[8] Wortman BG, Astreinidou E, Laman MS, van der Steen-Banasik EM, Lutgens LCHW, Westerveld H, et al. Brachytherapy quality assurance in the PORTEC-4a trial for molecular-integrated risk profile guided adjuvant treatment of endometrial cancer. Radiother Oncol 2021;155:160–6.
[9] Kirisits C, Federico M, Nkiwane K, Fidarova E, Jürgenliemk-Schulz I, de Leeuw A, et al. Quality assurance in MR image guided adaptive brachytherapy for cervical cancer: final results of the EMBRACE study dummy run. Radiother Oncol 2015;117:548–54.
[10] Seppenwoolde Y, Assenholt MS, Georg D, Nout R, Tan LT, Rumpold T, et al. Importance of training in external beam treatment planning for locally advanced cervix cancer: report from the EMBRACE II dummy run. Radiother Oncol 2019;133:149–55.
[11] De Leeuw AAC, Nout RA, Van Leeuwen RGH, Mans A, Verhoef LG, Jürgenliemk-Schulz IM. Implementation of state-of-the-art (chemo)radiation for advanced cervix cancer in the Netherlands: a quality improvement program. Tech Innovat Patient Supp Radiat Oncol 2019;9:1–7.
[12] Cornell M, Kaderka R, Hild SJ, Ray XJ, Murphy JD, Atwood TF, et al. Noninferiority study of automated knowledge-based planning versus human-driven optimization across multiple disease sites. Int J Radiat Oncol Biol Phys 2020;106:430–9.
[13] Zhu X, Ge Y, Li T, Thongphiew D, Yin F-F, Wu QJ. A planning quality evaluation tool for prostate adaptive IMRT based on machine learning. Med Phys 2011;38:719–26.
[14] Yusufaly TI, Kallis K, Simon A, Mayadev J, Yashar CM, Einck JP, et al. A knowledge-based organ dose prediction tool for brachytherapy treatment planning of patients with cervical cancer. Brachytherapy 2020;19:624–34.
[15] Damato AL, Viswanathan AN, Cormack RA. Validation of mathematical models for the prediction of organs-at-risk dosimetric metrics in high-dose-rate gynecologic interstitial brachytherapy. Med Phys 2013;40:101711.
[16] Shiraishi S, Moore KL. Knowledge-based prediction of three-dimensional dose distributions for external beam radiotherapy. Med Phys 2016;43:378.
[17] Wang Y, Heijmen BJM, Petit SF. Prospective clinical validation of independent DVH prediction for plan QA in automatic treatment planning for prostate cancer patients. Radiother Oncol 2017;125:500–6.
[18] Oud M, Kolkman-Deurloo I-K, Mens J-W, Lathouwers D, Perkó Z, Heijmen B, et al. Fast and fully-automated multi-criterial treatment planning for adaptive HDR brachytherapy for locally advanced cervical cancer. Radiother Oncol 2020;148:143–50.
[19] Kazhdan M et al. A shape relationship descriptor for radiation therapy planning. International conference on medical image computing and computer-assisted intervention. Springer; 2009.
[20] Breiman L. Random forests. Machine Learn 2001;45:5–32.
[21] Kramer O. Scikit-learn. In: Machine learning for evolution strategies. Springer; 2016. p. 45–53.
[22] Viswanathan AN, Erickson B, Gaffney DK, Beriwal S, Bhatia SK, Lee Burnett O, et al. Comparison and consensus guidelines for delineation of clinical target volume for CT- and MR-based brachytherapy in locally advanced cervical cancer. Int J Radiat Oncol Biol Phys 2014;90:320–8.
[23] Lessard E, Pouliot J. Inverse planning anatomy-based dose optimization for HDR-brachytherapy of the prostate using fast simulated annealing algorithm and dedicated objective function. Med Phys 2001;28:773–9.
[24] Mao X, Pineau J, Keyes R, Enger SA. Rapid radiation dose calculations in brachytherapy via deep learning. Int J Radiat Oncol Biol Phys 2020;108:802–12.
[25] Shen C, Gonzalez Y, Klages P, Qin N, Jung H, Chen L, et al. Intelligent inverse treatment planning via deep reinforcement learning, a proof-of-principle study in high dose-rate brachytherapy for cervical cancer. Phys Med Biol 2019;64:115013. https://doi.org/10.1088/1361-6560/ab18bf.