RESEARCH

Approach and assessment of automated stereotactic radiotherapy planning for early stage non-small-cell lung cancer

Xue Bai, Guoping Shan, Ming Chen and Binbing Wang[*]

*Correspondence:
wangbb@zjcc.org.cn
Department of Radiation
Physics, Zhejiang Key
Laboratory of Radiation
Oncology, Zhejiang Cancer
Hospital, Hangzhou 310022,
Zhejiang, People's Republic
of China

Abstract

Background: Intensity-modulated radiotherapy (IMRT) and volumetric-modulated arc therapy (VMAT) are standard physical technologies of stereotactic body radiotherapy (SBRT) that are used for patients with non-small-cell lung cancer (NSCLC). The treatment plan quality depends on the experience of the planner and is limited by planning time. An automated planning process can save time and ensure a high-quality plan. This study aimed to introduce and demonstrate an automated planning procedure for SBRT for patients with NSCLC based on machine-learning algorithms. The automated planning was conducted in two steps: (1) determining patient-specific optimized beam orientations; (2) calculating the organs at risk (OAR) dose achievable for a given patient and setting these dosimetric parameters as optimization objectives. A model was developed using data of historical expertise plans based on support vector regression. The study cohort comprised patients with NSCLC who were treated using SBRT. A training cohort ($N = 125$) was used to calculate the beam orientations and dosimetric parameters for the lung as functions of the geometrical feature of each case. These plan–geometry relationships were used in a validation cohort ($N = 30$) to automatically establish the SBRT plan. The automatically generated plans were compared with clinical plans established by an experienced planner.

Results: All 30 automated plans (100%) fulfilled the dose criteria for OARs and planning target volume (PTV) coverage, and were deemed acceptable according to evaluation by experienced radiation oncologists. An automated plan increased the mean maximum dose for ribs (31.6 ± 19.9 Gy vs. 36.6 ± 18.1 Gy, $P < 0.05$). The minimum, maximum, and mean dose; homogeneity index; conformation index to PTV; doses to other organs; and the total monitor units showed no significant differences between manual plans established by experts and automated plans ($P > 0.05$). The hands-on planning time was reduced from 40–60 min to 10–15 min.

Conclusion: An automated planning method using machine learning was proposed for NSCLC SBRT. Validation results showed that the proposed method decreased planning time without compromising plan quality. Plans generated by this method were acceptable for clinical use.

Keywords: Machine learning, Non-small-cell lung cancer radiotherapy planning, Stereotactic body radiotherapy, Machine learning

Bai *et al. BioMed Eng OnLine*   (2019) 18:101

Page 2 of 15

Background

Stereotactic body radiotherapy (SBRT) is an attractive alternative to lobar resection in patients with early stage non-small-cell lung cancer (NSCLC) not eligible for lobectomy. Compared with surgery, SBRT is noninvasive; it does not involve postoperative complications and can realize good local control [1, 2]. Intensity-modulated radiotherapy (IMRT) and volumetric-modulated arc radiotherapy (VMAT) are typical technologies used in NSCLC SBRT [3]. IMRT enables good conformity to tumor volume and low doses for healthy tissues [4]. Through the development of dynamic delivery using multi-leaf collimators (MLC), VMAT can deliver dose efficiently to reduce uncertainties in the intra-fraction setup [5, 6]. In IMRT, dose is delivered through a serial of static segments which is made up of modulated MLCs' shapes. IMRT employs variable intensity across multiple radiation beams leading to the construction of highly conformal dose distributions. VMAT is an advanced IMRT technology [7]. The linac rotates around the patient and the MLCs continuously reshape and change the intensity of beams during dose delivery. Giving the radiotherapy in VMAT shortens the treatment time. VMAT is an alternative to fixed-gantry angle IMRT delivery. It allows the simultaneous variation of gantry rotation speed, treatment aperture shape via movement of MLC leaves, and dose rate during treatment delivery. One of the most important factors that affects the prognosis of IMRT and VMAT is the quality of the treatment plan, which depends on the experience and skill of the planner [8, 9]. Moreover, a plan is often developed through trial-and-error, necessitating a long planning time, which can be a limitation.

Several studies have focused on decreasing the interactions between the planner and the computer to reduce the planning time and improve the consistency of the plan quality. Methods such as template-based planning with the use of "scripting" tools [10–13], which imitate the trial-and-error process involved in manual planning, as well as multi-criteria optimization (MCO) [14–18] and knowledge-based treatment planning [19–26] have been investigated. Knowledge-based automated planning uses historical expertise plans as training data to learn the optimal strategies for new patients. This method has developed rapidly with the advancement of machine-learning algorithms. However, (1) only a few previous automated planning strategies have focused on SBRT, and (2) most existing studies have used the training data with an identical beam arrangement for different patients, such as seven or nine coplanar beams with equally spaced gantry angles in IMRT or a full-arc beam in VMAT [19–26]. These are listed in Table 1. Patient-specific beam arrangement has not been integrated in the knowledge-based automated planning studies.

Although identical beam arrangements are widely used in the traditional radiotherapy, customized partial arc radiation is generally better than an identical beam arrangement in an SBRT plan owing to variation in the tumor location and the low-dose radio sensitivity of healthy lung tissues. The shape, size, and position of tumors are important features that need to be considered for an optimized beam arrangement. Therefore, a strategy including either beam orientation optimization or available dose prediction is required to build automated SBRT planning tools.

Machine learning is a widely used method for analyzing big data from medical images and encapsulating it into task-specific information. Radiomics, medical image segmentation, and mass lesion classification based on machine-learning algorithms

Table 1 Studies on knowledge-based automated radiotherapy treatment plans

	Database	Site	Prediction method	Dosimetric parameter predictions	Beam angle predictions
Moore [19]	25 IMRT	Head-and-neck prostate	Analytical formulas	Mean doses of esophagus, larynx, parotid gland, bladder and rectum	No
Wu [20]	91 IMRT	Head-and-neck	Minimal dose approximation	Dose-volume objectives for head-and-neck organs	No
Zhu [21]	198 IMRT	Prostate	Machine learning	DVH of bladder and rectum	No
Petit [22]	25 IMRT	Pancreas	Minimal dose approximation	Dose-volume objectives for kidney and liver	No
Yang [24]	21 IMRT	Prostate	Linear regression	Rectal dose D_{15}, D_{20}, D_{25}, D_{35}, D_{50}, and bladder dose D_{15}	No
Nwankwo [25]	95 VMAT	Prostate	Machine learning	Uniformity index, D_{10}, D_{30}, D_{50}, D_{70} and D_{90} in bladder and rectum	No
Wang [26]	80 IMRT	Esophagus	Machine learning	Mean heart dose and mean lung dose	No

have brought about several breakthroughs in the diagnosis and treatment of cancer [27–31]. Machine learning-based automated planning systems have also been proposed by several studies [21, 25, 32, 33]. These methods have been proven efficient in head-and-neck cancer [34, 35] and prostate cancer [21, 32, 36]. In this study, the feasibility of a machine learning-based automated planning method was studied in NSCLC SBRT. We used a machine-learning method to determine the optimal strategy for an SBRT plan for lung cancer, including patient-specific objective function and beam arrangement. Anatomical features were extracted from individual patients, and the final beam arrangement and dosimetric results were predicted. Machine learning-based mathematical models were proposed and the feasibility of using this method in automated planning was validated. The objective of our study was to develop a fully automated NSCLC SBRT plan that aims to reduce planning time and meet the requirement for consistency in planning quality.

Results

Feature selection

The results of the feature selection are listed in Table 2. V_{Heart} and Y_{PH} were excluded for model building, because they had no significant correlation with beam and dosimetric features. D_{PL}, D_{PH}, X_{PL}, Y_{PL}, and X_{PH} were used in beam angle prediction, while V_{PTV}, V_{Lung}, D_{PL}, D_{PH}, OVZ_{PL}, and OVZ_{PH} were used in dosimetric feature prediction.

Validation of the prediction model

With regard to the CV of the LOO method, the root-mean-squared error (RMSE) of prediction for the start and stop gantry angles was 22.20° and 17.44°, respectively

Bai *et al. BioMed Eng OnLine*   (2019) 18:101

Page 4 of 15

Table 2 Spearman's rank correlation test

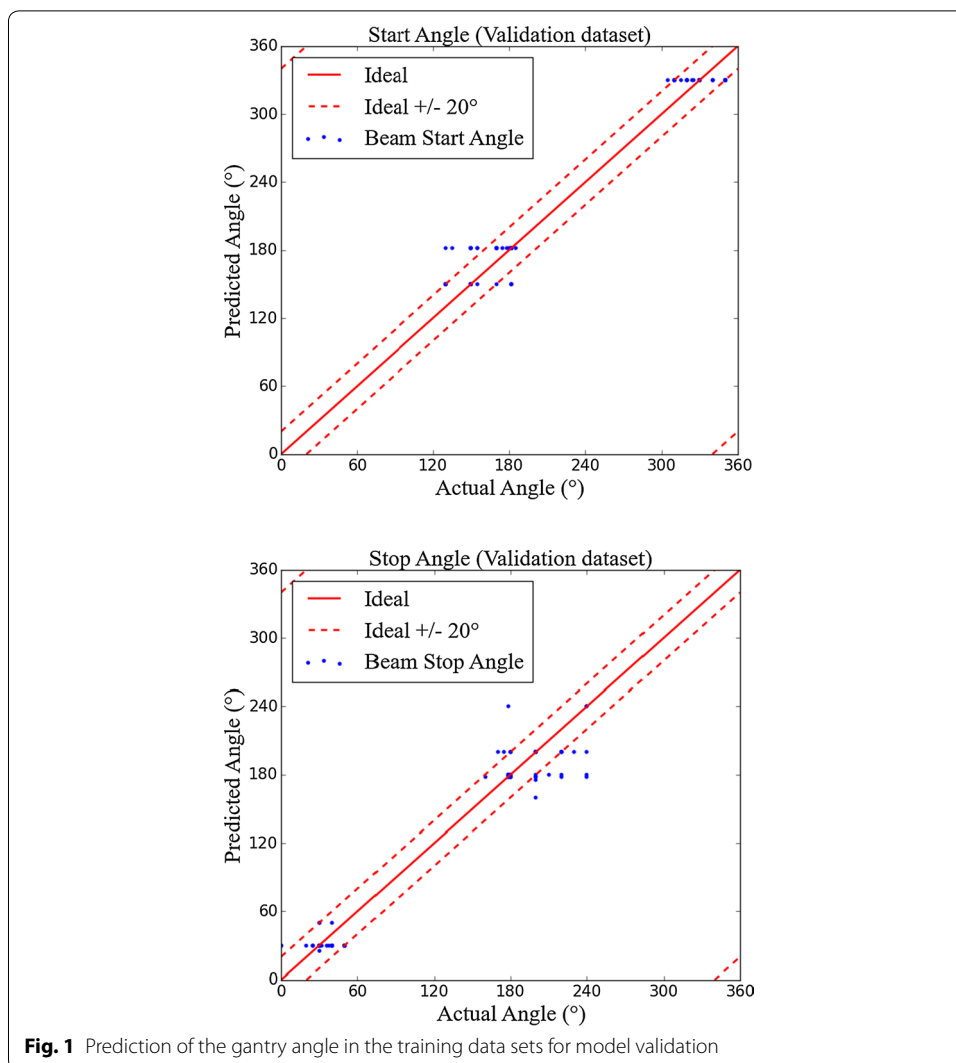|	Lung dose			Beam angle	
	Mean	V_{20}	V_{10}	Start	Stop
V_{PTV}					
ρ	0.747*	0.693*	0.731*	−0.033	−0.099
P	0.000	0.000	0.000	0.717	0.273
V_{Lung}					
ρ	−0.196*	−0.205*	−0.235*	0.073	0.021
P	0.029	0.022	0.008	0.416	0.813
V_{Heart}					
ρ	0.014	0.005	0.023	−0.077	-0.124
P	0.876	0.953	0.795	0.396	0.168
D_{PL}					
ρ	−0.508*	−0.519*	−0.467*	0.182*	0.234*
P	0.000	0.000	0.000	0.042	0.009
D_{PH}					
ρ	−0.241*	−0.233*	−0.203*	−0.432*	−0.412*
P	0.007	0.009	0.023	0.000	0.000
OVZ_{PL}					
ρ	0.654*	0.639*	0.583*	−0.33	−0.66
P	0.000	0.000	0.000	0.713	0.466
OVZ_{PH}					
ρ	0.385*	0.381*	0.254*	−0.075	−0.061
P	0.000	0.000	0.004	0.404	0.496
X_{PL}					
ρ	−0.098	−0.154	−0.075	0.762*	0.745*
P	0.277	0.087	0.405	0.000	0.000
Y_{PL}					
ρ	−0.100	−0.155	−0.074	0.758*	0.740*
P	0.265	0.085	0.415	0.000	0.000
X_{PH}					
ρ	−0.113	−0.169	−0.106	0.756*	0.729*
P	0.211	0.059	0.240	0.000	0.000
Y_{PH}					
ρ	−0.068	−0.114	−0.109	−0.169	−0.078
P	0.449	0.207	0.225	0.059	0.385

*Significant correlation

(Fig. 1). The RMSE prediction for the start and stop gantry angles for external validation data set was 18.5° and 9.6°, respectively.

The RMSE of the predicted mean lung dose (MLD), V_{10}, and V_{20} of the lung was 0.83 Gy, 3.10%, and 1.71%, respectively, for CV (Fig. 2). For external validation, the RMSE of the predicted MLD, V_{10}, and V_{20} of the lung was 1.00 Gy, 3.64%, and 2.40%, respectively.

Comparison of the quality of plans

Table 3 shows a comparison of all 60 plans for 30 patients in the testing set. According to the radiation oncologists, 30/30 automated plans fulfilled the dose criteria for OARs and PTV coverage for Zhejiang Cancer Hospital and were acceptable for clinical treatment.

Bai *et al. BioMed Eng OnLine* (2019) 18:101

Page 5 of 15

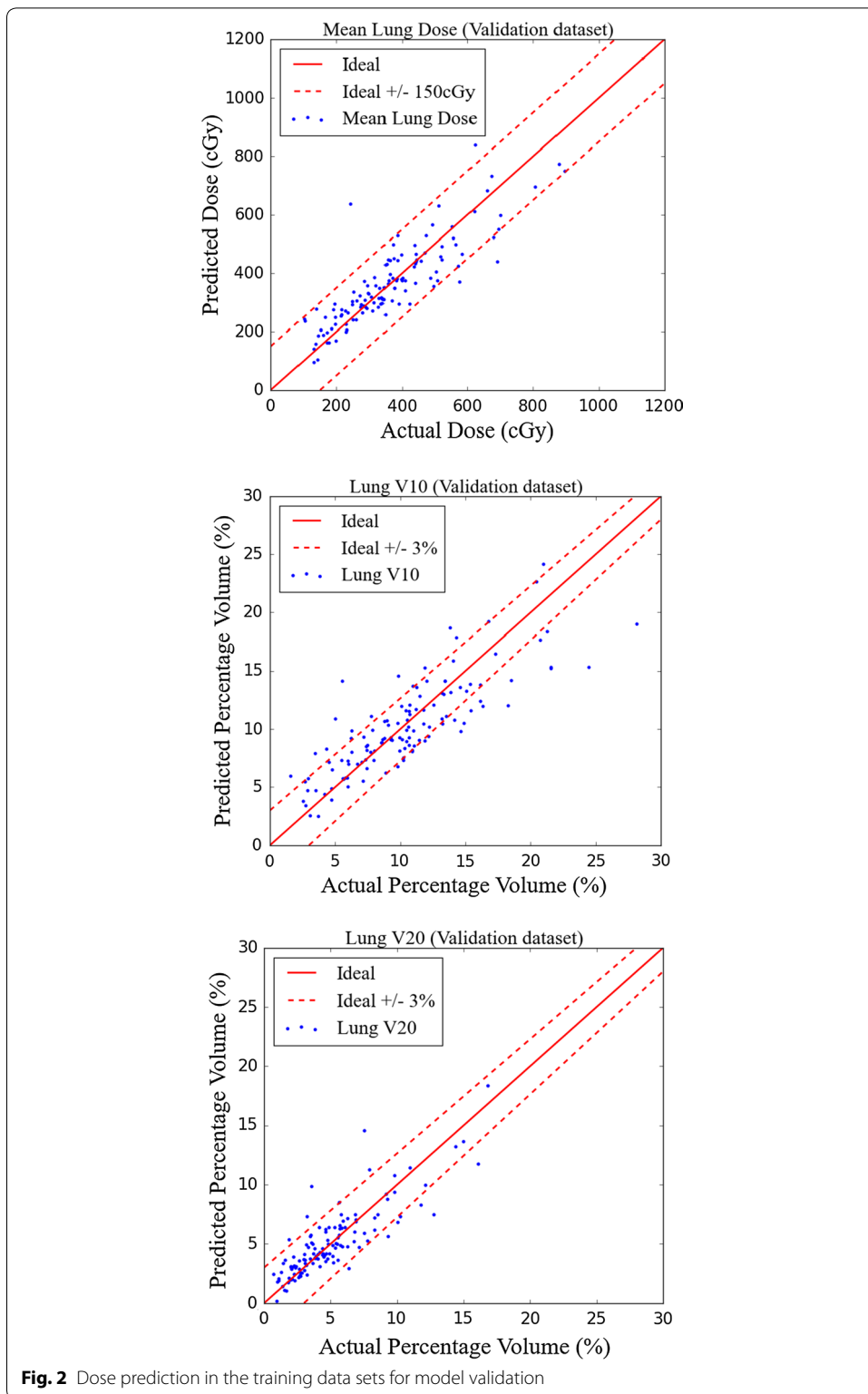**Fig. 1** Prediction of the gantry angle in the training data sets for model validation

A significant difference was observed only in the mean D_{max} of the rib: 31.6 ± 19.9 Gy in the manual plan and 36.6 ± 18.1 Gy in the automated plan ($P < 0.05$). Doses to other organs and PTV, as well as the total monitor units, showed no significant differences between the two planning strategies ($P > 0.05$). Figure 3 shows a boxplot of dose differences between the manual plan and the automated plan for each ROI.

The hands-on planning time was estimated and could be reduced from approximately 40–60 min to 10–15 min using the automated planning procedure.
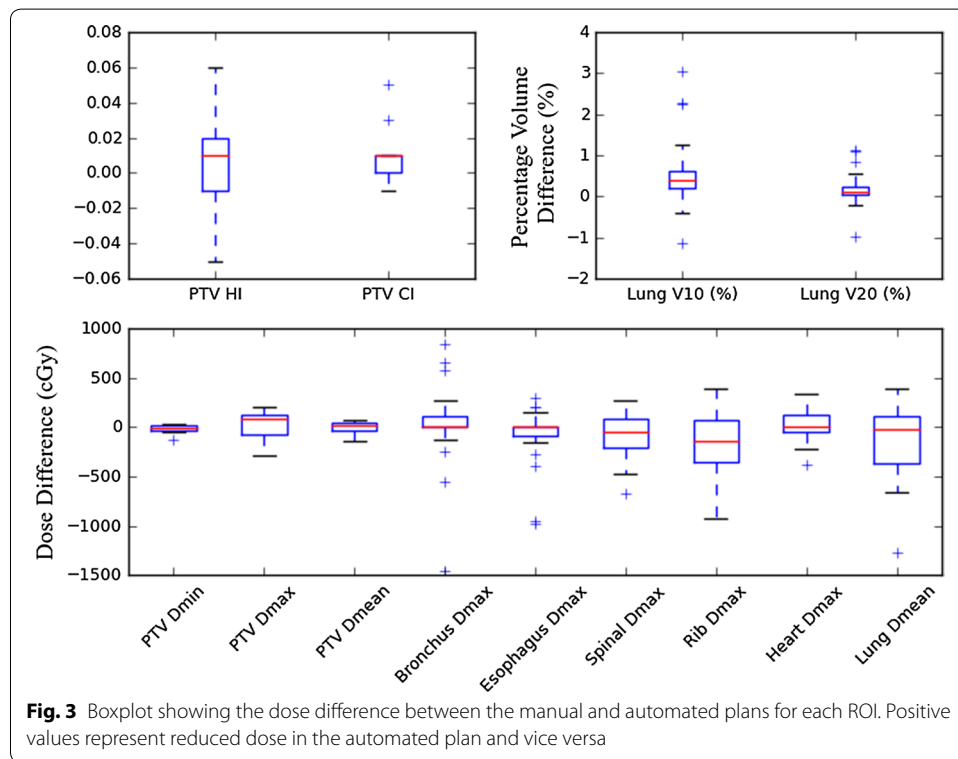
Discussion

This study demonstrates that machine-learning method can be used to generate clinically acceptable treatment plans automatically for NSCLC SBRT patients. All the plans were evaluated and recognized by two radiation oncologists. The anatomical geometry features of patients were extracted from 125 CT scans. The SVR was used to find the relationship between these features and the optimized plan parameters for training plans. Thereafter, the regressive models were used to generate radiotherapy plans

Fig. 2 Dose prediction in the training data sets for model validation

automatically. As shown in Table 3, for all the 30 validation patients, the automated plans generated were comparable with the plans created by expert planners, and were acceptable for clinical treatment.

Table 3 Clinical constraints and dose results for various indices

Index	Constraints	Manual plan		Automated plan		P value
		Mean	SD	Mean	SD	
PTV D_{min} (Gy)	>45.5	48.4	0.5	48.4	0.3	0.837
PTV D_{max} (Gy)	<70.0	66.7	2.5	69.8	2.1	0.594
PTV D_{mean} (Gy)	–	58.1	1.3	58.5	1.3	0.515
PTV HI	–	0.31	0.04	0.32	0.03	0.478
PTV CI	>0.8	0.87	0.04	0.88	0.03	0.157
Bronchus D_{max} (Gy)	<30.0	9.3	12.5	9.8	10.7	0.841
Esophagus D_{max} (Gy)	<32.5	5.9	6.0	6.5	5.8	0.472
Spinal D_{max} (Gy)	<30.0	9.7	5.1	8.7	4.2	0.082
Rib D_{max} (Gy)	<54.0	31.6	19.9	36.6	18.1	0.005
Heart D_{max} (Gy)	<30.0	13.1	12.2	13.0	11.6	0.658
Lung D_{mean} (Gy)	<5.0	3.4	1.4	3.3	1.4	0.082
Lung V_{10} (%)	<15.0	9.7	4.3	9.2	4.1	0.050
Lung V_{20} (%)	<10.0	4.6	2.6	4.4	2.5	0.658
Total MU	–	1078	217	1054	198	0.594

Fig. 3 Boxplot showing the dose difference between the manual and automated plans for each ROI. Positive values represent reduced dose in the automated plan and vice versa

In this study, the SVR was used to fit geometry features and plan parameters instead of the classical regression method, because the former has the advantage of preventing over-fit via stopping penalty when the error between predicted and actual values is less than a given value. Zhu [21] implemented SVR to establish the correlation between the features of the DVH and the anatomical information in a database consisting of 198

Bai *et al. BioMed Eng OnLine* (2019) 18:101

Page 8 of 15

high-quality prostate plans. SVR can manage over-fit when the data size increases. Furthermore, the best plan parameters were not identical owing to the subjective variations introduced by the planners. An expert plan could be achieved by setting the parameters in different ways, while its robust quality could be attained through minor changes in parameters. The error tolerance of SVR was appropriate in the prediction of plan parameters.

Although beam orientation optimization is essential and widely discussed [37–39], previous studies regarding automated plans often used full-arc beam arrangements rather than an optimized one. One of the reasons for this is that PTV with an irregular shape is more difficult for beam orientation prediction [26]. However, because of the variations in tumor position and low-dose sensitivity of healthy lung tissues, optimized beam arrangement may potentially reduce the complications in normal tissue. In that case, the beam arrangement strategy must be different from previous auto plan studies which focus on either head-and-neck [20, 23] or pelvic [24, 25, 40] regions. Data set involved in this study includes convex PTVs in early stages of NSCLCs, making the prediction of beam orientation feasible. The auto-beam arrangement based on anatomy was proposed and validated in this study. *iCycle* used a priori multi-criteria approach to combine beam angle choice into automated plan by sequentially adding beams in an iterative procedure [16]. Our SVR method could attain the optimized beam arrangement in a single step.

Feature selection in this study is different from that reported in previously published studies that employed overlap-volume histogram (OVH) to infer the dose-volume levels [20, 23, 24, 41]. The OVH represents percentage volumes of an OAR that is within a specified distance of PTV boundary. The OVH fails to describe the influence of space coordinates in the voxel dose. Moreover, a greater number of features increase the accuracy of the model but also increase the complexity of modeling and the risk of over-fitting, requiring a large number of training examples for model construction. Besides the features commonly used in dose prediction, such as ROI volumes and distances between PTV and OARs, three-dimensional coordinate information for each OAR was used to describe their relative position in the present study. Using this method of choosing anatomy factors, the calculated beam angle and predicted OAR dose can be obtained in real time.

The purpose of this study was developing an automated planning procedure for VMAT plans. IMRT plans were also enrolled to enlarge the size of training data set. Because for most tumor site, VMAT and fixed IMRT could produce largely equivalent target volume coverage, dose conformity, and homogeneity [42, 43], using a hybrid training data set is reasonable. Although there were 59.2% IMRT plans (74 in 125) in training data set, while the test data sets were all VMAT plans, the prediction results were acceptable for the test data set. The result proves that static IMRT data could be used for training VMAT models.

In this study, predictions of only three parameters, V_{10}, V_{20}, and D_{mean} of the lung, were made to ensure that the procedure for the automated treatment plan remained uncomplicated. Since the expected benefit from further reducing the D_{max} of serial organs compared with the constraint value is limited and not aimed for clinical practice, the optimization objectives for the heart, bronchus, spine, esophagus, and rib were

programmed as a template and kept constant for all patients in the study. The key dosimetric difference between the automated treatment plan and manual plan in this study was the maximum dose for the ribs. This was because a number of patients for whom the distance between the PTV mass center and the ribs mass center was short were enrolled. The aim of future studies would be to predict the doses for serial organs and combine the results into an objective function.

Our study is novel in that it calculated the best start and stop angles for VMAT using a machine-learning method, which can be a valuable supplement for the automated VMAT plan. The geometrical features used to build the prediction model could be readily extracted from CT scans using the scripts tool in RayStation, enabling the automated tool to be applicable to regular practice. In theory, the study proved the relationship between anatomy geometry and the treatment plan parameters for NSCLC SBRT. Based on this relationship, an automated planning tool was developed to reduce planning time and meet the requirements for plan quality.

There are several limitations to this study. First, only coplanar beam plans with 0° collimator were considered for the training data set. Studies on SBRT considering advanced models with non-coplanar fields and patient-specific collimator angle will be conducted in the future. Second, although comparison of the automated and manual plans revealed that the former could achieve acceptable sparing of critical structures for the 30 patients in the test data set, the results could not prove that those plans are the most optimal ones in the present study. Other optimization parameters, such as objective weight, gantry spacing, and template-based objective functions, must also be studied to determine optimal values. This work is a foundation for further study, and better training data sets should be used to produce better quality plans in the future.

Conclusions

Using features of the anatomy of patients to generate a predicted arc angle and an objective function for an automated treatment plan is feasible. The procedure for an automated treatment plan was developed in this study involving set start and stop angles for the beam arc as well as objective functions that would operate without manual intervention. The dosimetric impact of the automated plan on the PTV, bronchus, esophagus, spinal cord, heart, and lung was insignificant ($P < 0.05$). In the automated plan, the mean maximum dose for the ribs was increased (31.6 ± 19.9 Gy vs. 36.6 ± 18.1 Gy, $P < 0.05$). All the 30 validation cases produced results acceptable for applying this method to clinical treatments based on oncologists' evaluation. This procedure could reduce the planning time by nearly three-quarters of the time required to formulate manual plans while generating plans that are as good as ones designed by an experienced planner. Therefore, the plans generated by this method are acceptable for clinical use.
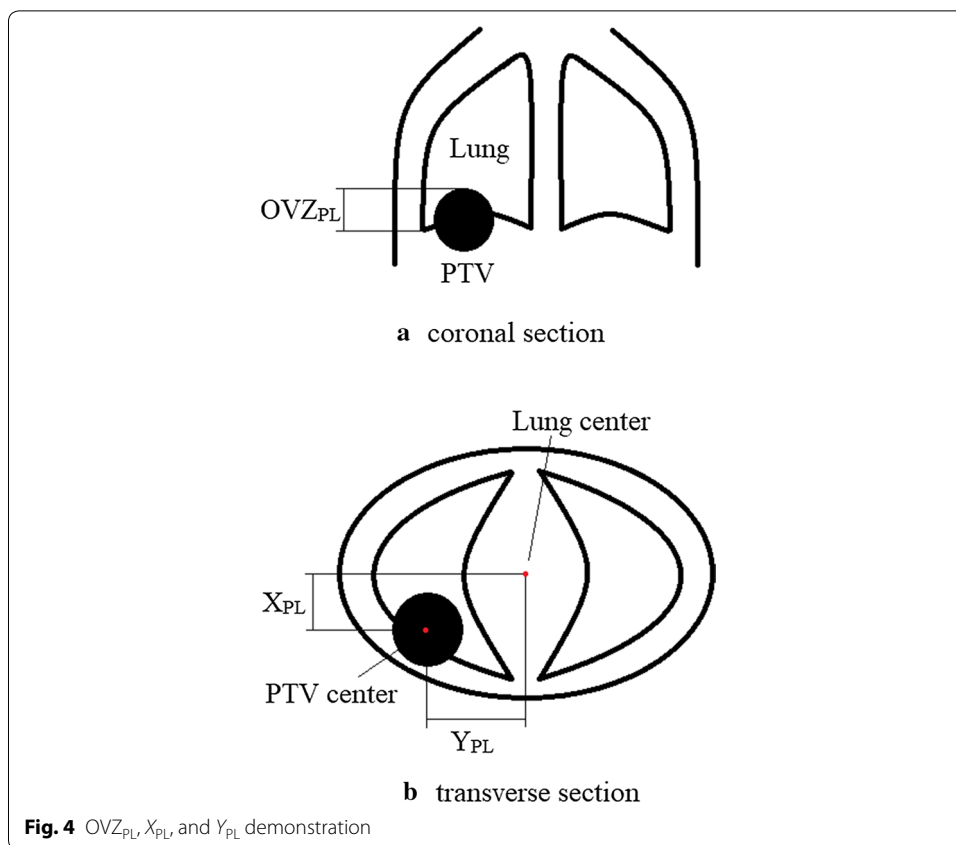
Methods

Patient data

A total of 125 patients with stage-I NSCLC underwent SBRT in 2015–2017 according to the protocols of the Zhejiang Cancer Hospital, Zhejiang, China. Clinical treatment plans for all patients were generated using the RayStation TPS (RaySearch Laboratories, Stockholm, Sweden).

The clinical target volume, planning target volume (PTV), and organs at risk (OARs) were delineated by experienced radiation oncologists and reviewed by senior physicians. The data comprised 74 IMRT plans and 51 VMAT plans. The IMRT plans were delivered using 11–13 step-and-shoot coplanar beams with a gantry spacing of 20° between the beams and arranged in a fan shape; the plans had dosimetric features similar to those of the VMAT plans. The latter were delivered using two arcs with a gantry spacing of 4° between the control points, with the distance between the start and stop angles varying from 220° to 260°. The start and stop angles of the arcs were decided by expert planners based on the anatomy of individual patients. The PTV was 3.19–357.20 cm^3 (mean, 36.92 cm^3). Patients were treated using five fractions and prescribed 50 Gy to the PTV. The prescription dose covered at 95% of the PTV, and the maximum dose did not exceed 150% of the prescription dose. The dosimetric constraints of the OAR partly consulted Radiation Therapy Oncology Group (RTOG) protocols 0813 and 0915, and are listed in Table 3. To conform the ALARA principle, all plans were optimized further using a trial-and-error process to achieve optimal sparing of OARs and were considered expert plans. These plans were used for clinical treatments and for the present study.

Characteristics of plans: geometry features, beam angles, and achievable dose for organs at risk (OARs)

In radiotherapy, the parameters of treatment plans are determined by the planners according to anatomical data based on computed tomography (CT) images. Intuitively, the beam orientation and constraints of OAR dose tend to correlate with the anatomic features of images from patients.

In the present study, 11 anatomical features were extracted from digital imaging and communications in medicine documents: (1) PTV volume (V_{PTV}); (2) lung volume (V_{Lung}); (3) heart volume (V_{Heart}); (4) distance between the PTV mass center and the lung-mass center (D_{PL}); (5) distance between the PTV mass center and the heart-mass center (D_{PH}); (6) overlap length of the PTV and the lung in the *z*-axis (OVZ_{PL}, introduced by Wang et al. to predict the Pareto front in esophageal cancer [26]); (7) overlap length of the PTV and the heart in the *z*-axis (OVZ_{PH}); (8) distance between the PTV mass center and the lung-mass center in the *x*-axis (X_{PL}); (9) distance between the PTV mass center and the lung-mass center in the *y*-axis (Y_{PL}); (10) distance between the PTV mass center and the heart-mass center in the *x*-axis (X_{PH}); and (11) distance between the PTV mass center and the heart-mass center in the *y*-axis (Y_{PH}). PTV is major concerned in treatment planning, in that 9 PTV related features were extracted. Meanwhile, delivered tumor cause inevitable dose to lung and heart which may cause radiation toxicity in normal tissue. Five lung-related features and five heart-related features were also extracted to evaluate delivery dose in this study. The OVZ_{PL}, X_{PL}, and Y_{PL} are shown in Fig. 4. These data could describe the volume, relative position, and shape of the regions of interest (ROIs), because the tissue anatomy of each patient was similar. The start and end angles of the arc or IMRT fan were recorded as features of the beam angle. The couch and collimator angles were 0° for all cases. The V_{10} (percentage lung volume of 10 Gy) and V_{20} (percentage lung volume of 20 Gy) of the lung and mean lung dose were recorded to represent the dose features. These dosimetric parameters were exported from the TPS using Python scripts.
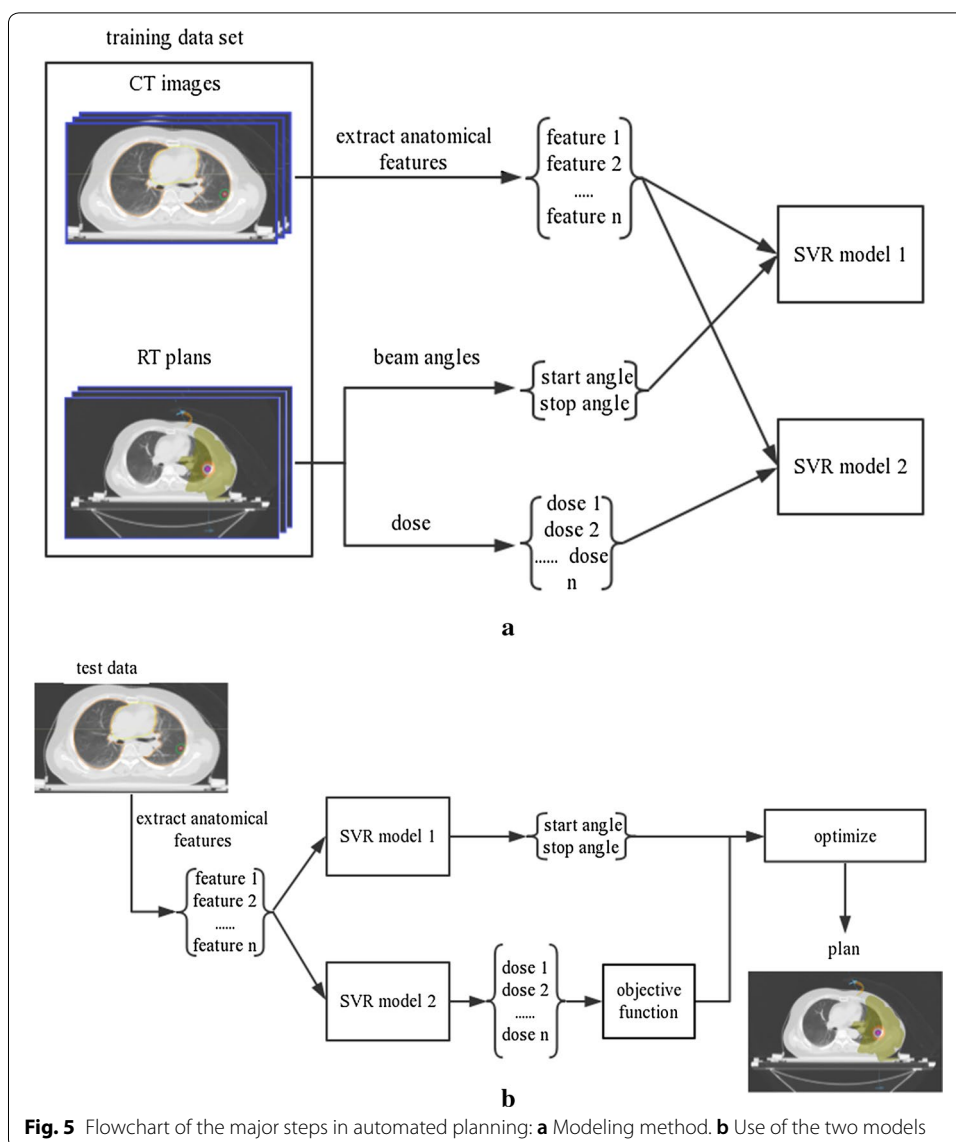
Fig. 4 OVZ_{PL}, X_{PL}, and Y_{PL} demonstration

Feature selection

Spearman's rank correlation test was used to evaluate the correlation between the anatomical features and the beam angle and dosimetric features. Spearman's rank correlation coefficient is a non-parametric rank statistic proposed as a measure of the strength of the association between two variables. It can be used in feature selection without making any assumptions about the frequency distribution of the variables [26]. If the *P* value was > 0.05, no significant correlation was found between the two variables. Irrelevant anatomical features were excluded from prediction modeling. The reserved features were used to predict beam angle and lung dose before determining the beam angle and objective function parameters for an automated plan using a machine-learning model.

Prediction and validation

Figure 5 is a flowchart of the major steps in the automated planning. The goal of training is to establish two mathematic correlations. One maps the anatomic information extracted from patient images and the selection of the beam angle. The other maps the anatomic information and V_{10}, V_{20}, and the mean dose of the lung. For convenience in the modeling, all plans were normalized at 95% of PTV in 50 Gy. All training data were standardized by removing the mean values and scaling to unit variance as a common requirement for machine-learning estimators.

Support vector regression (SVR) was implemented as the modeling method. SVR is a supervised learning method used for data regression. For complicated problems that

Fig. 5 Flowchart of the major steps in automated planning: **a** Modeling method. **b** Use of the two models

are not regressed to a linear function, SVR introduces a kernel function that projects the data into a higher dimensional space where it can be regressed to a linear function. By introducing the kernel, SVR gains flexibility in the choice of the form of regression function, which needs not be linear and even needs not have the same functional form for all data, since its function is non-parametric and operates locally. As a consequence, they can work with geometry features which show a non- linearly relation to the beam angles and OAR doses. SVR also introduces a tube of width ε; and finding a function that is at most ε deviations from the targets actually obtained for all the training data becomes problematic [44, 45]. By choosing an appropriate ε, SVR can be robust even when the training data have some bias; for example, the results of training plans are slightly variation according to the discretion of planers. In machine learning, a hyperparameter is a parameter whose value is set before the learning process begins. Hyperparameters include ε and the kernel function, which are not learned directly within the estimators

Bai *et al. BioMed Eng OnLine* (2019) 18:101

Page 13 of 15

in SVR. An exhaustive grid-search method was introduced to find the appropriate values for the two hyperparameters, by searching the parameter space for the best cross-validation (CV) score. The CV method is used to determine how the results of statistical analysis generalize to an independent data set. The leave-one-out (LOO) method was used in each model for CV. In LOO, an entire data set with n patients was separated into a training data set with $n-1$ patients and a validation data set with 1 patient [46]. The SVR model was developed using the training data set and applied to the validation data set. The SVR and Gridsearch algorithms were implemented using Scikit-learn [47].

A total of 125 cases were used in the training data set. With an LOO method, in each iteration, one case was randomly chosen as validation set and other 124 cases were training sets for cross validation. After models training, 30 cases outside the training pool were used as a test set for external validation. The actual values of the gantry start angle and stop angles, V_{10}, V_{20}, and mean dose of the lung were collected from the treatment plans generated by expert planners. The corresponding predicted value was calculated using a prediction model and the standard deviation of the resulting error was calculated.

Automated planning approach and assessment

Two factors of planning were determined automatically: gantry angles and objective functions. The start and stop gantry angles were predicted and used as an arc parameter for VMAT. Objective function parameters were calculated from the machine-learning model and individualized. For each patient, an automated plan was generated based on the predicted arc start and stop angles and the optimization objectives.

As a test for the automated planning procedure, two strategies were used to develop SBRT plans for the 30 cases in the testing set: (1) manual plan (designed by an experienced planner through trial-and-error) and (2) automated plan (designed by the automated planning procedure). All 60 plans were normalized at a 50-Gy dose (for 5 fractions) covering 95% of the PTV.

Two experienced radiation oncologists at Zhejiang Cancer Hospital reviewed the dose–volume histograms (DVHs) and dose distributions of the 30 automated plans and judged the acceptability of the plans for clinical treatments. The radiation oncologists was asked to decide each plan was clinical acceptable or not. For the PTV, the mean dose (D_{mean}), maximum dose (D_{max}), minimum dose (D_{min}), and homogeneity index (HI) as defined by the International Commission on Radiation Units and Measurements 83 [48], and the conformity index (CI) as defined by Paddick et al. [49] were evaluated. For OARs, D_{max} for the bronchus, esophagus, spine, ribs, and heart, as well as D_{mean}, V_{10}, and V_{20} for the lung, were evaluated for comparison.

Statistical analyses

Statistical analyses of dosimetric differences were performed using the Wilcoxon rank test based on the correlation between the manual plan and the automated plan using SPSS v21 (IBM, NY, USA). A P value < 0.05 was considered significant.

Abbreviations
IMRT: intensity-modulated radiotherapy; VMAT: volumetric-modulated arc therapy; SBRT: stereotactic body radiotherapy; NSCLC: non-small-cell lung cancer; OAR: organs at risk; PTV: planning target volume; MCO: multi-criteria optimization; CT:

computed tomography; RTOG: radiation therapy oncology group; CI: conformity index; SVR: support vector regression; CV: cross validation; LOO: leave one out; OVH: overlap-volume histogram; RMSE: root-mean-squared error; MLD: mean lung dose.

Availability of data and materials
The data sets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate
This study was approved by the Ethics Committee of Zhejiang Cancer Hospital with number 2014-03-12.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

References
1. Baardwijk AV, et al. Is high-dose stereotactic body radiotherapy (SBRT) for stage I non-small cell lung cancer (NSCLC) overkill? A systematic review. Radiother Oncol. 2012;105(2):145–9.
2. Chang JY, et al. Stereotactic ablative radiotherapy versus lobectomy for operable stage I non-small-cell lung cancer: a pooled analysis of two randomised trials. Lancet Oncology. 2015;16(6):630.
3. Peter M, Crispen C. A dosimetric retrospective planning study comparing volumetric arc therapy (VMAT) and stereotactic body radiotherapy (SBRT) treatment plans for non-small cell lung cancer (NSCLC). J Radiother Pract. 2015;14:260–7.
4. Bortfeld T. IMRT: a review and preview. Phys Med Biol. 2006;51(13):R363–79.
5. Bedford JL. Treatment planning for volumetric modulated arc therapy. Med Phys. 2009;36(11):5128–38.
6. Mutanga TF, et al. Day-to-day reproducibility of prostate intrafraction motion assessed by multiple kV and MV imaging of implanted markers during treatment. Int J Radiat Oncol Biol Phys. 2012;83(1):400–7.
7. Karl OJMP. Volumetric modulated arc therapy: IMRT in a single gantry arc. Med Phys. 2008;35(1):310–7.
8. Das IJ, et al. Intensity-modulated radiation therapy dose prescription, recording, and delivery: patterns of variability among institutions and treatment planning systems. J Natl Cancer Inst. 2008;100(5):300–7.
9. Nelms BE, et al. Variation in external beam treatment plan quality: an inter-institutional study of planners and planning systems. Pract Radiat Oncol. 2012;2(4):296–305.
10. Zhang X, et al. A sensitivity-guided algorithm for automated determination of IMRT objective function parameters. Med Phys. 2006;33(8):2935–44.
11. Tol JP, et al. Automatic interactive optimization for volumetric modulated arc therapy planning. Radiat Oncol. 2015;10:75.
12. Song Y, et al. Fully automatic volumetric modulated arc therapy plan generation for rectal cancer. Radiother Oncol. 2016;119(3):531–6.
13. Mitchell RA, et al. Improving the efficiency of breast radiotherapy treatment planning using a semi-automated approach. J Appl Clin Med Phys. 2017;18(1):18–24.
14. Craft D, et al. An approach for practical multiobjective IMRT treatment planning. Int J Radiat Oncol Biol Phys. 2007;69(5):1600–7.
15. Breedveld S, et al. A novel approach to multi-criteria inverse planning for IMRT. Phys Med Biol. 2007;52(20):6339–53.
16. Breedveld S, et al. iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans. Med Phys. 2012;39(2):951–63.
17. Pardo-Montero J, Fenwick JD. An approach to multiobjective optimization of rotational therapy. Med Phys. 2009;36(7):3292–303.
18. Pardo-Montero J, Fenwick JD. An approach to multiobjective optimization of rotational therapy. II. Pareto optimal surfaces and linear combinations of modulated blocked arcs for a prostate geometry. Med Phys. 2010;37(6):2606–16.
19. Moore KL, et al. Experience-based quality control of clinical intensity-modulated radiotherapy planning. Int J Radiat Oncol Biol Phys. 2011;81(2):545–51.

Bai *et al. BioMed Eng OnLine*   (2019) 18:101

Page 15 of 15

20. Wu B, et al. Data-driven approach to generating achievable dose-volume histogram objectives in intensity-modulated radiotherapy planning. Int J Radiat Oncol Biol Phys. 2011;79(4):1241–7.
21. Zhu X, et al. A planning quality evaluation tool for prostate adaptive IMRT based on machine learning. Med Phys. 2011;38(2):719–26.
22. Petit SF, et al. Increased organ sparing using shape-based treatment plan optimization for intensity modulated radiation therapy of pancreatic adenocarcinoma. Radiother Oncol. 2012;102(1):38–44.
23. Wu B, et al. Using overlap volume histogram and IMRT plan data to guide and automate VMAT planning: a head-and-neck case study. Med Phys. 2013;40(2):021714.
24. Yang Y, et al. An overlap-volume-histogram based method for rectal dose prediction and automated treatment planning in the external beam prostate radiotherapy following hydrogel injection. Med Phys. 2013;40(1):011709.
25. Nwankwo O, et al. A global quality assurance system for personalized radiation therapy treatment planning for the prostate (or other sites). Phys Med Biol. 2014;59(18):5575–91.
26. Wang JZ, et al. Patient feature based dosimetric Pareto front prediction in esophageal cancer radiotherapy. Med Phys. 2015;42(2):1005–11.
27. Aerts HJ, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun. 2014;5:4006.
28. Men K, Dai J, Li Y. Automatic segmentation of the clinical target volume and organs at risk in the planning CT for rectal cancer using deep dilated convolutional neural networks. Med Phys. 2017;44(12):6377–89.
29. Wang Q, et al. Radiomics nomogram building from multiparametric MRI to predict grade in patients with glioma: a cohort study. J Magn Reson Imaging. 2019;49(3):825–33.
30. Ma Z, et al. Nasopharyngeal carcinoma segmentation based on enhanced convolutional neural networks using multi-modal metric learning. Phys Med Biol. 2019;64(2):025005.
31. Jiang F et al. Breast mass lesion classification in mammograms by transfer learning, in Proceedings of the 5th international conference on bioinformatics and computational biology—ICBCB '17. 2017. p. 59–62.
32. Nwankwo O, et al. 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.
33. Masi K, et al. Knowledge-based treatment planning and its potential role in the transition between treatment planning systems. Med Phys. 2015;42(6Part20):3457.
34. Krayenbuehl J, et al. Evaluation of an automated knowledge based treatment planning system for head and neck. Radiat Oncol. 2015;10:226.
35. Tol JP, et al. Evaluation of a knowledge-based planning solution for head and neck cancer. Int J Radiat Oncol Biol Phys. 2015;91(3):612–20.
36. Cagni E, et al. Knowledge-based treatment planning: an inter-technique and inter-system feasibility study for prostate cancer. Phys Med. 2017;36:38–45.
37. Pugachev A, et al. Role of beam orientation optimization in intensity-modulated radiation therapy. Int J Radiat Oncol Biol Phys. 2001;50(2):551–60.
38. Hou Q, et al. Beam orientation optimization for IMRT by a hybrid method of the genetic algorithm and the simulated dynamics. Med Phys. 2003;30(9):2360–7.
39. Bangert M, Oelfke U. Spherical cluster analysis for beam angle optimization in intensity-modulated radiation therapy treatment planning. Phys Med Biol. 2010;55(19):6023–37.
40. Good D, et al. A knowledge-based approach to improving and homogenizing intensity modulated radiation therapy planning quality among treatment centers: an example application to prostate cancer planning. Int J Radiat Oncol Biol Phys. 2013;87(1):176–81.
41. Wu BB, et al. Patient geometry-driven information retrieval for IMRT treatment plan quality control. Med Phys. 2009;36(12):5497–505.
42. Holt A, et al. Volumetric-modulated arc therapy for stereotactic body radiotherapy of lung tumors: a comparison with intensity-modulated radiotherapy techniques. Int J Radiat Oncol Biol Phys. 2011;81(5):1560–7.
43. Brock J, et al. Optimising stereotactic body radiotherapy for non-small cell lung cancer with volumetric intensity-modulated arc therapy—a planning study. Clin Oncol. 2012;24(1):68–75.
44. Smola AJ, Schölkopf BJA. On a Kernel-based method for pattern recognition, regression, approximation, and operator inversion. Algorithmica. 1998;22(1–2):211–31.
45. Awad M, Khanna RJNIPL, and Reviews, Support Vector Regression. 2007;11(10):203–224.
46. Fushiki T. Estimation of prediction error by using K-fold cross-validation. Dordrecht: Kluwer Academic Publishers; 2011. p. 137–46.
47. Pedregosa F, et al. Scikit-learn: machine Learning in Python. J Mach Learn Res. 2012;12(10):2825–30.
48. Hodapp N. The ICRU Report No. 83: Prescribing, recording and reporting photon-beam intensity-modulated radiation therapy (IMRT). Strahlentherapie Und Onkologie. 2012;188(1):97–9.
49. Paddick I. A simple scoring ratio to index the conformity of radiosurgical treatment plans Technical note. J Neurosurg. 2000;93(Suppl 3):219.

Publisher's Note

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