



OPEN ACCESS

EDITED BY
John Varlotto,
Marshall University, United States

REVIEWED BY
Saravanan Kandasamy,
Jawaharlal Institute of Postgraduate
Medical Education and Research
(JIPMER), India
Ramin Jaberi,
Tehran University of Medical
Science, Iran

*CORRESPONDENCE
Jie Fu
fujie74@sjtu.edu.cn

[†]These authors have contributed
equally to this work

SPECIALTY SECTION
This article was submitted to
Radiation Oncology,
a section of the journal
Frontiers in Oncology

RECEIVED 12 June 2022
ACCEPTED 02 August 2022
PUBLISHED 30 August 2022

CITATION
Li Z, Chen K, Yang Z, Zhu Q, Yang X,
Li Z and Fu J (2022) A personalized
DVH prediction model for HDR
brachytherapy in cervical
cancer treatment.
Front. Oncol. 12:967436.
doi: 10.3389/fonc.2022.967436

COPYRIGHT
© 2022 Li, Chen, Yang, Zhu, Yang, Li
and Fu. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

A personalized DVH prediction model for HDR brachytherapy in cervical cancer treatment

Zhen Li^{1†}, Kehui Chen^{2†}, Zhenyu Yang³, Qingyuan Zhu¹,
Xiaojing Yang¹, Zhaobin Li¹ and Jie Fu^{1*}

¹Shanghai Jiao Tong University Affiliated Sixth People's Hospital, Shanghai, China, ²Shuguang Hospital Affiliated to Shanghai University of Traditional Chinese Medicine, Shanghai, China, ³Duke University, Durham, NC, United States

Purpose: Although the knowledge-based dose-volume histogram (DVH) prediction has been largely researched and applied in External Beam Radiation Therapy, it is still less investigated in the domain of brachytherapy. The purpose of this study is to develop a reliable DVH prediction method for high-dose-rate brachytherapy plans.

Method: A DVH prediction workflow combining kernel density estimation (KDE), k-nearest neighbor (kNN), and principal component analysis (PCA) was proposed. PCA and kNN were first employed together to select similar patients based on principal component directions. 79 cervical cancer patients with different applicators inserted was included in this study. The KDE model was built based on the relationship between distance-to-target (DTH) and the dose in selected cases, which can be subsequently used to estimate the dose probability distribution in the validation set. Model performance of bladder and rectum was quantified by $|\Delta D_{2cc}|$, $|\Delta D_{1cc}|$, $|\Delta D_{0.1cc}|$, $|\Delta D_{max}|$, and $|\Delta D_{mean}|$ in the form of mean and standard deviation. The model performance between KDE only and the combination of kNN, PCA, and KDE was compared.

Result: 20, 30 patients were selected for rectum and bladder based on kNN and PCA, respectively. The absolute residual between the actual plans and the predicted plans were 0.38 ± 0.29 , 0.4 ± 0.32 , 0.43 ± 0.36 , 0.97 ± 0.66 , and 0.13 ± 0.99 for $|\Delta D_{2cc}|$, $|\Delta D_{1cc}|$, $|\Delta D_{0.1cc}|$, $|\Delta D_{max}|$, and $|\Delta D_{mean}|$ in the bladder, respectively. For rectum, the corresponding results were 0.34 ± 0.27 , 0.38 ± 0.33 , 0.63 ± 0.57 , 1.41 ± 0.99 and 0.23 ± 0.17 , respectively. The combination of kNN, PCA, and KDE showed a significantly better prediction performance than KDE only, with an improvement of 30.3% for the bladder and 33.3% for the rectum.

Conclusion: In this study, a knowledge-based machine learning model was proposed and verified to accurately predict the DVH for new patients. This model is proved to be effective in our testing group in the workflow of HDR brachytherapy.

KEYWORDS

machine learning, brachytherapy, cervical cancer, dose prediction, radiation oncology

1 Introduction

Cervical cancer is the fourth most common cancer in women globally (1). The treatment of cervical cancer relies on a combination of external radiotherapy and HDR brachytherapy to increase the dose being delivered to the primary tumor (2, 3). Numerous studies have shown that high-dose-rate brachytherapy (HDR-BT) is strongly correlated with patients' survival rates and plays an essential curative role in cervical cancer (4, 5). It allows delivery of highly localized doses to the target and excellent sparing of surrounding organs at risk (OARs). High-quality planning is a critical component in gynecologic BT treatment. However, unlike in external-beam radiation therapy (EBRT), the planning workflow in BT necessitates the collaboration of multidisciplinary teamwork in the shortest possible time to minimize the patient's discomfort and movement. The pressures of a fast-paced and accurate planning will put the entire workflow under stress, raising the risk of planning errors. Additionally, the experience and preference of brachytherapy planners, as well as clinical expertise and intuition would result in large inter- and intra-plan quality variations, further introducing more uncertainties to the BT treatment.

In the last decade, Knowledge-based planning (KBP), a new set of data-driven methodologies has been developed to improve the quality and efficiency of EBRT planning based on the previous high-quality clinical plans (6–9). Many researchers have demonstrated KBP's strength and validity in guiding planners to achieve optimal dose-volume histograms (DVHs) for OARs in treatment planning (10–12). KBP methods in EBRT are usually classified into two major categories: (a) case and atlas-based methods; and (b) statistical modeling and machine learning methods. In general, case and atlas-based methods utilize geometric features to find the best-matched prior cases from a database to improve the current planning. The statistical modeling and machine learning approaches try to build dose prediction models based on regression models such as stepwise regression (13, 14), multivariate linear regression (15–17), support vector regression (18, 19), and logistic regression (20, 21). After years of development, KBP methods have been widely investigated and even clinically implemented in commercial treatment planning systems. RapidPlan™, a commercial software module integrated into Eclipse released by Varian (Varian Medical Systems, Palo Alto, CA, USA), is one such system that predicts the achievable doses and specifies the optimization objectives needed to achieve them based on the KBP method. The success of RapidPlan™ has demonstrated the KBP's practicability and effectiveness in EBRT (22, 23).

All these achievements verified the success KBP achieved in model accuracy, stability, and feasibility in clinical application in the area of EBRT. As mentioned above, the unique clinical workflow in BT makes it vulnerable to suboptimal plans. Quality

control tools such as KBP are especially important, and promising in detecting suboptimal plans in BT. Moreover, current recommendations and guidelines have provided various dose constraints for BT planning. These population-based guides mainly concentrate on ensuring normal tissues do not exceed the dose limits, rather than guaranteeing individual patient receive the optimal dose distribution for their anatomy (24). Therefore, patient anatomy-based KBP, which could be used as patient-specific dose predictions to quantify currently unknown quality variations, would be particularly useful for brachytherapy.

However, despite great success achieved in EBRT, KBP is still far relatively unexplored in HDR brachytherapy due to some exceptional challenges in brachytherapy. The main reason is that the dose distribution in BT plans is highly constrained around the inserted applicator, resulting in insufficient freedom of dose modulation in planning. To the best of our knowledge, only two studies have investigated the possible application of KBP in brachytherapy, and both focused on tandem and ovoid applicator in an intracavitary setting. In Yusufaly et al.'s study (25), the authors used an established external-beam knowledge-based DVH estimation method to predict D_{2cc} in OARs. Zhou et al. (26) proposed a support vector machine model for dose prediction and effectively predicted D_{2cc} . Both studies demonstrated good results in predicting critical brachytherapy dose metrics using intracavitary brachytherapy applicators. Currently, only Tandem and Ovoid applicators were investigated in the knowledge-based planning strategies. However, various applicators with different geometry and source position would increase the dose distribution's versatility and complexity.

In this study, we proposed a KBP method to estimate the OAR dose distribution in HDR-BT treatment. To our knowledge, this work is the first application of knowledge-based dose estimation in HDR-BT with both intracavitary and interstitial cases included.

2 Method

2.1 Prediction pipeline

The study consists of two main tasks: (a) training dataset selection and (b) model training and validation. Figure 1 shows an overview of the entire workflow. Briefly, we first performed the principal component analysis (PCA) for all the training cases. Then, we retrieved the top k (bladder: k=30, rectum: k=20) plans from the training database based on PCA results using the k-nearest neighbor (kNN) algorithm. Cases in the new dataset have the most similar features to the cases in the validation dataset. These selected cases are then used to create a new training dataset to build training models. Subsequently, we used kernel density estimation (KDE) to develop a robust

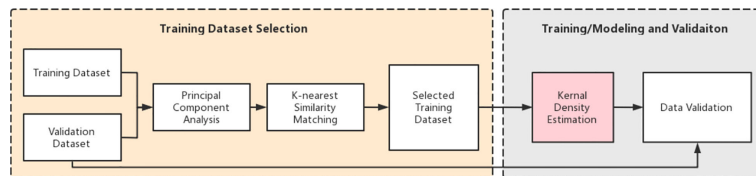


FIGURE 1 Pipeline of the DVH prediction.

prediction model based on the selected k plans. After training, a series of validation tests were conducted.

2.2 Patient database

79 HDR cervical patients previously treated in our center from 2019 to 2021 were included in this retrospective study. Each patient consists 2-4 fraction, with a total of 216 cases/fractions were involved. In this study, a ‘case’ indicates one single fraction containing a unique simulation CT and a treatment plan. All the cases were randomly divided into two datasets, with 170 cases in the training dataset and 46 cases in the validation dataset.

OARs and HRCTV were contoured based on CT images acquired at a GE 128 slice CT (Discovery, GE Healthcare, Inc.) using a tube setting of 120kV and 60mAs. All of the scans used identical parameters for image acquisition and reconstruction. The slice thickness and slice increment were 2.5*2.5mm. Each patient was treated with an applicator set among Tandem and Ovoid applicator (T+O), Ovoid applicator alone, Vaginal Multi-Channel applicator, 3D printed applicators, free needles, and a tandem applicator with up to 10 interstitial needles (T+N). The overview of the treatment characteristics is shown in Table 1.

According to the OARs dose constraint and the prescribed dose for the target (HR-CTV) recommended by the GEC-ESTRO, and the American Brachytherapy Society (ABS), all the treatment plans were created using the Oncentra treatment planning (Elekta, Stockholm, Sweden) system and based on TG 43 algorithm. The HRCTV was given an 80–90Gy EQD2 (biologically equivalent dose in 2-Gy fractions) prescription dose, assuming $\alpha/\beta = 10$. The bladder was given a maximum D_{2cc} of 90Gy EQD2, and the rectum was given a maximum D_{2cc} of 75 Gy EQD2, assuming $\alpha/\beta = 3$. The prescription dose in HDR brachytherapy ranges from 4-6 Gy/fraction (Table 1). In the treatment planning, we use a graphical optimization approach to repeatedly optimize the plan until the dose administered to 90% of the HRCTV reached the prescription dose. The dose distribution in each plan was scrutinized by physicians and physicists to ensure the dose distribution was clinically acceptable before treatment.

2.3 DVH prediction modeling

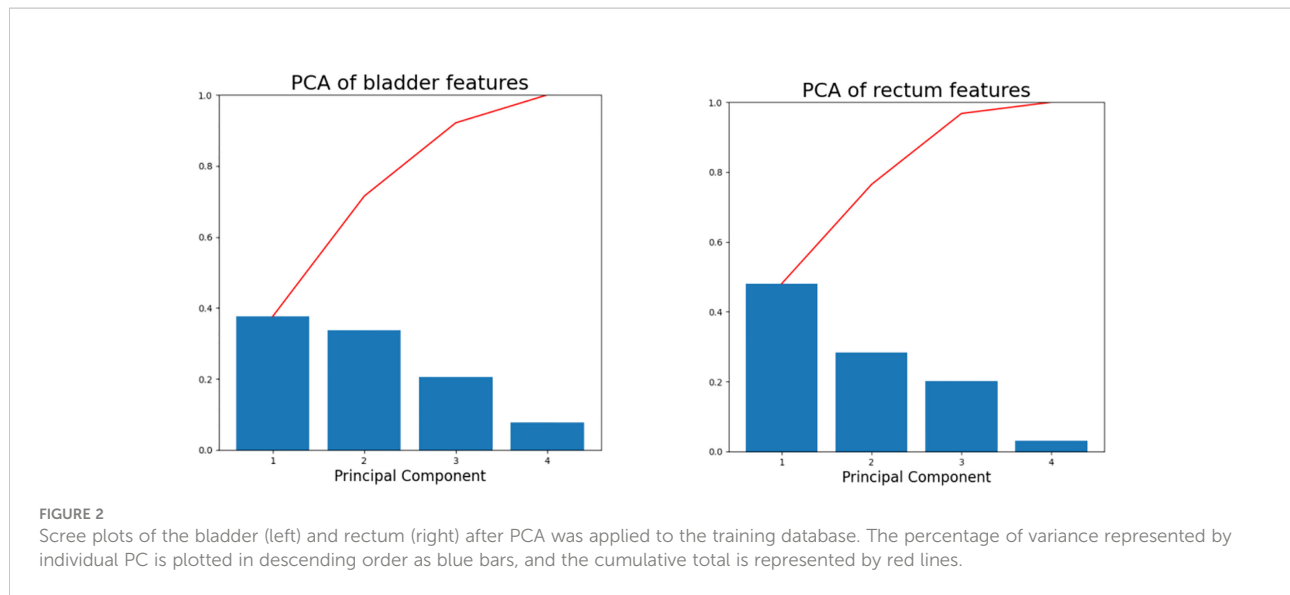
2.3.1 Feature reduction

Principal component analysis (PCA) has been applied to the brachytherapy plans to reduce the data dimensions and the model complexity (18, 27). PCA enables the transformation of a large set of variables into a smaller one containing most of the information. In this study, four features closely related to the dose distribution were processed using PCA to further reduce their dimensions: HRCTV volume, the distance between the centroids of the D_{2cc} and the HRCTV, prescription dose, and the average distance from D_{2cc} to the margin of the HRCTV. PCA simplifies the diversity of the dose distribution into a few principal component directions, and the individual variations can be represented by a small number of principal components. In our study, the first three principal components account for more than 90% of the variance (Figure 2) in both bladder and rectum, and thus were employed in the subsequent kNN similarity matching process for case selection.

TABLE 1 Treatment characteristics of included patients.

	Training Dataset	Test Dataset
Applicator Type		
Tandem + Needles	26	8
Vaginal	34	8
Tandem + Ovoid	45	15
Ovoid	28	5
Free Needles	26	6
3D_Printed	11	4
Total	170	46
Prescription Dose		
4	10	5
4.5	5	2
5	40	12
5.5	15	3
6	100	24
Total	170	46

Prescription dose and applicator types are grouped based on data from all included fractions(cases) per patient.



2.3.2 k-Nearest similarity matching

To take full advantage of the proposed feature, we incorporated k-Nearest Neighbors (kNN) similarity matching (28), a known robust non-parametric regression method to reduce model complexity. The purpose of kNN is to select a subset of training cases that matches closest to the validation case to create a new training dataset for subsequent model training. After recasting the training dataset along the principal component axes, the Euclidean distance between the single case in the validation set and each case in the training dataset was calculated separately in three principal component dimensions. kNN was thus used to retrieve k cases having the smallest Euclidean distance (bladder: k=30, rectum: k=20) in the training dataset.

2.3.3 Kernel density estimation

Considering that the determinant factor for dose levels near the HRCTV is the distance to the HRCTV, the histogram of distances from voxels in OARs to the HRCTV surface, which is the distance-to-target histogram (DTH), is a natural choice for predictive features. Thus, we implemented kernel density estimation (29, 30) for model training in this study. In KDE, for each voxel inside the OAR, we measure the closest distance from the voxel to the HRCTV boundary, and the distance was denoted by $x_i = x_1, x_2, \dots, x_i$ (Figure 3). The second step was to estimate the joint probability $P_{D, x_i}(d, x_i)$ of the dose d and the distance x , in which d was the corresponding dose in each voxel inside the OAR. Thirdly, we estimated the distance probability distribution $P_{x_i}(x_i)$, and calculated the conditional probability through: $P_{D|x_i}(d|x_i) = P_{D, x_i}(d, x_i)/P_{x_i}(x_i)$

In the prediction part, we estimated the new distance distribution probability $P_{x_i}^{\#}(x_i)$ for each case from the

validation set. Based on the calculated conditional probability in the training process, the prediction of the dose distribution for a new patient in the validation set can be calculated as: $P_D^{\#}(d) = \sum P_{D|x_i}(d|x_i) \cdot P_{x_i}^{\#}(x_i)$. In the final step, DVH was defined as a function of the dose d and the probability that a random variable D was larger than or equal to d : $DVH(d) = 1 - P(D \leq d) = 1 - \int_0^d P_D^{\#}(s) ds$.

2.4 Model validation

To quantitatively measure the prediction accuracy of the proposed model, we set the actual clinical plan DVHs as a baseline for comparison. Specific dose-volume indices including D_{2cc} , D_{1cc} , $D_{0.1cc}$, D_{max} , and D_{mean} , were extracted and analyzed, where D_{xcc} represented the minimum dose received by x cm³ of an OAR. Absolute residuals of predicted value and the actual value ($|\Delta D_{2cc}|$, $|\Delta D_{1cc}|$, $|\Delta D_{0.1cc}|$, $|\Delta D_{max}|$, and $|\Delta D_{mean}|$) were calculated to assess the level of agreement between the predicted DVHs and the actual plan DVHs. Standard deviation (σ) over the residuals was considered a measure of model error. The mean squared error (MSE) was calculated for all planned and predicted D_{2cc} values.

2.5 Statistical analysis

Significant differences were determined using a two-sided paired t-test. Correlations between predicted parameters and actual parameters were tested by performing the Pearson correlation test. Kruskal-Wallis ANOVA was

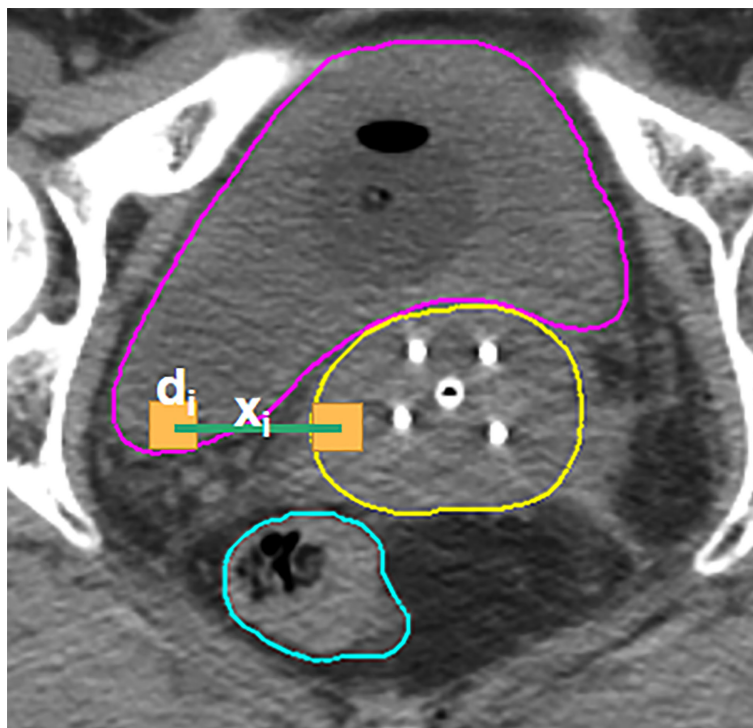


FIGURE 3

Illustration of the KDE method in the brachytherapy treatment planning. For each voxel in the OAR, x_i is the distance between this voxel and its closest voxel on the HRCTV surface and d_i is the dose received in each voxel inside the OAR.

employed to identify differences among various applicators. All statistical data analyses were performed using Python.

3 Results

For each test case, we built a model trained with the selected k cases (bladder: $k=30$, rectum: $k=20$) from a total of 170 training cases. The model accuracy was evaluated using a separate validation dataset consisting of 46 cases. The kernel density model based on 170 cases is shown in Figure 4. There was no statistically significant difference between the predicted and actual D_{2cc} values for the bladder ($p = 0.74$), rectum ($p = 0.57$) in the validation group. As shown in Table 2, the MSE value of D_{2cc} for the bladder was 0.23, and for rectum was 0.18. In bladder, the absolute difference between predicted plans and actual plans in D_{2cc} , D_{1cc} , $D_{0.1cc}$, D_{max} and D_{mean} , were 0.38 ± 0.29 , 0.4 ± 0.32 , 0.43 ± 0.36 , 0.97 ± 0.66 and 0.13 ± 0.09 , respectively. For the corresponding rectum, the absolute difference values were 0.34 ± 0.27 , 0.38 ± 0.33 , 0.63 ± 0.57 , 1.41 ± 0.99 and 0.23 ± 0.17 , respectively. The model prediction error in D_{2cc} to the bladder, and rectum was within 0.3 Gy, as quantified by the standard deviation (Figure 5).

The model-predicted DVHs were compared to their corresponding DVHs in the actual plans to assess the model's prediction accuracy. The DVHs of one example in the validation set were plotted and compared, as shown in Figure 6.

4 Discussion

Knowledge-based planning in EBRT has demonstrated its effectiveness across multiple disease sites. However, it is still less investigated in high-dose-rate brachytherapy. This study developed a KBP method for DVH prediction in HDR-BT. Theoretically, such a KBP model can be trained using gold-standard datasets and serve as quality assurance tools in the clinic to identify suboptimal plans in treatment planning prospectively. Therefore, this KBP method has a great potential to assess the treatment plan quality and offer guidance for following plan optimization.

Different from previous studies, in which only T+O applicator was investigated (25, 26, 31), we involved different applicator sets including T+O, Vaginal Multi-Channel applicator, Ovoid applicator, free needles, 3D printed applicators, and T+N applicators. Overall, our results show slightly better accuracy in D_{2cc} than Yusufaly et al. (25): bladder σ was 0.36 Gy vs. 0.46 Gy,

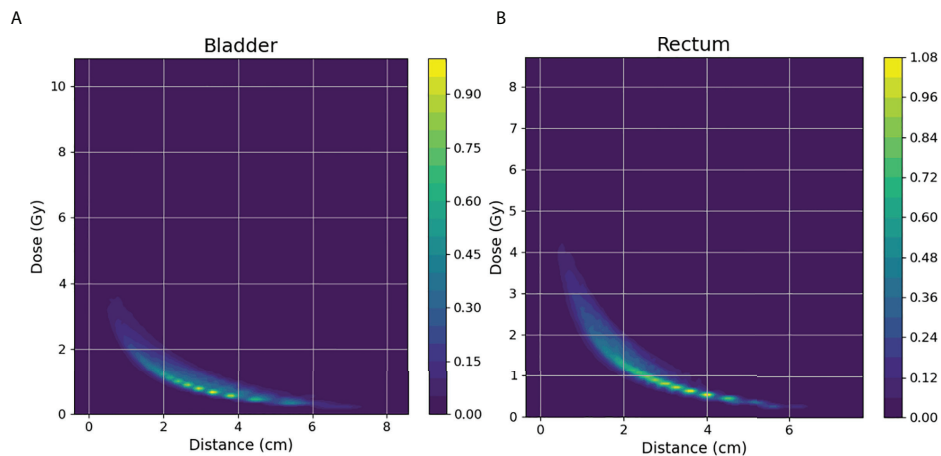


FIGURE 4 (A, B) depict the KDE model for bladder(left) and rectum(right) of 170 training cases.

rectum σ was 0.42 Gy vs. 0.47 Gy. (Since Yusufaly used relative error, here we also used the results of relative error for comparison.) To evaluate whether the applicator types would affect the predictive accuracy, we applied the Kruskal-Wallis analysis of variance (ANOVA) test in the validation dataset. The ANOVA revealed

that D_{2cc} has no significant statistical difference among different applicators ($p=0.109$). A possible explanation for this might be that the combination of kNN and PCA selected similar training cases for each validation case, reducing the variation in dose distribution caused by different applicators.

TABLE 2 Model performances for bladder and rectum.

	MSE (D_{2cc})	$ \Delta D_{2cc} $	$ \Delta D_{1cc} $	$ \Delta D_{0.1cc} $	$ \Delta D_{max} $	$ \Delta D_{mean} $
Bladder	0.23	0.38 ± 0.29	0.4 ± 0.32	0.43 ± 0.36	0.97 ± 0.66	0.13 ± 0.09
Rectum	0.18	0.34 ± 0.27	0.38 ± 0.33	0.63 ± 0.57	1.41 ± 0.99	0.23 ± 0.17

The mean squared error and the absolute difference between predicted values and actual values were calculated.

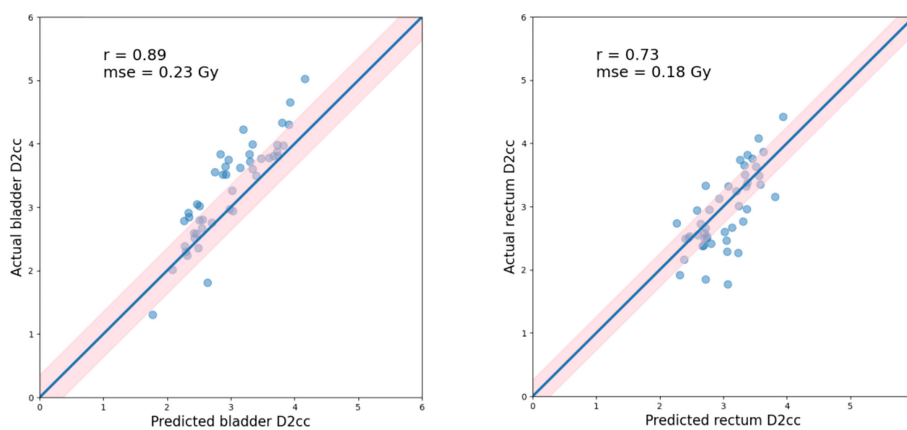


FIGURE 5 Actual vs. predicted D_{2cc} for the bladder(left) and rectum(right), as well as the Pearson correlation coefficients (r), standard deviation (indicated by σ as well as light pink shaded area). Blue lines indicate the theoretically ideal predictions.

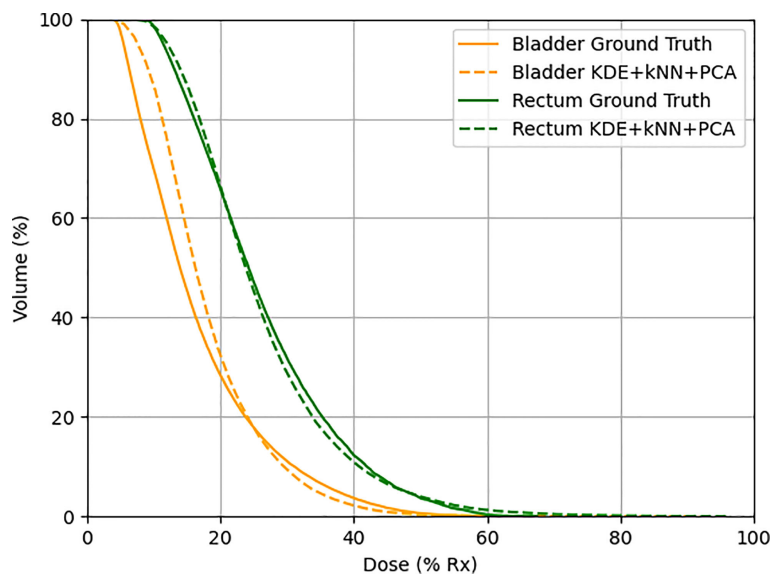


FIGURE 6
One example in the validation set demonstrating actual (solid line) and predicted (dotted line) DVHs. Orange lines represent the bladder, and green lines represent the rectum.

kNN was used to select *k* plans in the training dataset that mostly resemble the case in the validation dataset for model training. Once we determined the *k* value, we would neglect or hardly consider the information of other nearest neighbors. Thus, reducing the number of cases for training naturally comes at the expense of accuracy. The benefit of the case reduction is that smaller data set is easier to explore and analyze. It eliminates redundant and irrelevant variables and gives rise to an easier and faster training process in machine learning. A proper *k* value is crucial for model training and an inappropriate *k* value may yield unstable performance. To select the most appropriate *k* value, we run the kNN algorithm on 20 test cases with *k* values ranging from 10-40. The *k* value achieved the best model performance in D_{2cc} was used in model training (Table 3). Thus, we select *k*=30 for bladder and 20 for the rectum.

PCA was combined with kNN for better similarity matching in case selection. The main function of PCA was to reduce feature dimensionality in an interpretable way, while preserving

TABLE 3 Mean square error of predicted D_{2cc} using different *k* values.

k	Bladder	Rectum
10	0.283	0.094
20	0.215	0.091
30	0.210	0.106
40	0.215	0.110

as much information as possible at the same time. In our study, the first three principal components were used for kNN similarity matching. To verify the effectiveness of the PCA and kNN in model performance, we tested two different methods, KDE only, and KDE combined with kNN and PCA, separately. As shown in Table 4, results showed that the all-inclusive approach which combined KDE, kNN and PCA achieved the best D_{2cc} performance for both bladder and rectum. The all-inclusive method showed a significantly better prediction performance than KDE only, with an improvement of 30.3% for bladder and 33.3% for rectum. Figure 7 depicts one example predicted DVHs and actual DVHs (ground truth) for bladder and rectum. The DVHs predicted using all-inclusive method had the lowest difference from the actual DVHs.

DVH prediction modeling is often a complex, not a simple problem. There are still several sources of error limiting the model's ability to predict a satisfactory DVH. In this investigation, the main error came from the nonconformity of the HRCTV dose distribution, and 100% prescription dose line (contour of $V_{100\%}$). In KDE, since we do not have the contour of 100% prescription dose line for validation cases, we use HRCTV instead. We calculated the probability distribution based on the assumption that HRCTV achieved exactly 100% prescription dose. However, the dose distribution inside the HRCTV was determined by the source dwell position and dwell time, and the dose distribution in brachytherapy are highly constrained around applicators. The HRCTV and 100% prescription dose line can be slightly

TABLE 4 Comparison of model performances using different methods.

		MSE (D2cc)	\Delta D2cc	\Delta D1cc	\Delta D0.1cc	\Delta Dmax	\Delta Dmean
Bladder	KDE	0.27	0.44 ± 0.28	0.47 ± 0.29	0.5 ± 0.42	2.1 ± 1.9	0.32 ± 0.2
	KDE+kNN+PCA	0.23	0.38 ± 0.29	0.4 ± 0.32	0.43 ± 0.36	0.97 ± 0.66	0.13 ± 0.09
Rectum	KDE	0.27	0.43 ± 0.31	0.45 ± 0.34	0.54 ± 0.49	0.93 ± 0.72	0.31 ± 0.23
	KDE+kNN+PCA	0.18	0.34 ± 0.27	0.38 ± 0.33	0.63 ± 0.57	1.41 ± 0.99	0.23 ± 0.17

The mean squared error and the absolute residual between predicted values and ground truth were calculated.

inconsistent if the case has inadequate or inappropriate applicator positions in some slices, especially in the slices at the beginning or the end of HRCTV. That is to say, the bias between the HRCTV contour and the 100% prescription dose line would cause error in the subsequent dose probability estimation. Thus, we applied a contour correction method to adjust the HRCTV contour.

$$\text{Mean} = \frac{1}{n} \sum \frac{1}{r_{ij}^2}$$

where $i = 1, 2, 3 \dots n$, n is the number of points in the HRCTV boundary, $j = 1, 2, 3 \dots m$, m is the number of applicators, r_{ij} represents the distance between the point in the HRCTV boundary to the applicator.

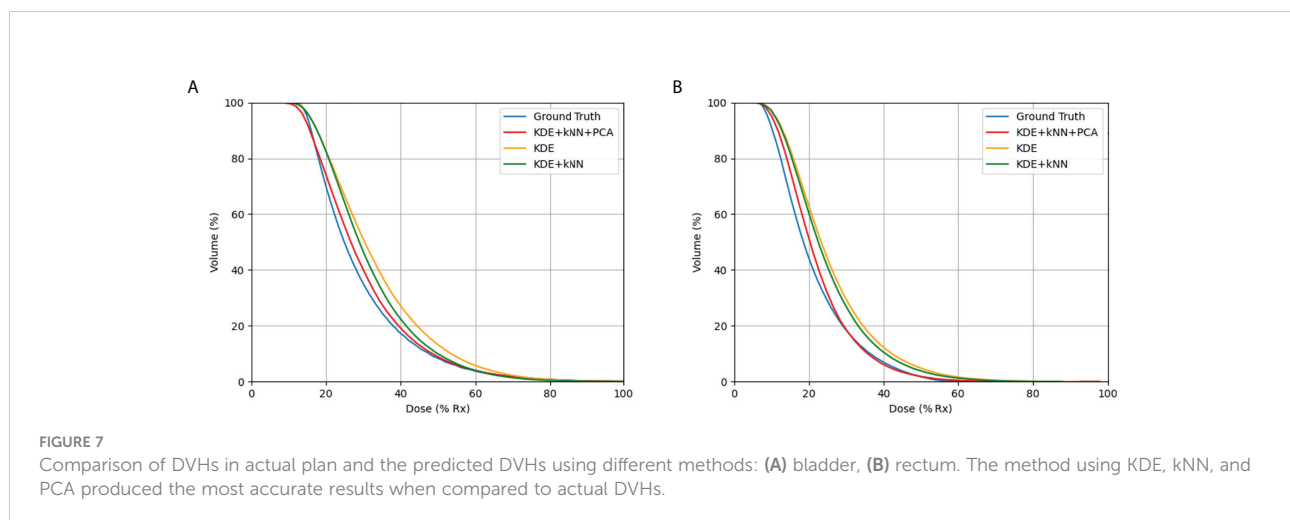
Figure 8 shows the optimized HRCTV contour based on the correction method. The Dice Similarity Coefficient (DSC) value was used to evaluate the matching degree after correction. The DSC values between the optimized contour and 100% prescription dose line indicate a slightly better match degree than the original HRCTV contour, but ultimately these corrections made unstable and modest improvements ranging from -0.023-0.057 in D_{2cc} prediction accuracy. Thus, we did not integrate this correction into our machine learning algorithm. This limitation will become the direction of our future research.

To sum up, this work presents a KBP method to predict DVH for OARs in brachytherapy treatment. Patient anatomical features in previously treated cases were learned to predict

DVHs for new patients. The predictions based on the individual patient geometry could motivate planners to go beyond the dosimetric constraints imposed by protocols to improve planning and provide better dose sparing for OARs. In our study, the entire process, including case selection, model construction, and DVH prediction, can be completed within one minute, which is acceptable in clinical application. Initial results have shown great potential in making this KBP model a quality control tool in treatment planning. Future studies will focus on the feasibility verification of using this model as a quality control tool in clinical practice. Another research direction is to re-plan those plans with D_{2cc} exceeding the prediction interval to improve plan quality and facilitate customized treatment planning for each patient.

5 Conclusion

In this paper, we proposed a machine learning method based on KDE, kNN, and PCA to predict the DVH in HDR-BT. The DVH for a new treatment plan was estimated using patient-specific anatomical information and an estimation model trained from prior plans. To our knowledge, this is the first KBP method that can predict the DVHs in patient who was treated with interstitial applicators, intracavitary applicators or both. The preliminary results have verified the



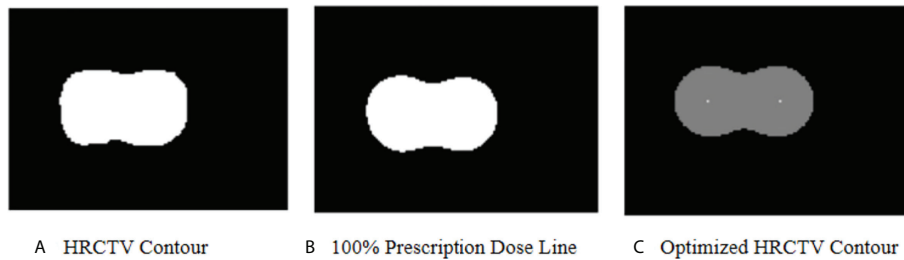


FIGURE 8

(A) shows the HRCTV contoured by a radiation oncologist in an actual plan. (B) shows the 100% prescription dose line in an actual plan. (C) shows the optimized HRCTV contour based on the correction method. In KDE method, we assume (A) matches with (B) perfectly. In this case, the optimized HRCTV has a better similarity with the 100% prescription dose line.

model's effectiveness in OAR dose estimation and its potential in providing guidance for brachytherapy planning in the future.

Data availability statement

The original contributions presented in the study are included in the article. Further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by Shanghai Sixth People's Hospital. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

ZL and KC : data acquisition, data processing, statistics, and writing; ZY, QZ, XY, ZL and JF: reviewing and editing. All authors contributed to the article and approved the submitted version.

References

1. Buskwofe A, David-West G, Clare CA. A review of cervical cancer: Incidence and disparities. *J Natl Med Assoc* (2020) 112(2):229–32. doi: 10.1016/j.jnma.2020.03.002
2. Chargari C, Deutsch E, Blanchard P, Gouy S, Martelli H, Guerin F, et al. Brachytherapy: An overview for clinicians. *CA Cancer J Clin* (2019) 69(5):386–401. doi: 10.3322/caac.21578
3. Holschneider CH, Petereit DG, Chu C, Hsu IC, Ioffe JY, Klopp AH, et al. Brachytherapy: A critical component of primary radiation therapy for cervical cancer: From the society of gynecologic oncology (SGO) and the American

Funding

This research was supported by the Science and Technology Project of Shanghai Municipal Science and Technology Commission (No.22Y31900500).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

brachytherapy society (ABS). *Brachytherapy* (2019) 18(2):123–32. doi: 10.1016/j.brachy.2018.11.009

4. Eifel PJ, Thoms WW Jr, Smith TL, Morris M, Oswald MJ. The relationship between brachytherapy dose and outcome in patients with bulky endocervical tumors treated with radiation alone. *Int J Radiat Oncol Biol Phys* (1994) 28(1):113–8. doi: 10.1016/0360-3016(94)90148-1

5. Tanderup K, Eifel PJ, Yashar CM, Pötter R, Grigsby PW. Curative radiation therapy for locally advanced cervical cancer: brachytherapy is NOT optional. *Int J Radiat Oncol Biol Phys* (2014) 88(3):537–9. doi: 10.1016/j.ijrobp.2013.11.011

6. Ge Y, Wu QJ. Knowledge-based planning for intensity-modulated radiation therapy: A review of data-driven approaches. *Med Phys* (2019) 46(6):2760–75. doi: 10.1002/mp.13526
7. Jiao SX, Wang ML, Chen LX, Liu XW. Evaluation of dose-volume histogram prediction for organ-at risk and planning target volume based on machine learning. *Sci Rep* (2021) 11(1):3117. doi: 10.1038/s41598-021-82749-5
8. Zhang J, Ge Y, Sheng Y, Yin FF, Wu QJ. Modeling of multiple planning target volumes for head and neck treatments in knowledge-based treatment planning. *Med Phys* (2019) 46(9):3812–22. doi: 10.1002/mp.13679
9. Nwankwo O, Mekdash H, Sihono DS, Wenz F, Glatting G. Knowledge-based radiation therapy (KBRT) treatment planning versus planning by experts: validation of a KBRT algorithm for prostate cancer treatment planning. *Radiat Oncol* (2015) 10:111. doi: 10.1186/s13014-015-0416-6
10. Zhang J, Wu QJ, Xie T, Sheng Y, Yin FF, Ge Y. An ensemble approach to knowledge-based intensity-modulated radiation therapy planning. *Front Oncol* (2018) 8:57. doi: 10.3389/fonc.2018.00057
11. Tol JP, Dahele M, Delaney AR, Slotman RJ, Verbakel WF. Can knowledge-based DVH predictions be used for automated, individualized quality assurance of radiotherapy treatment plans? *Radiat Oncol* (2015) 10:234. doi: 10.1186/s13014-015-0542-1
12. Janssen TM, Kusters M, Wang Y, Wortel G, Monshouwer R, Damen E, et al. Independent knowledge-based treatment planning QA to audit pinnacle autoplanning. *Radiother Oncol* (2019) 133:198–204. doi: 10.1016/j.radonc.2018.10.035
13. Wang J, Jin X, Zhao K, Peng J, Xie J, Chen J, et al. Patient feature based dosimetric pareto front prediction in esophageal cancer radiotherapy. *Med Phys* (2015) 42(2):1005–11. doi: 10.1118/1.4906252
14. Millunchick CH, Zhen H, Redler G, Liao Y, Turian JV. A model for predicting the dose to the parotid glands based on their relative overlapping with planning target volumes during helical radiotherapy. *J Appl Clin Med Phys* (2018) 19(2):48–53. doi: 10.1002/acm2.12203
15. Sheng Y, Ge Y, Yuan L, Li T, Yin FF, Wu QJ, et al. Outlier identification in radiation therapy knowledge-based planning: A study of pelvic cases. *Med Phys* (2017) 44(11):5617–26. doi: 10.1002/mp.12556
16. Faught AM, Olsen L, Schubert L, Rusthoven C, Castillo E, Castillo R, et al. Functional-guided radiotherapy using knowledge-based planning. *Radiother Oncol* (2018) 129(3):494–8. doi: 10.1016/j.radonc.2018.03.025
17. Younge KC, Marsh RB, Owen D, Geng H, Xiao Y, Spratt DE, et al. Improving quality and consistency in NRG oncology radiation therapy oncology group 0631 for spine radiosurgery via knowledge-based planning. *Int J Radiat Oncol Biol Phys* (2018) 100(4):1067–74. doi: 10.1016/j.ijrobp.2017.12.276
18. Zhu X, Ge Y, Li T, Thongphiew D, Yin FF, Wu QJ. A planning quality evaluation tool for prostate adaptive IMRT based on machine learning. *Med Phys* (2011) 38(2):719–26. doi: 10.1118/1.3539749
19. Zhang HH, Meyer RR, Shi L, D'Souza WD. The minimum knowledge base for predicting organ-at-risk dose-volume levels and plan-related complications in IMRT planning. *Phys Med Biol* (2010) 55(7):1935–47. doi: 10.1088/0031-9155/55/7/010
20. Boutilier JJ, Lee T, Craig T, Sharpe MB, Chan TC. Models for predicting objective function weights in prostate cancer IMRT. *Med Phys* (2015) 42(4):1586–95. doi: 10.1118/1.4914140
21. Lee T, Hammad M, Chan TC, Craig T, Sharpe MB. Predicting objective function weights from patient anatomy in prostate IMRT treatment planning. *Med Phys* (2013) 40(12):121706. doi: 10.1118/1.4828841
22. Yuan L, Ge Y, Lee WR, Yin FF, Kirkpatrick JP, Wu QJ. Quantitative analysis of the factors which affect the interpatient organ-at-risk dose sparing variation in IMRT plans. *Med Phys* (2012) 39(11):6868–78. doi: 10.1118/1.4757927
23. Li N, Carmona R, Sirak I, Kasaova L, Followill D, Michalski J, et al. Highly efficient training, refinement, and validation of a knowledge-based planning quality-control system for radiation therapy clinical trials. *Int J Radiat Oncol Biol Phys* (2017) 97(1):164–72. doi: 10.1016/j.ijrobp.2016.10.005
24. Yusufaly TI, Meyers SM, Mell LK, Moore KL. Knowledge-based planning for intact cervical cancer. *Semin Radiat Oncol* (2020) 30(4):328–39. doi: 10.1016/j.semradonc.2020.05.009
25. 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(5):624–34. doi: 10.1016/j.brachy.2020.04.008
26. Zhou P, Li X, Zhou H, Fu X, Liu B, Zhang Y, et al. Support vector machine model predicts dose for organs at risk in high-dose rate brachytherapy of cervical cancer. *Front Oncol* (2021) 11:619384. doi: 10.3389/fonc.2021.619384
27. Söhn M, Alber M, Yan D. Principal component analysis-based pattern analysis of dose-volume histograms and influence on rectal toxicity. *Int J Radiat Oncol Biol Phys* (2007) 69(1):230–9. doi: 10.1016/j.ijrobp.2007.04.066
28. Salvador-Meneses J, Ruiz-Chavez Z, Garcia-Rodriguez J. Compressed kNN: K-nearest neighbors with data compression. *Entropy (Basel)* (2019) 21(3):234. doi: 10.3390/e21030234
29. Skarpman Munter J, Sjölund J. Dose-volume histogram prediction using density estimation. *Phys Med Biol* (2015) 60(17):6923–36. doi: 10.1088/0031-9155/60/17/6923
30. Fan J, Wang J, Zhang Z, Hu W. Iterative dataset optimization in automated planning: Implementation for breast and rectal cancer radiotherapy. *Med Phys* (2017) 44(6):2515–31. doi: 10.1002/mp.12232
31. Cortes KG, Kallis K, Simon A, Mayadev J, Meyers S, Moore KL. Knowledge-based three-dimensional dose prediction for tandem-and-ovoid brachytherapy. *Brachytherapy* (2022) 21(4):532–42. doi: 10.1016/j.brachy.2022.03.002