


Incorporating Case-Based Reasoning for Radiation Therapy Knowledge Modeling: A Pelvic Case Study

Technology in Cancer Research & Treatment
Volume 18: 1-13
© The Author(s) 2019
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1533033819874788
journals.sagepub.com/home/tct


Yang Sheng, PhD¹ , Jiahan Zhang, PhD¹ , Chunhao Wang, PhD¹ , Fang-Fang Yin, PhD¹, Q. Jackie Wu, PhD¹ , and Yaorong Ge, PhD²

Abstract

Knowledge models in radiotherapy capture the relation between patient anatomy and dosimetry to provide treatment planning guidance. When treatment schemes evolve, existing models struggle to predict accurately. We propose a case-based reasoning framework designed to handle novel anatomies that are of same type but vary beyond original training samples. A total of 105 pelvic intensity-modulated radiotherapy cases were analyzed. Eighty cases were prostate cases while the other 25 were prostate-plus-lymph-node cases. We simulated 4 scenarios: Scarce scenario, Semiscarce scenario, Semiample scenario, and Ample scenario. For the Scarce scenario, a multiple stepwise regression model was trained using 85 cases (80 prostate, 5 prostate-plus-lymph-node). The proposed workflow started with evaluating the feature novelty of new cases against 5 training prostate-plus-lymph-node cases using leverage statistic. The case database was composed of a 5-case dose atlas. Case-based dose prediction was compared against the regression model prediction using sum of squared residual. Mean sum of squared residual of case-based and regression predictions for the bladder of 13 identified outliers were 0.174 ± 0.166 and 0.459 ± 0.508 , respectively ($P = .0326$). For the rectum, the respective mean sum of squared residuals were 0.103 ± 0.120 and 0.150 ± 0.171 for case-based and regression prediction ($P = .1972$). By retaining novel cases, under the Ample scenario, significant statistical improvement was observed over the Scarce scenario ($P = .0398$) for the bladder model. We expect that the incorporation of case-based reasoning that judiciously applies appropriate predictive models could improve overall prediction accuracy and robustness in clinical practice.

Keywords

radiation therapy, knowledge modeling, case-based reasoning, prostate cancer

Abbreviations

AI, artificial intelligence; DTH, distant-to-target histogram; DVH, dose-volume histogram; IMRT, intensity-modulated radiation therapy; LN, lymph node; NN, nearest neighbor; OAR, organ-at-risk; OVH, overlap volume histogram; PC, principal component; PTV, planning target volume; SSR, sum of squared residual; 2D, 2-dimensional; 3D, 3-dimensional

Received: February 28, 2019; Revised: June 25, 2019; Accepted: August 15, 2019.

Introduction

Knowledge modeling in radiation therapy treatment planning has been heavily researched¹⁻¹⁹ and clinically implemented in recent years via commercial products such as RapidPlan, a clinical treatment planning guidance solution provided by Varian Medical Systems (Palo Alto, California). Knowledge models predict dosimetric parameters, such as 1-dimensional dose-volume points or 3-dimensional (3D) dose distribution, based upon patient anatomical features. Depending on how dose end points are generated, there are 2 major categories of

¹ Department of Radiation Oncology, Duke University Medical Center, Durham, NC, USA

² Department of Software and Information Systems, University of North Carolina at Charlotte, Charlotte, NC, USA

Corresponding Author:

Yang Sheng, Department of Radiation Oncology, Duke University Medical Center, 201 Trent Dr, Durham, NC 27710, USA.

Email: yang.sheng@duke.edu



knowledge modeling. The first category is anatomy-based method. It utilizes nearest neighbors (NN) to find the most similar case for the query case based on some similarity metrics. Atlas-based treatment planning modeling^{20,21} falls into this category. Plan-related parameters such as the fluence map used by Good et al²⁰ or 3-D dose distribution used by Sheng et al²¹ are transferred to the query case's anatomy with appropriate fine-tuning. High similarity between the atlas case and the query case could guarantee decent dose prediction accuracy. Another category of knowledge modeling is statistics-based methods. It utilizes statistical regression and other machine learning models to learn the closed-form solution to the anatomy-dosimetry relation. Yuan et al¹ used distant-to-target histogram (DTH) to describe the geometric features between the organ-at-risk (OAR) and the planning target volume (PTV). The DTH-based piece-wise regression model is the fundamental solution employed by RapidPlan.²² Wu et al³ developed overlap volume histogram (OVH) to describe the geometric relation between OAR and PTV. Appenzoller et al² investigated multiple normal regression model to predict the dose level at certain distance. All aforementioned methods rely on the substantial amount of training cases to fully capture the relation between anatomy and dosimetry features.

Translating knowledge modeling into clinical practice requires substantial effort to guarantee the overall robustness of the model. Thorough analysis and validation of the knowledge model require efforts in both treatment planning and statistical modeling. Knowledge models are typically approximations of highly nonlinear maps. They require a sufficient number of cases for training and usually perform accurately when new cases are inliers in the feature space. Since there is intrinsic plan quality variation even within a single institution, sifting out high-quality plans to be included in model training is required. Delaney et al,²³ Tol et al,²⁴ and Sheng et al²⁵ analyzed the effect of outliers existing in the knowledge models. These studies found that the existence of outliers could adversely impact the model quality. Delaney et al and Tol et al focused on dosimetric outliers, while Sheng et al focused on the anatomical outliers. Planning target volume target delineation changes, such as moving from treating prostate only to treating pelvic lymph node (LN) as well, could introduce anatomical outliers to the knowledge model and subsequently deteriorate the model prediction accuracy as demonstrated by Sheng et al. However, no comprehensive strategy has been provided yet to handle these anatomical outliers in practice. On the other hand, improving the robustness of the statistical model is also needed to account for the variation in the training data set. One solution is to identify outliers and exclude them from modeling. One limitation of this approach is that it does not increase the model's capability to handle such outlier cases. If such similar outlier case does occur again, it may be excluded from modeling or prediction.

In order to improve the model's overall robustness, especially when dealing with novel anatomy, we propose a case-based reasoning framework for knowledge modeling. In this article, cases with novel anatomy, for example, the prostate-plus-LN cases in the present study, are also referred as outliers, or geometric outliers as referred by Sheng et al.²⁵ Geometric outlier is often referred to as anatomical outlier. It contrasts with the dosimetric outliers which are similar in anatomy but vary in plan quality. The novel anatomy, or anatomy outlier, contrasts with the inlier prostate cases primarily in the PTV definition. A prostate intensity-modulated radiation therapy (IMRT) plan generally treats the PTV which includes the prostate and seminal vesicles.²⁶ A prostate-plus-LN plan, however, requires treatment of the pelvic LN in addition to the prostate and seminal vesicles.²⁵ An example of a major difference in PTV delineation is depicted in Figure 1. Therefore, while they are both considered prostate cases, prostate and prostate-plus-LN cases are different in the sense of target delineation, which could affect the dose distribution as well as the OAR dose.

Case-based reasoning originates in artificial intelligence (AI) research as an effective framework to provide a solution to novel tasks. It consists of a closed-loop 4-R steps, namely "Retrieve", "Reuse", "Revise", and "Retain".²⁷ Retrieve aims to recall the most relevant experience to solving the current task by identifying a matching case that is in some sense the most similar to the current query case. Reuse refers to employing the solution from previous experience recorded in the matching case to the current task. Revise means adopting certain modifications to the previous solution to better solve the current task. Retain refers to storing the current experience and possibly updating the available knowledge for future practice. This closed-loop solution implemented in knowledge modeling could accumulate valuable knowledge over time to improve the capability of the knowledge model to handle more anatomy variation. The proposed case-based reasoning framework could provide a better understanding and utilization of machine learning and AI models in radiation therapy. As the general AI application evolves from shallow learning to deep learning, these AI tools turn into black box from the user's perspective. It is now of supreme importance that the user understands the model better in terms of when and how to use the tool for each individual scenario. This step needs to be well studied and understood before the machine learning, and AI tools can be safely deployed in the clinical application. This study is the first attempt to introduce case-based reasoning framework in radiation therapy knowledge modeling.

In this study, we implemented the case-based reasoning framework to radiation therapy knowledge modeling. This case-based reasoning framework is capable of handling different scenarios, evolving from when the novel anatomy available is scarce to when the novel anatomy has accumulated ample amount of cases. We used pelvic cases in this study to demonstrate the concept.

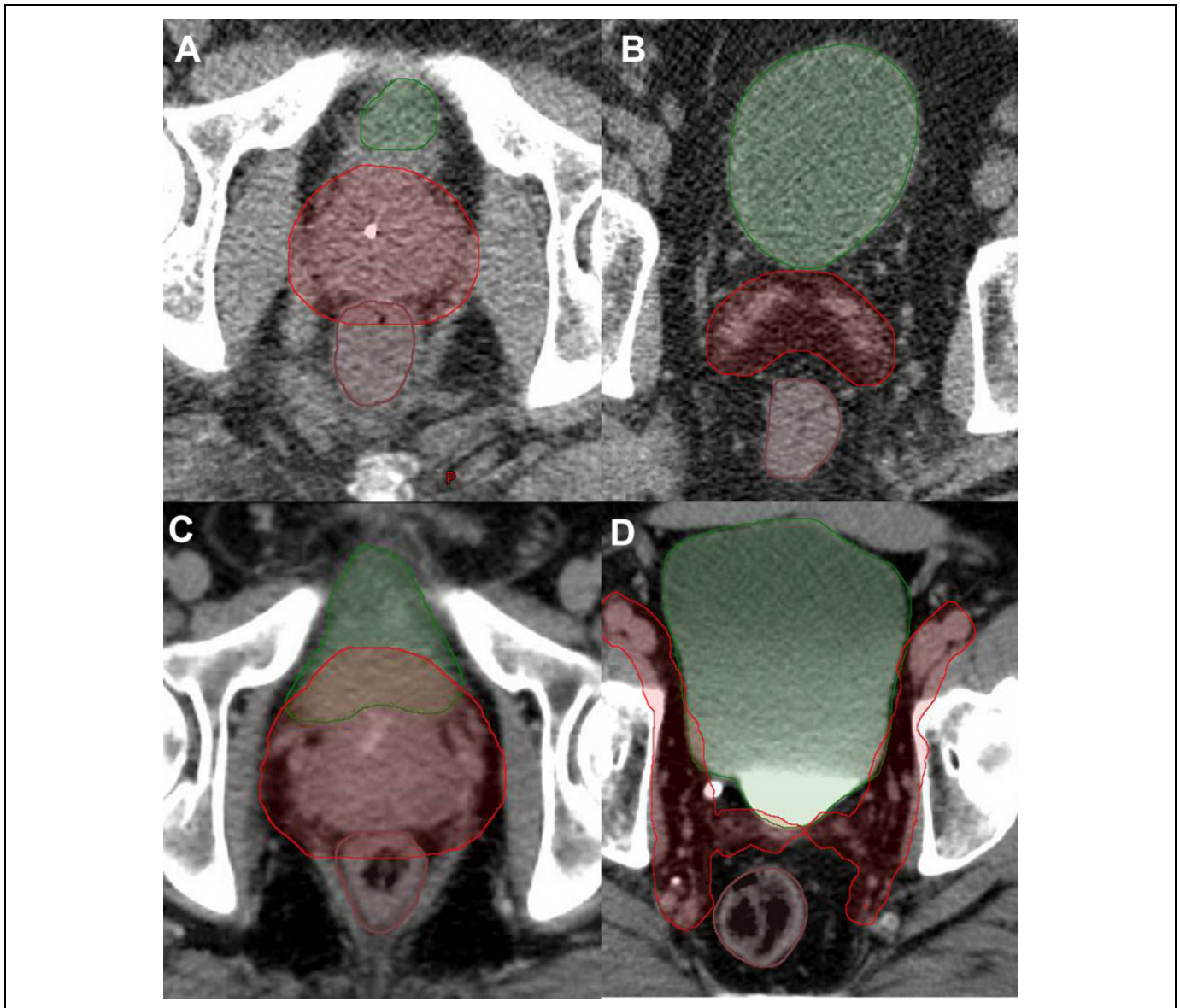


Figure 1. Example of anatomy comparison between prostate (A and B) PTV and prostate-plus-LN PTV (C and D). A, An example of axial slice of a prostate case which goes through the middle of the prostate. B, The axial slice of the same case as (A), which goes through the seminal vesicles. C, An example of axial slice of a prostate-plus-LN case which goes through the middle of the prostate. D, The axial slice of the same case as (C), which goes through the pelvic LN. LN indicates lymph node; PTV, planning target volume.

Materials and Methods

Materials

A total of 105 pelvic IMRT plans were retrospectively selected for this study. Eighty plans are clinical prostate IMRT plans. The other 25 plans are clinical prostate-plus-LN IMRT plans. Prostate-plus-LN IMRT plans were included to mimic novel anatomy with respect to prostate-only plans.

Knowledge Model Design

The current framework involves 2 types of predictive models, namely the multiple stepwise regression model and the

atlas-based model. The multiple stepwise regression model¹ has been commonly used as the methodology for knowledge modeling. For most inlier cases, that is, the prostate cases in this study, the framework will retrieve a regression-based model for prediction. For initial cases with novel anatomy, that is, the prostate-plus-LN cases, the framework adopts a dose atlas constructed from limited available training cases for prediction. When the number of cases with novel anatomy grows to a sufficient size, the Retain step of the case-based reasoning framework generates a regression-based model for future retrieval. Similar cases in the future will no longer be considered novel anatomy by the framework. In order to demonstrate the versatile workflow for various scenarios, we simulated 4

scenarios, namely *Scarce scenario*, *Semiscarce scenario*, *Semiample scenario*, and *Ample scenario*. The *Scarce scenario* occurs when the novel anatomy available is limited (5 cases). *Semiscarce scenario*, *Semiample scenario*, and *Ample scenario* occur when the novel anatomy has accumulated 10, 15, and 20 cases, respectively. Based on the simulation results, we finalized the workflow to deal with novel anatomy as the number of cases accumulates.

Regression model. In this study, we implemented the stepwise multiple-regression knowledge model proposed by Yuan et al¹ to predict OAR's dose-volume histograms (DVHs). The regression model predicts DVH's first three principal components (PCs) based on a set of carefully designed anatomical features, which include first three PCs of the DTH, OAR-PTV overlapping ratio, OAR outside treatment field ratio, OAR volume, and PTV volume. The stepwise multiple-regression model selects the features in forward selection manner. A feature is included in the model if the inclusion of it can provide statistically significant improvement of fit. Once features are selected, a multiple linear regression using all selected features is performed to establish the model.

Case-based reasoning using atlas. In addition to conventional regression knowledge modeling, we proposed a case-based reasoning framework that incorporates dose atlas guidance to supplement the regression model. We hypothesize that the introduction of case-based reasoning using atlas-based method could improve the overall model prediction accuracy. Atlas-based dose guidance has been implemented by Sheng et al for prostate cases.²¹ In this study, we construct a case atlas for prostate-plus-LN cases, which ask for different anatomy features/descriptors. The case-based dose atlas was constructed using N prostate-plus-LN cases, which was used for case-based dose prediction for novel anatomy in relation to the feature space of the regression model. For subsequent validation of prostate-plus-LN cases, each were first evaluated for geometric novelty by calculating the leverage metric²⁵ against N training prostate-plus-LN cases. The leverage score of each training case is defined as

$$h_i = (H)_i \quad (1)$$

where h_i is the i th diagonal element of the hat matrix $H = X(X^T X)^{-1} X^T$, and X is the feature matrix. If the leverage metric value for the query case was the largest among all $N + 1$ cases (N training + 1 query), the dose atlas was then used for dose prediction. The leverage has shown its potential in detecting cases that will perturb the regression model or be inaccurately predicted by the model as shown in a previous study.²⁵ We elect to turn to an atlas-based dose prediction method which is not sensitive to statistical noise. Otherwise, the regression model was used for prediction. The proposed regression model plus dose atlas strategy was compared against using regression model alone for prediction.

To build the atlas, the anatomy pattern was extracted via 3 anatomical features: topological connectivity, nodal separation, and nodal length, as illustrated in Figure 2. Topological connectivity C is a categorical descriptor for 2 major patterns: *disconnected*, as shown in the first row in Figure 2 and *connected* as shown in the second row in Figure 2. Nodal separation is defined as the distance between the center-of-mass of the left and right branch of the LN in the left-right direction. Nodal length is defined as the length of LN in the superior-inferior direction. These 3 features were identified as clinically relevant for dosimetric features around the target. In addition, such shape-based features help find the most similar atlas case available to generate case-based dose guidance.

Each of N training prostate-plus-LN cases served as a dose atlas case. For each of query prostate-plus-LN case i , the matched atlas case j is identified by Equation (2).

$$\min_j \{ \mathcal{L}(i) - \mathcal{L}(j)_2 + \delta(|C(i) - C(j)| - 1) \} \quad (2)$$

where \mathcal{L} is the 2-dimensional (2D) feature vector of nodal separation and nodal length; δ is the delta function. C is the topological connectivity (0 for disconnected and 1 for connected). The matched atlas case j was then linked with query case i using the deformable registration method in MIM Maestro System (MIM Software Inc, Cleveland, Ohio). The atlas case dose was subsequently transferred to the query anatomy through the deformation field, creating the goal dose. The goal dose served as the case-based dose guidance for the query novel anatomy. The workflow of the atlas construction and case-based reasoning guidance is shown in Figure 3. Dose guidance from the atlas was applied toward novel cases with deformable dose warping to account for anatomy variation between atlas and query cases. In this article, we refer "atlas" as a set of prototypical cases and in this study is composed of 5 prostate-plus-LN cases and later expanded as the cases accumulate. When we refer to "atlas case", it refers to an individual case in the atlas.

Scarce Scenario Simulation

The *Scarce scenario* was simulated with an initial case pool composed of 80 prostate cases and 5 prostate-plus-LN cases. Such 5 prostate-plus-LN plans were randomly selected from the 25 cases pool to represent the previously stored clinical cases, which were subsequently recruited to build the regression knowledge model together with the 80 prostate cases and used to construct the dose atlas, while the other 20 prostate-plus-LN cases were reserved to serve as query/validation cases. The case-based reasoning incorporated knowledge model prediction was compared against the prediction of the regression model. Figure 4 demonstrates the overall workflow of *Scarce scenario*.

Semiscarce, Semiample, and Ample Scenario Simulation

Retaining experience into memory, by adding current query case into the knowledge model training pool, should ideally

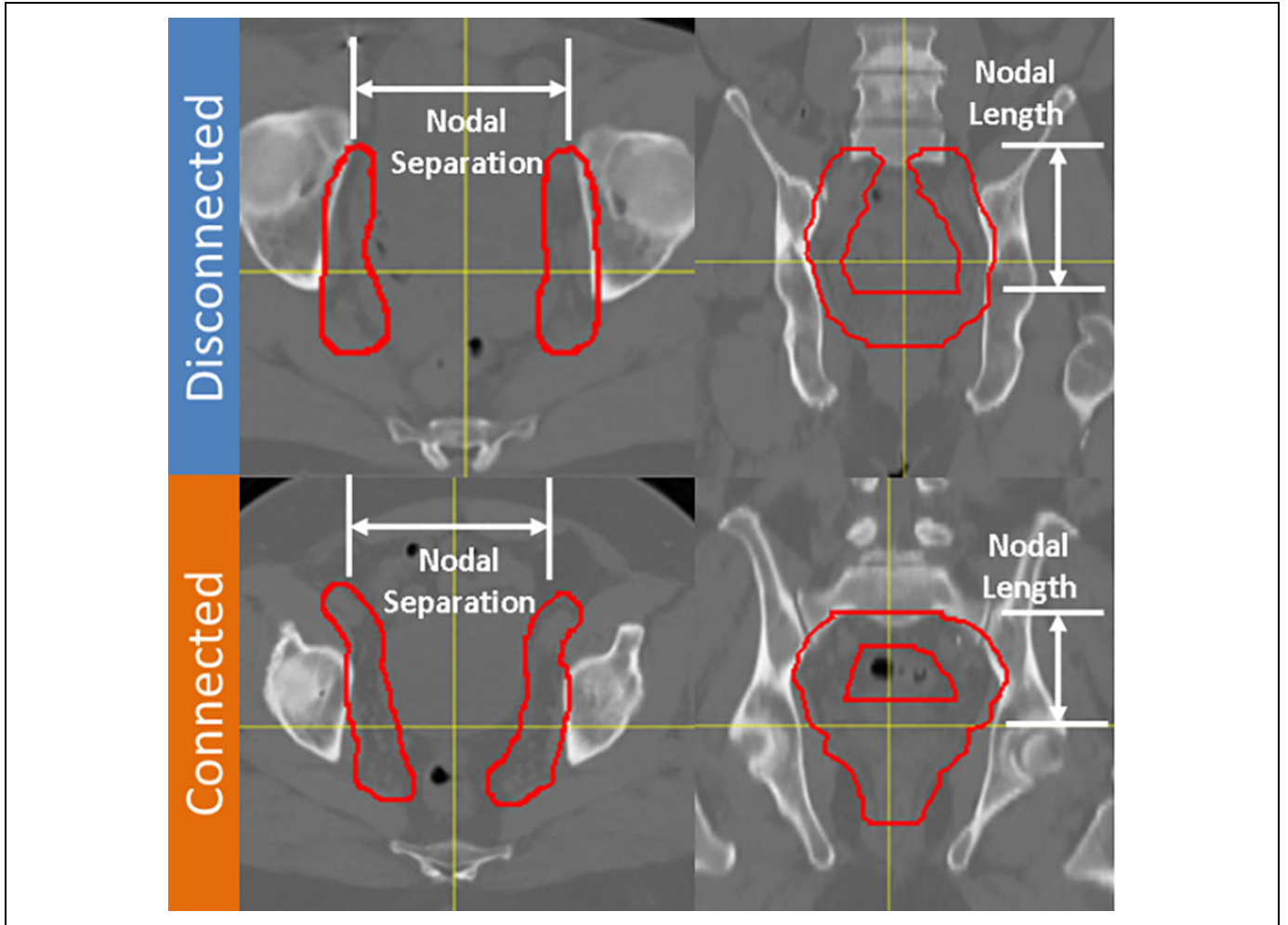


Figure 2. Illustration of parameterization of anatomy pattern for prostate-plus-LN anatomy. First row denotes disconnected LN in the superior portion of PTV; second row denotes connected LN in the superior portion of PTV. The other 2 features are nodal separation which is the center-of-mass distance between left and right branch of nodal PTV, and nodal length which is defined as the superior–inferior length of the nodal PTV. LN indicates lymph node; PTV, planning target volume.

boost model performance in the future practice. In order to validate this hypothesis, we simulated the *Semiscarce*, *Semiample*, and *Ample scenario* by retaining novel anatomy case in the knowledge model pool. An initial regression model \mathcal{M}_0 was trained using 80 prostate cases and 5 prostate-plus-LN cases. For *Semiscarce scenario*, the rest of the 20 prostate-plus-LN cases were divided into 4 folds. One fold was reserved for validation and another random fold was selected and added to the training pool to establish the regression model \mathcal{M}_1 , which is trained with 80 prostate cases and 10 prostate-plus-LN cases. For *Semiample scenario*, the rest of the 20 prostate-plus-LN cases were divided into 2 folds. One fold was reserved for validation and the other fold was selected and added to the training pool to establish the regression model \mathcal{M}_2 , which is trained with 80 prostate cases and 15 prostate-plus-LN cases. For the *Ample scenario*, the rest of the 20 prostate-plus-LN cases were divided into 4 folds. One fold was reserved for validation while the other 3 folds were added to the initial model to simulate the process of retaining. The *Expanded*

regression model \mathcal{M}_3 was trained using the original 80 prostate cases and 5 prostate-plus-LN cases, together with the newly added 15 prostate-plus-LN cases. Twenty cases were commonly regarded as minimally sufficient for statistical regression-based modeling, as demonstrated in RapidPlan’s Q&A document.²² In the alternative route of using case-based reasoning approach, novel cases were similarly retained into the atlas case pool under aforementioned 3 scenarios. The atlas-based model was invoked when necessary to generate the prediction for validation cases. The validation was repeated for all 4 folds. The prediction of each validation case from case-based reasoning assisting \mathcal{M}_0 , \mathcal{M}_1 , \mathcal{M}_2 , and \mathcal{M}_3 was individually compared against the 5-case atlas prediction using 1-tailed Wilcoxon Rank-Sum test. The null hypothesis is that the atlas-based method is not better (equal or inferior) than the regression-based method. In addition, the atlas-based prediction under various scenarios was compared against 5-case atlas prediction. The workflow for retaining is shown in Figure 5.

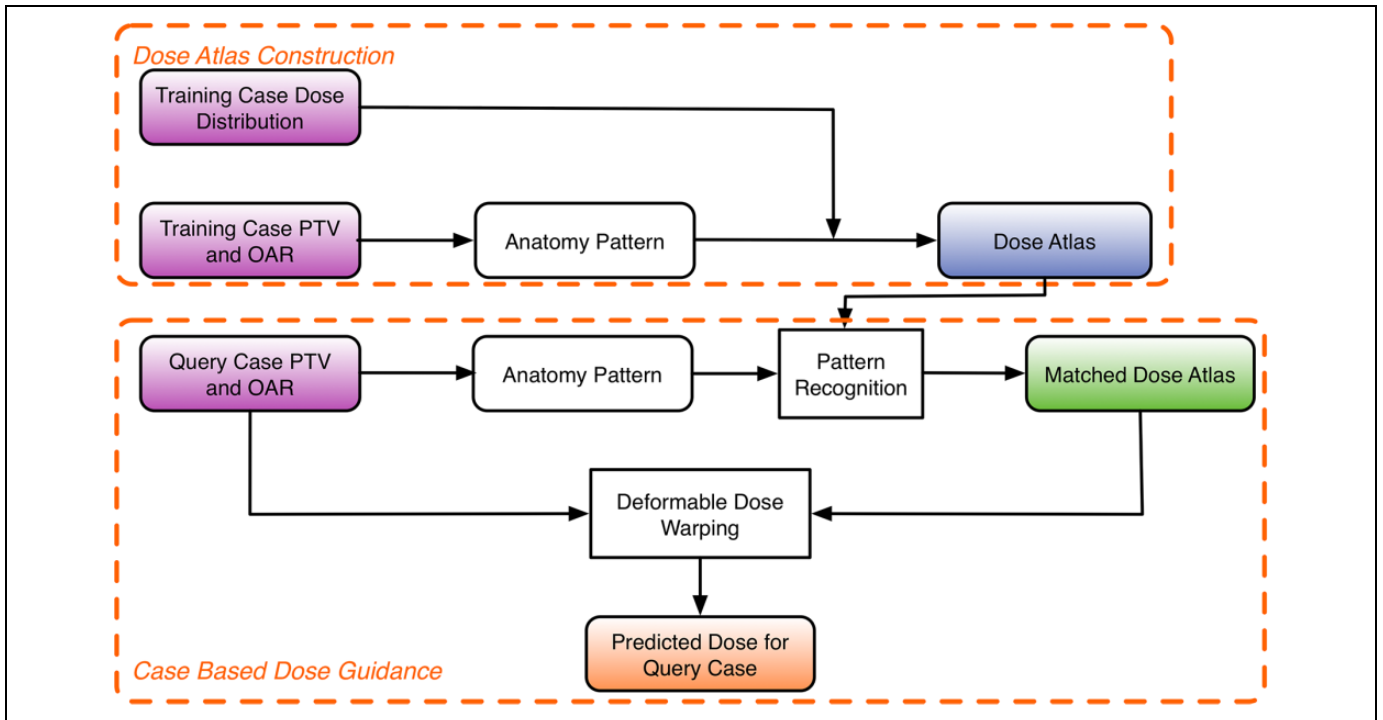


Figure 3. Flowchart for dose atlas construction (top box) and case-based reasoning dose guidance (bottom box).

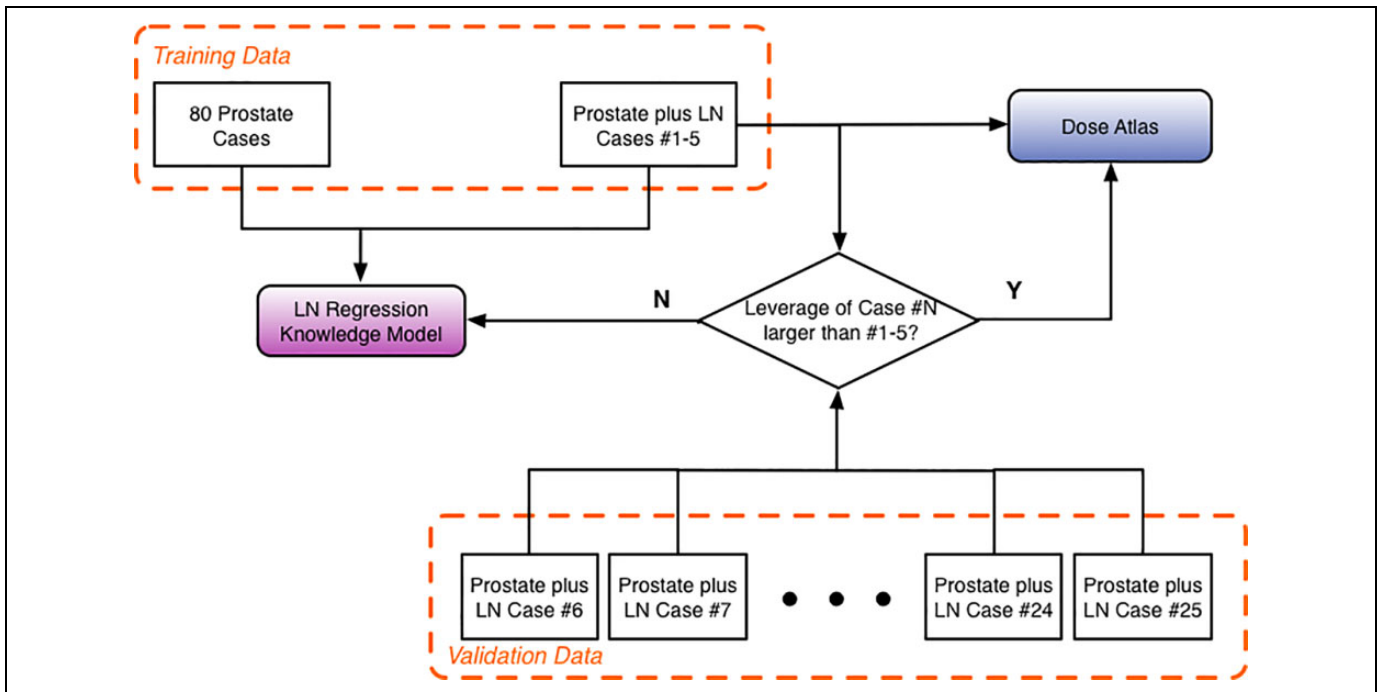


Figure 4. Flowchart of case-based reasoning framework for knowledge modeling using atlas-based guidance for prostate-plus-LN anatomy. LN indicates lymph node.

We also evaluated if the regression model can take over the atlas-based model under different scenarios. The regression models can offer several advantages over the atlas-based model, such as the fact that the regression model can be easily

transferrable once the model is trained and validated since the key information is the fitting parameters. However, the atlas-based model needs to carry the image, structure, and dose information at all times, making the model transfer difficult

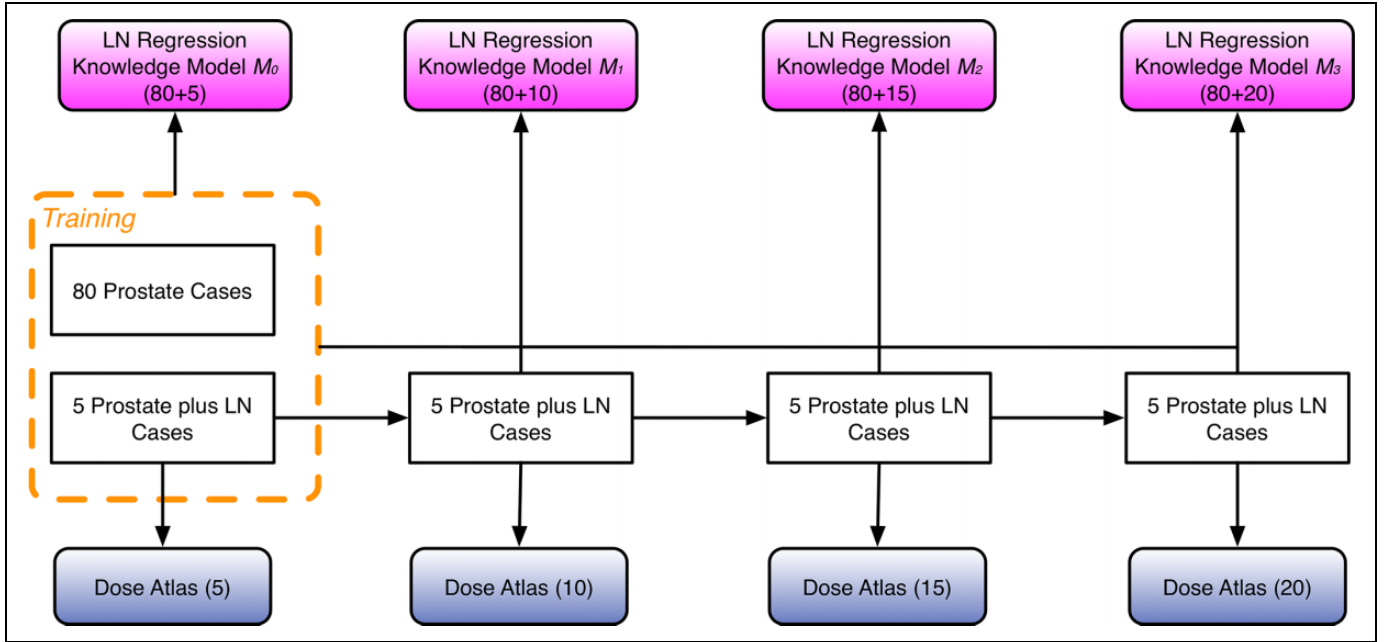


Figure 5. Workflow for retaining cases in regression model and atlas-based model. The number in the parenthesis indicates the number of cases available in that model.

without protocols. We hypothesize that the regression model can take over the atlas-based model if (1) statistical significant improvement over the 5-case regression model is established and (2) no statistical significance is observed between the regression model and the corresponding size atlas-based model. One-tailed Wilcoxon Rank-Sum test was performed for each comparison.

Performance Evaluation

The regression model and atlas-guided prediction were evaluated for DVH accuracy using the sum of squared residual (SSR):

$$SSR = \sum_{D=1}^{100} (V_{p,D} - V_{C,D})^2 \cdot \Delta D \quad (3)$$

where $V_{p,D}$ is the dose volume point for the predicted DVH at $\text{bin}D$; $V_{C,D}$ is the dose volume point for the clinical DVH at $\text{bin}D$.

Results

Scarce Scenario

Of the 20 validation cases, 13 prostate-plus-LN cases were identified as outliers and were subsequently guided by the dose atlas (8 cases needed case-based reasoning guidance for bladder, 7 cases needed case-based reasoning guidance for rectum, and 2 cases needed both). The 2D feature map (nodal length and nodal separation) for all training and validation prostate-plus-LN cases is shown in Figure 6. Connectivity is color-coded with connected group colored by red and nonconnected group colored by blue. Atlas-to-query match is shown by the

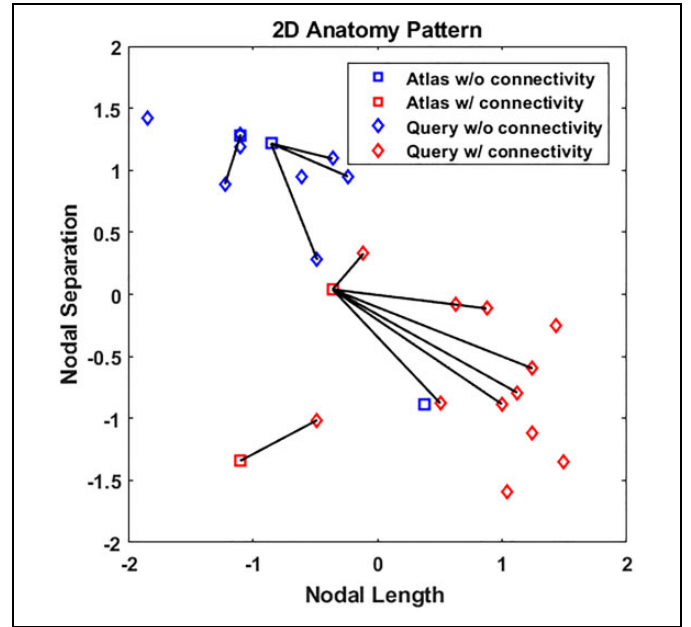


Figure 6. Two-dimensional feature space (nodal length and nodal separation) showing atlas and query cases. Atlas case is square mark and query case is diamond mark. Node connectivity is color coded (blue is disconnected and red is connected). Line connecting atlas and query cases denotes that atlas-based model was invoked and the atlas and query cases were matched.

black line connecting the atlas case (square) and the query case (diamond). Query cases without line connection mean that they were identified as inliers and were subsequently predicted using the regression model. Figure 7 shows the DVH SSR comparison between the regression model and dose atlas



Figure 7. Boxplot of DVH SSR comparison between regression knowledge model prediction and case-based dose atlas prediction. Left boxplot shows the comparison for the bladder and right boxplot shows the prediction for the rectum. The blue box denotes interquartile range. Red bar in the box denotes the median value. DVH indicates dose–volume histogram; SSR, sum of squared residual.

guidance for all 13 outlier validation cases. For the bladder, the DVH SSRs from 5-case atlas guidance (0.174 ± 0.166) were significantly lower than those (0.459 ± 0.508) from the regression model trained with 5 prostate-plus-LN cases added ($P = .0326$, 1-sided Wilcoxon Rank-Sum test). For the rectum, there was no significant difference (0.103 ± 0.120 and 0.150 ± 0.171) for case-based (5-case atlas) and regression prediction (5 prostate-plus-LN cases added), $P = .1972$, 1-sided Wilcoxon Rank-Sum test).

Figure 8 shows DVH of one example case guided by case-based reasoning. Green lines are DVHs for the bladder and brown lines are DVHs for the rectum. Solid lines are clinical plan's DVHs. Long-dashed lines are DVH predictions given by atlas-based guidance as a part of the case-based reasoning framework. Short-dashed lines are regression model-based predictions. For low-dose and high-dose regions, all 3 DVH groups (clinical plan, case-based prediction, and regression model prediction) agreed with each other for both OARs. For example, in intermediate dose region, regression model prediction overpredicted ($\sim 10\%$) for the bladder and rectum when compared to the clinical DVH. Case-based prediction, on the other hand, agreed well with the clinical DVH.

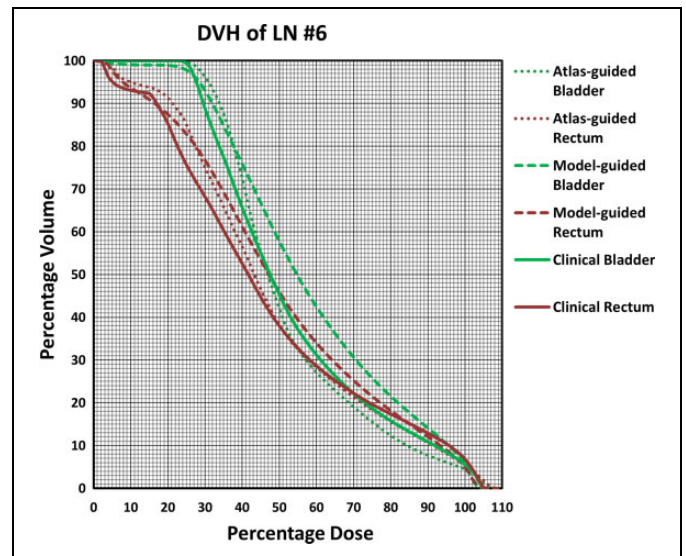


Figure 8. DVH comparison among clinical plan (solid line), regression model prediction (dashed line), and atlas-guided prediction (dotted line). Green DVH is bladder and brown DVH is rectum. Atlas-guided prediction agrees better with clinical DVH than regression model prediction, especially for median dose level. DVH indicates dose–volume histogram.

Semiscarce, Semiample, and Ample Scenario

For the *Semiscarce*, *Semiample*, and *Ample scenario*, Figure 9 shows the boxplot of the DVH error in the validation cases that compare predictions by the regression model \mathcal{M}_0 , \mathcal{M}_1 , \mathcal{M}_2 , and \mathcal{M}_3 , and the dose atlas guidance from 5, 10, and 15-case atlas, respectively. For a total of 13, 6, and 5 cases, the atlas-based model was invoked by case-based reasoning for *Scarce*, *Semiscarce*, and *Semiample scenario*, respectively. Twenty-case atlas was not invoked by case-based reasoning because no outlier was ever detected when 20 cases existed in the case pool. Thus, the result for the 20-case atlas was not shown in the plot. For the bladder, mean SSR of 5-case atlas, 10-case atlas, and 15-case atlas was 0.173 ± 0.166 , 0.119 ± 0.136 , and

0.201 ± 0.191 , respectively. The corresponding mean SSR of \mathcal{M}_0 , \mathcal{M}_1 , \mathcal{M}_2 , and \mathcal{M}_3 was 0.459 ± 0.508 , 0.346 ± 0.552 , 0.311 ± 0.516 , and 0.320 ± 0.577 , respectively. For the rectum, mean SSR of 5-case atlas, 10-case atlas, and 15-case atlas was 0.103 ± 0.120 , 0.097 ± 0.080 , and 0.142 ± 0.168 , respectively. The corresponding mean SSR of \mathcal{M}_0 , \mathcal{M}_1 , \mathcal{M}_2 , and \mathcal{M}_3 was 0.150 ± 0.171 , 0.142 ± 0.183 , 0.135 ± 0.171 , and 0.138 ± 0.184 , respectively. By retaining and accumulating novel cases to update the regression model, the median prediction accuracy is improved over original regression model for the bladder to reach the similar level as atlas-based prediction. Interquartile range is overall comparable between

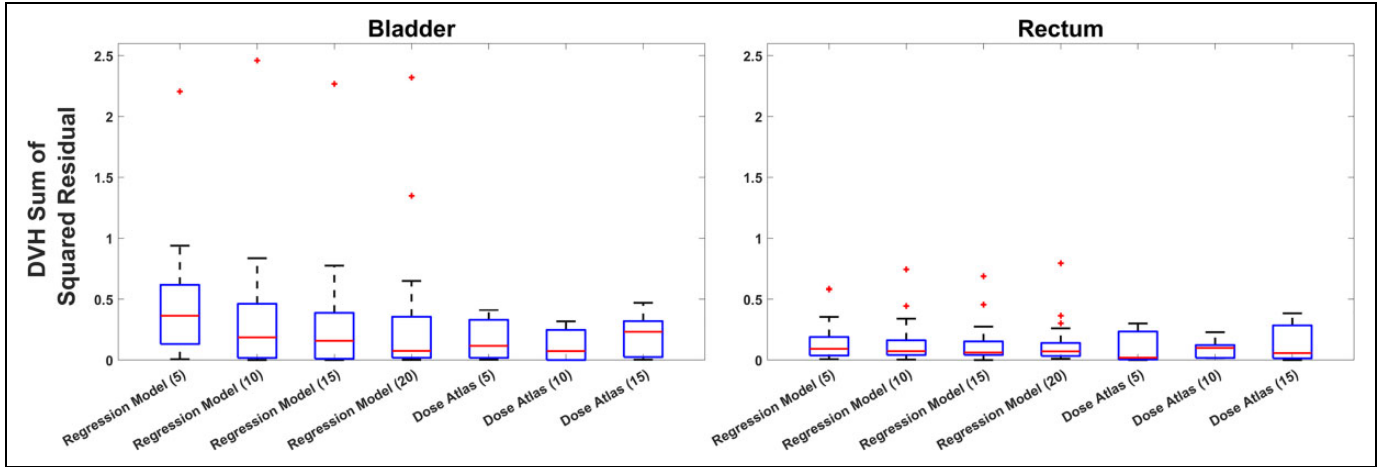


Figure 9. Boxplots of DVH error, sum of squared residual, for regression model \mathcal{M}_0 , \mathcal{M}_1 , \mathcal{M}_2 , and \mathcal{M}_3 and dose atlas guidance from 5, 10, and 15-case atlas. Bladder and rectum predictions are shown in left figure and right figure, respectively. The number in parenthesis is the number of prostate-plus-LN cases in respective model.

the *Expanded* regression model \mathcal{M}_3 and the dose atlas guidance for both the bladder and rectum and no significant difference was observed between medians. The result suggests that retaining novel geometry can improve the overall regression model prediction accuracy. Under the *Ample scenario*, the regression model achieved similar performance as the atlas-based method through retaining novel cases. Among regression models, statistical significant difference was only observed between \mathcal{M}_0 and \mathcal{M}_3 for the bladder ($P = .0398$), which indicates significant model performance improvement by retaining 20 cases. Among atlas-based models, no statistical significance was observed between different atlas sizes. For the comparison between \mathcal{M}_1 and 10-case atlas-based model, no statistical significance was observed for the bladder ($P = .2235$) or rectum ($P = .8551$). Similarly, no statistical significance was observed for bladder ($P = .9458$) or rectum ($P = .8385$) in the comparison between \mathcal{M}_2 and 15-case atlas-based model. Results showed that regression model could successfully take over from the atlas-based prediction when the novel anatomy accumulates over 20 pelvic cases.

Workflow for Using Case-Based Reasoning

We propose here a complete workflow to use case-based reasoning-assisted knowledge modeling for pelvic cases. When novel anatomy initially arises, for example, for the initial 5 prostate-plus-LN cases, the regression model does not predict well for these outlier cases as demonstrated in the studies by Delaney et al,²³ Tol et al,²⁴ and Sheng et al.²⁵ Since no prior knowledge exists, we recommend human intelligence to help provide clinical solutions in these instances. Human's interaction with planning novel cases can help feed new knowledge back to the knowledge model in the future model training/refining process. When novel cases accumulate to 5, an atlas size recommended by Sheng et al,²¹ the case-based reasoning framework will adopt dose atlas to provide prediction guidance as proposed in Knowledge Model Design Section. As more

novel cases arise, the case-based framework will retain them as training pool for the regression model while still maintaining the atlas-based method. As the number of novel cases reaches 20, the regression model can be learned and independently functions with satisfactory accuracy. The entire workflow is illustrated in Figure 10.

Discussion

In this study, we proposed a case-based reasoning framework for a radiation therapy knowledge model. Results showed that case-based prediction achieved better accuracy than the regression model when dealing with novel anatomy cases. Results also showed that retaining these novel cases into the regression model did boost the prediction accuracy of the regression model for future query cases. This study demonstrated that case-based reasoning that judiciously combines the use of an atlas-based prediction and regression-based prediction can help improve the overall robustness of the knowledge-based modeling especially when the existing data in the system are sparse or the new observation is novel to the existing system. In addition, the closed-loop feedback Retain step helps the knowledge-based model learn the novel anatomy pattern in order to be able to generalize for more cases. This study demonstrated that the 4-R steps of case-based reasoning can be implemented under the knowledge-based modeling framework to make it more robust and less prone to erroneous generalization for novel unseen cases. We also provided a systematic workflow to guide generating and/or predicting dose for novel anatomy under various scenarios. When the number of novel cases is small (eg, less than 5), manual planning is encouraged to leverage human knowledge for the interpretation of novel anatomy. As novel cases accumulate to a sufficient size (eg, more than 20), a regression model provides good prediction accuracy. An atlas-based model is primarily useful between 5 and 20 novel cases, a range where the novel knowledge is rapidly growing from the regression model's perspective. The proposed

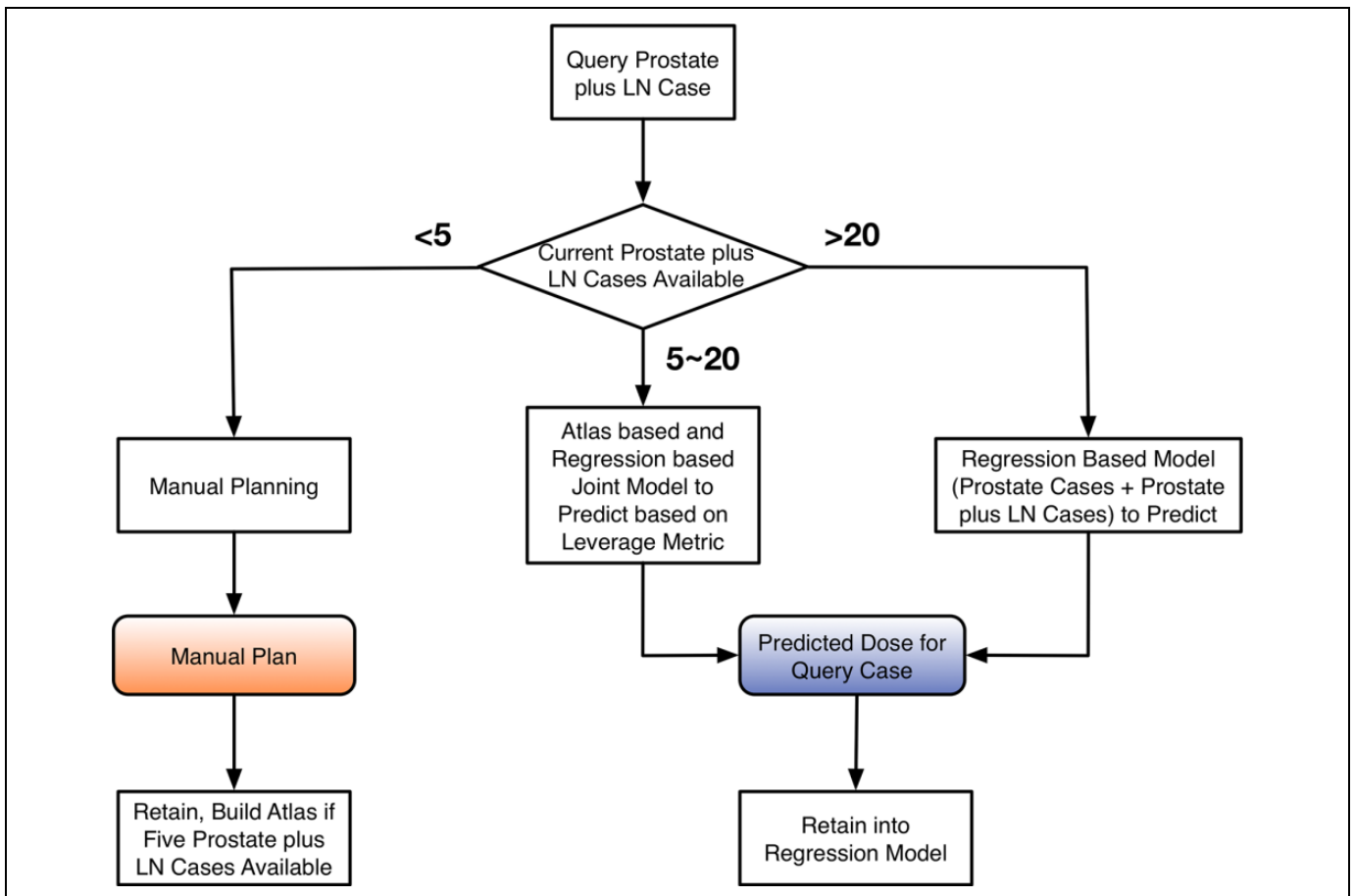


Figure 10. Case-based reasoning workflow for handling pelvic novel anatomy cases for different scenarios when using knowledge models. When the available prostate-plus-LN cases are less than 5, manual planning is encouraged. Atlas-based case-based reasoning is effective when the number of available novel cases is between 5 and 20. Case-based reasoning framework retires after 20 novel cases are accumulated and the regression model is solely responsible for prediction.

case-based reasoning framework addresses a major drawback of the conventional case-based and atlas-based knowledge models that require a large database of prior cases and are usually specific to one type of treatment sites or scenarios. The case-based reasoning framework could potentially integrate multiple regression models and multiatlas- (or case) based models into 1 overall knowledge modeling framework that can provide treatment planning guidance for various cancer sites. With the case-based reasoning framework, each case can be assigned to a specific local model, which is part of the general model. We are actively working along this direction.

The rationale of case-based reasoning originates from mimicking human planner's behavior when dealing with novel cases. Human planner's behavior is based on memory of training with similar cases. A good planner is capable of creating effective strategies based on past experience. Nowadays, machine modeling is repeating the first step by analytically parsing the anatomy and dosimetry relation, and as long as the anatomy pattern is within range of the training data, the prediction is mostly reliable. However, it is common that many patients have to be analyzed case by case, and they are

often referred as new knowledge. This is where case-based reasoning is helpful in terms of improving the system's overall robustness. And we need to deal with "Scarce scenario" which is also commonly seen in a clinical setting. Therefore, we believe the case-based reasoning framework provides a systematic approach to taking advantages of both the regression model and atlas-based method to build an overall enhanced and dynamically adaptive modeling scheme. The regression-based model requires sufficient numbers of training cases to reach optimal prediction accuracy, while the atlas-based approach can provide case-by-case guidance even if the novel knowledge is scarce. On the other hand, as the number of novel cases increases, both approaches show similar prediction accuracy with the regression model showing advantages. Once the regression model is trained, the training cases can be released from the model and the model can be easily transferred as a combination of model parameters. The overall prediction speed is faster as the atlas-based approach needs deformable registration and transferring dose. We believe the dual-model system is versatile and can adapt as the case available evolves.

This study is the first attempt to introduce case-based reasoning in radiation therapy knowledge modeling. The proposed case-based reasoning framework also fills the gap in translating knowledge models into effective clinical applications. While the specific design of the 4-R steps could vary for different knowledge models and for different clinical scenarios, the general principles of an intelligent system that learns from novel cases and accumulates new knowledge should remain the same and are well captured in the 4-R steps. We hope the introduction of case-based reasoning framework will provide a valuable foundation and inspire the future practice of handling knowledge models in complex clinical settings that will inevitably encounter novel scenarios. We anticipate that in the near future AI-based tool would be widely implemented and accepted in the clinic, and this study completes the final step of translating the tool into the clinic.

The 4-R steps of case-based reasoning framework add a layer on top of the original knowledge models which are known for inferior performance when predicting novelties. Specifically, the first 3-R steps address predicting and generating guidance for novel cases and the last R step, Retaining, is responsible for feeding new knowledge back to the knowledge model behind the scene. The 4-R steps within the case-based reasoning framework work collaboratively with each other and should not be separated.

We noticed in the result that there was less improvement provided by case-based reasoning for the rectum than for the bladder. The geometry change from treating prostate only to treating prostate-plus-LN affects the bladder more than the rectum. As shown in Figure 1, pelvic LN wraps around the bladder and changes the dose gradient inside the bladder entirely when compared to prostate cases. On the other hand, the rectum is less affected since the PTV shape around the rectum remains similar even with the inclusion of pelvic LN in the PTV although treating more superior component on top of the prostate results in a scaling effect of the DVH for the rectum. This is probably why the prostate model can still acceptably predict for the rectum for the prostate plus LN cases.

Case-based prediction showed superior accuracy for outlier/novel geometry than the regression model as shown by pelvic cases in this study. One limitation for the statistical regression knowledge-based model is that it needs certain amount of training cases to saturate for accurate prediction.²⁴ This number could vary for different treatment sites. This makes the regression model difficult to adapt to new patient cases when deployed clinically. The model has to be thoroughly evaluated and validated for all possible anatomy geometry before released for use and even after this, the generalizability of the model will always have limits. Sometimes it is not feasible due to the lack of cases from particular treatment sites. Alternatively, we can implement the case-based reasoning framework that incorporates an atlas-based model to boost the overall performance. In this study, we used 3 shape descriptors to cluster the high-dimensional shape feature space. The entire space was clustered into 5 subspaces,

with each atlas case responsible for predicting cases falling into its NN. Combined with deformable image registration, the warped dose from the atlas case can serve as a reasonable and clinically relevant prediction for the query novel anatomy. The regression model plus the case-based reasoning framework is the overall robust whether the data are sparse or not.

We constructed the current atlas based solely on the PTV's geometry. We did not include the shape feature of the OAR into constructing the atlas. The reason is 2-fold. First, the PTV shape is highly variant for prostate-plus-LN cases. Since the intermediate-to-high dose level should be conformal to the PTV, shape descriptors for the PTV could best categorize all cases to better guide the subsequent dose warping, making the warped dose with reasonable fall-off around the target, and achievable for the optimization. Second, the OAR spatial location relative to the target is relatively similar for pelvic cases. This assumption may not hold true for other treatment sites such as gastrointestinal cases where the bowel can form any shape around the target. To expand the framework to other treatment sites, special consideration such as site-specific handcrafted feature is needed when constructing the atlas to best reflect the relation between the dose and the target/OAR shape features. Substantial amount of effort is needed in this regard, and it could be a limitation for deployment in many clinics as it stands. Developing transferable case-based reasoning framework which respects patient privacy and data transfer protocol is an option. Future research along this line is warranted.

This study demonstrated the feasibility of implementing case-based reasoning framework using pelvic cases. The case-based reasoning framework could potentially be more important and meaningful for other treatment sites. The anatomy commonly has more variation than pelvic cases, which results in the fact that more cases are needed for the regression model to saturate. However, often the cases available are extremely sparse, such as for the liver or pancreatic stereotactic body radiation therapy. Case-based reasoning would be helpful in this context to make decision about dose constraints for the OAR or even dose sparing tradeoff among OARs. These are current challenges for clinical implementation of knowledge-based modeling, and case-based reasoning offers a solution. Further research along this line is under way.

Retaining novel case did show performance improvement for the regression model. This observation echoes the fact that the regression model needs to reach a certain number to saturate for predicting accurately. Based on the results, interestingly, we found that when retaining up to 10 or 15 cases, the regression model was not statistically different than the atlas-based model. However, the regression model continued to improve (red median bar in Figure 9) as more cases were retained, and with 20 cases, statistical significance was observed in the difference. These results suggest that the regression model could replace the atlas-based model when the number of cases reaches 20. Adding novel cases into the regression model adds to the feature space covered by the regression model and subsequently reduces the chance of seeing novel

anatomy in the future practice, which makes the regression model more robust against outliers. As more and more cases are retained, the chance of seeing outlier case is so small that the regression model reaches saturation for the specific treatment site. However, as new treatment techniques and treatment modalities arise, novel dose-anatomy patterns could appear again in the current model's context. The case-based reasoning's 4-R steps allow the framework to repeat learning and accumulating new knowledge.

Conclusion

In this study, a case-based reasoning framework was proposed and constructed that properly combines the use of a regression model for inlier cases (eg, prostate cases) and a dose atlas for novel cases (eg, prostate-plus-LN cases). The dose atlas served as a better prediction model when regression-based knowledge model is not suitable for prediction. Results showed that dose atlas guidance had superior prediction accuracy over the regression model when the number of novel case available is limited. A versatile workflow was provided to handle novel anatomy at different case number levels for pelvic plans. Establishing the case-based reasoning framework has the potential to improve the overall robustness of the clinical application of knowledge models.

Acknowledgments

The authors thank Hunter Stephens for assistance in proof reading this article.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.





Ethics Statement

This study has been approved by Duke University Health System Institutional Review Board under Amendment ID Amd002_Pro00034599. Consent was not needed since this study is retrospective analysis and operates under the protocol of "Retrospective Plan Feature Extraction and Modeling for Intensity Modulated Radiation Therapy".

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was partially supported by NIH/NCI 1R01CA201212 and a master research grant from Varian Medical Systems.

ORCID iD

Yang Sheng  <https://orcid.org/0000-0003-3380-1966>
 Jiahua Zhang  <https://orcid.org/0000-0002-4288-6503>
 Chunhao Wang  <https://orcid.org/0000-0002-6945-7119>
 Q. Jackie Wu  <https://orcid.org/0000-0001-8235-2322>

References

1. Yuan L, Ge Y, Lee WR, Yin FF, Kirkpatrick JP, Wu QJ. Quantitative analysis of the factors which affect the inter-patient organ-at risk dose sparing variation in IMRT plans. *Med Phys*. 2012;39(11):6868-6878.
2. Appenzoller LM, Michalski JM, Thorstad WL, Mutic S, Moore KL. Predicting dose-volume histograms for organs-at-risk in IMRT planning. *Med Phys*. 2012;39(12):7446-7461.
3. Wu B, Ricchetti F, Sanguineti G, 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):1241-1247.
4. Chanyavanich V, Das SK, Lee WR, Lo JY. Knowledge-based IMRT treatment planning for prostate cancer. *Med Phys*. 2011;38(5):2515.
5. Moore K, Scott Brame R, Low D, Mutic S. Experience based quality control of clinical intensity modulated radiotherapy planning. *Int J Radiat Oncol Biol Phys*. 2011;81(2):545-551.
6. Shiraishi S, Moore KL. Knowledge-based prediction of three-dimensional dose distributions for external beam radiotherapy. *Med Phys*. 2016;43(1):378-387.
7. Zhang J, Wu QJ, Xie T, Sheng Y, Yin F-F, Ge Y. An ensemble approach to knowledge-based intensity-modulated radiation therapy planning. *Front Oncol*. 2018;8:57.
8. 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):719-726.
9. Ziemer BP, Shiraishi S, Hattangadi-Gluth JA, Sanghvi P, Moore KL. Fully automated, comprehensive knowledge-based planning for stereotactic radiosurgery: Preclinical validation through blinded physician review. *Pract Radiat Oncol*. 2017;7(6):e569-e578.
10. Li N, Carmona R, Sirak I, 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):164-172.
11. Yuan L, Wu QJ, Yin FF, Jiang Y, Yoo D, Ge Y. Incorporating single-side sparing in models for predicting parotid dose sparing in head and neck IMRT. *Med Phys*. 2014;41(2):021728.
12. Yuan L, Wu QJ, Yin F, et al. Standardized beam bouquets for lung IMRT planning. *Phys Med Biol*. 2015;60(5):1831-1843.
13. Yuan L, Zhu W, Ge Y, et al. Lung IMRT planning with automatic determination of beam angle configurations. *Phys Med Biol*. 2018;63(13):135024.
14. Sheng Y, Li T, Yoo S, et al. Development of an ultra-fast, high-quality whole-breast radiation therapy treatment planning system. *Int J Radiat Oncol Biol Phys*. 2016;96(2):S228.
15. Lu L, Sheng Y, Donaghue J, et al. Three IMRT advanced planning tools: A multi-institutional side-by-side comparison. *Journal of Applied Clinical Medical Physics*. 2019;20(8):65-77.
16. Sheng Y, Li T, Yoo S, et al. Automatic planning of whole breast radiation therapy using machine learning models. *Frontiers in Oncology*. 2019;9(750):1-8.
17. Wang W, Sheng Y, Yoo S, Blitzblau RC, Yin F-F, Wu QJ. Goal-driven beam setting optimization for whole-breast radiation therapy. *Technology in cancer research & treatment*. 2019;18:1533033819858661.
18. Li X, Wu J, Palta M, et al. A collimator setting optimization algorithm for dual-arc volumetric modulated arc therapy in

- pancreas stereotactic body radiation therapy. *Technology in cancer research & treatment* 2019;18:1533033819870767.
19. Zhang J, Wu QJ, Ge Y, et al. Knowledge-based statistical inference method for plan quality quantification. *Technology in cancer research & treatment* 2019;18:1533033819857758.
 20. Good D, Lo J, Lee WR, Wu QJ, Yin FF, Das SK. 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):176-181.
 21. Sheng Y, Li T, Zhang Y, et al. Atlas-guided prostate intensity modulated radiation therapy (IMRT) planning. *Phys Med Bio.* 2015;60(18):7277-7291.
 22. Varian Medical Systems. Rapid Plan Knowledge-based Planning Frequently Asked Questions. https://www.varian.com/sites/default/files/resource_attachments/RapidPlanFAQs_RAD10321B.pdf
 23. Delaney AR, Tol JP, Dahele M, Cuijpers J, Slotman BJ, Verbakel WF. Effect of dosimetric outliers on the performance of a commercial knowledge-based planning solution. *Int J Radiat Oncol Biol Phys.* 2016;94(3):469-477.
 24. Tol JP, Delaney AR, Dahele M, Slotman BJ, Verbakel WF. Evaluation of a knowledge-based planning solution for head and neck cancer. *Int J Radiat Oncol Biol Phys.* 2015;91(3):612-620.
 25. Sheng Y, Ge Y, Yuan L, Li T, Yin F-F, Wu QJ. Outlier identification in radiation therapy knowledge-based planning: a study of pelvic cases. *Med Phys.* 2017;44(11):5617-5626.
 26. Sheng Y, Li T, Lee WR, Yin F-F, Wu QJ. Exploring the margin recipe for online adaptive radiation therapy for intermediate-risk prostate cancer: an intrafractional seminal vesicles motion analysis. *Int J Radiat Oncol Biol Phys.* 2017;98(2):473-480.
 27. De Mántaras RL, McSherry D, Bridge D, et al. Retrieval, reuse, revision and retention in case-based reasoning. *Knowledge Eng Rev.* 2005;20(03):215-240.