Modeling the dosimetry of organ-at-risk in head and neck IMRT planning: An intertechnique and interinstitutional study

Jun Lian[a)]
Department of Radiation Oncology, The University of North Carolina, Chapel Hill, North Carolina 27599

Lulin Yuan[b)]
Department of Radiation Oncology, Duke University, Durham, North Carolina 27710

Yaorong Ge
Department of Software and Information Systems, The University of North Carolina, Charlotte, North Carolina 28223

Bhishamjit S. Chera
Department of Radiation Oncology, The University of North Carolina, Chapel Hill, North Carolina 27599

David P. Yoo
Department of Radiation Oncology, Duke University, Durham, North Carolina 27710

Sha Chang
Department of Radiation Oncology, The University of North Carolina, Chapel Hill, North Carolina 27599

FangFang Yin and Q. Jackie Wu[c)]
Department of Radiation Oncology, Duke University, Durham, North Carolina 27710

(Received 25 June 2013; revised 13 October 2013; accepted for publication 14 October 2013; published 7 November 2013)

Purpose: To build a statistical model to quantitatively correlate the anatomic features of structures and the corresponding dose-volume histogram (DVH) of head and neck (HN) Tomotherapy (Tomo) plans. To study if the model built upon one intensity modulated radiation therapy (IMRT) technique (such as conventional Linac) can be used to predict anticipated organs-at-risk (OAR) DVH of patients treated with a different IMRT technique (such as Tomo). To study if the model built upon the clinical experience of one institution can be used to aid IMRT planning for another institution.
Methods: Forty-four Tomotherapy intensity modulate radiotherapy plans of HN cases (Tomo-IMRT) from Institution A were included in the study. A different patient group of 53 HN fixed gantry IMRT (FG-IMRT) plans was selected from Institution B. The analyzed OARs included the parotid, larynx, spinal cord, brainstem, and submandibular gland. Two major groups of anatomical features were considered: the volumetric information and the spatial information. The volume information includes the volume of target, OAR, and overlapped volume between target and OAR. The spatial information of OARs relative to PTVs was represented by the distance-to-target histogram (DTH). Important anatomical and dosimetric features were extracted from DTH and DVH by principal component analysis. Two regression models, one for Tomotherapy plan and one for IMRT plan, were built independently. The accuracy of intratreatment-modality model prediction was validated by a leave one out cross-validation method. The intertechnique and interinstitution validations were performed by using the FG-IMRT model to predict the OAR dosimetry of Tomo-IMRT plans. The dosimetry of OARs, under the same and different institutional preferences, was analyzed to examine the correlation between the model prediction and planning protocol.
Results: Significant patient anatomical factors contributing to OAR dose sparing in HN Tomotherapy plans have been analyzed and identified. For all the OARs, the discrepancies of dose indices between the model predicted values and the actual plan values were within 2.1%. Similar results were obtained from the modeling of FG-IMRT plans. The parotid gland was spared in a comparable fashion during the treatment planning of two institutions. The model based on FG-IMRT plans was found to predict the median dose of the parotid of Tomotherapy plans quite well, with a mean error of 2.6%. Predictions from the FG-IMRT model suggested the median dose of the larynx, median dose of the brainstem and D2 of the brainstem could be reduced by 10.5%, 12.8%, and 20.4%, respectively, in the Tomo-IMRT plans. This was found to be correlated to the institutional differences in OAR constraint settings. Re-planning of six Tomotherapy patients confirmed the potential of optimization improvement predicted by the FG-IMRT model was correct.
Conclusions: The authors established a mathematical model to correlate the anatomical features and dosimetric indexes of OARs of HN patients in Tomotherapy plans. The model can be used for the setup of patient-specific OAR dose sparing goals and quality control of planning results.

The institutional clinical experience was incorporated into the model which allows the model from one institution to generate a reference plan for another institution, or another IMRT technique.
 [http://dx.doi.org/10.1118/1.4828788]

Key words: IMRT, intertechnique, interinstitutional, prediction

1. INTRODUCTION

Intensity modulated radiation therapy (IMRT) has provided us with the ability to spare the surrounding critical organs-at-risk (OAR) while achieving prescription coverage to the planning target volume (PTV).[1–3] Although IMRT planning has become more intelligent and automatic,[4–6] the input of the planners is still the essential driving force for the quality of the plan. One major difficulty of IMRT planning is the unknown feature of best achievable dosimetry at the beginning of optimization. Currently a more optimal plan normally requires the diligent effort of trial and error of the planner to guide the optimization manually. Therefore the quality of the IMRT plan depends highly on the experience of the planner, time available for the case and institutional/individual dose constraints set by physicians. While quality assurance (QA) has become a main concern in the field of radiation therapy, the QA process of IMRT plans normally only checks the agreement between the directive of physician and the treatment plan, and between the calculated results by computer and measured results on the Linac. The dosimetry of a new plan can be compared with a standard protocol, but it cannot be verified for a particular patient whether the plan is as optimal as other plans or not.

It has become an important topic to quantify the geometrical distribution of organs and the ability of dosimetric avoidance of IMRT plans.[5,7–9] For example, in a study of the dose sparing of parotid glands for head-and-neck (HN) cancer patients, Hunt *et al.* found that percent gland volume overlapping with PTV could predict the mean gland dose.[7] In prostate cancer cases, Reddy *et al.* reported that the mean dose to rectum and bladder increased with increasing prostate volume.[8] Moore *et al.* demonstrated a correlation between the fraction of OAR volume overlapping with PTV and the OAR mean dose, and used the correlation to formulate a tool of quality control for the IMRT plans.[9] Further, Zhu *et al.* proposed the concept of distance-to-target histogram (DTH) to establish the correlation between the OAR-PTV anatomy and OAR dose-volume histograms (DVHs).[10] Wu and his colleagues constructed a database with target-OAR overlap volume histogram (OVH) and DVH information of 91 HN patients.[11] From the OVH analysis, the initial planning goals for IMRT optimization were generated from the OVH analysis of the database. Recently, Yuan *et al.* analyzed a database of high-quality prior plans including 64 prostate and 82 HN cases.[12] They quantified the effects of an array of patient anatomical features of the PTV and OARs and their spatial relationships on the interpatient OAR dose sparing variation in IMRT. This evidence-based model was able to reliably predict OAR dose sparing that was achievable based on the best available prior experience. The concept of "refined model" was introduced by Appenzoller *et al.* to exclude the outliers indicated by the "average model" from the training cohort of all patients.[13] The suboptimal plans were correctly identified by the model which guided the replanning with better OAR sparing achieved afterwards.

In these previous studies, all IMRT techniques used in modeling were the fixed gantry IMRT (FG-IMRT). Further, all the data used for modeling and validation were from the same institution and from the same patient group. Tomotherapy provides another way of intensity modulation and in theory can produce better dose distributions than static gantry IMRT for complex cases.[14] However, there is no published work to analyze the correlations between the anatomy and achieved dosimetry for Tomotherapy patients. The commonly used simultaneous integrated boost (SIB) treatment scheme of Tomotherapy also poses special difficulty of modeling the anatomy and dosimetry relations.[13] While a few publications have demonstrated that a well-established model can predict the DVH of a new patient treated by the same technique,[9–11] it is unknown whether the model is still valid when it is applied to another IMRT treatment modality. The planning process is normally governed by the institutional guidelines, and the quality of the plan is inevitably affected since the guidelines are often used as the objectives for which the planner strives. It has been found that the plan quality of IMRT was inconsistent among institutions.[15–17] Further, the training of model based on high quality plans to correlate the anatomic features and optimized dosimetry of structures naturally incorporates the best clinical experience of an institution. It is unclear how the experience of one institution can benefit the treatment planning of another institution which employs different protocols and planning approaches.

In this paper, we extended our previous work of quantitative analysis of the factors affecting the organ-at-risk dose sparing of FG-IMRT planning[12] to Tomotherapy planning. We applied our model built on the FG-IMRT plans of HN patients treated at Institution B to predict the dosimetry of Tomotherapy plans of patients treated at Institution A. The predicted and actual DVHs of all structures were compared. In this way, we investigated the extent to which a model's dependence on a particular IMRT treatment technique and the clinical experience of institutions with different protocols can be useful to each other.

2. METHODS AND MATERIALS

2.A. Patient treatment plans

Forty-four former Tomotherapy HN plans with concurrent boost were included in the study. These plans were randomly selected from the clinical plans used for patient treatment at

TABLE I. Template of OAR dose constraints for HN Tomotherapy planning in Institution A and range of OAR dose constraints for HN FG-IMRT planning in Institution B (Institution B specifies the constrict of patient case by case.).

Anatomic structures	Institution A (Tomotherapy)	Institution B (FG-IMRT)
Brainstem	Max: 54 Gy to point dose	Max: 25–35 Gy
Spinal cord	Max: 50 Gy to point dose	Max: 40–45 Gy
Larynx	Mean dose $\leq$ 41 Gy	Median dose: 15–30 Gy
	Volume receiving 60 Gy $\leq$ 24%	
Parotid	Mean dose $<$ 26 Gy	Median dose: 5–26 Gy
	50% receives $<$ 30 Gy	
Submandibular	Mean dose $<$ 35 Gy	Median dose: $<$ 30 Gy
Oral cavity	Mean dose $<$ 39 Gy	Median dose: 10–40 Gy
Cochlea	Mean dose $<$ 45 Gy	Max dose 30–45 Gy
Unspecified tissue outside of PTV	$\leq$1% receives $>$ 110% of the Rx	$\leq$1% receives $>$ 110% of the Rx

Institution A. The primary target dose was 54–60 Gy and the boost target dose was 60–70 Gy. Tomotherapy delivery was on a TomoTherapy HD machine (Accuray) in helical mode with simultaneous table translation and gantry rotation. Beams were 6 MV and distributed evenly from 51 directions. The modulation factor used was 0.287–0.35, pitch 2.6–3.5, and longitudinal field width 2.5 cm. Leaf width was 6.25 mm at the isoplane (SAD 85 cm). Beamlet calculation was carried out prior to the optimization to accelerate the dose calculation during each iteration of optimization. The beam-on time of average HN treatment was 7–10 min. Treatment planning was performed with TomoTherapy planning software (TomoHD 1.2.1, Accuray) under the guidance of an institutional prescription and OAR sparing template, as listed in Table I. When training the DVH model for an OAR, only the cases in which the OAR dose sparing satisfied the institutional template were used.

Fifty-three FG-IMRT HN plans were retrospectively analyzed. These plans were randomly selected from the clinical plans performed at Institution B. All patients were treated in sequential boost fashion. The prescription was 44 Gy or 50 Gy to the primary PTV and 70 Gy to boost PTV. Nine equally spaced coplanar 6 MV beams were used. Leaf width was 5 mm at the isoplane (SAD 100 cm). Treatment was delivered on Varian Trilogy and 2100EX linear accelerator with treatment time 6–10 min. Planning was performed with Eclipse Version 10.0 (Varian). Institution B does not use a default dose constraint protocol for HN IMRT planning. Instead, the dose constraints were prescribed case by case by the physicians and they were tighter than The Radiation Therapy Oncology Group (RTOG) guidelines.[18] The ranges of the dose constraints for major OARs are listed in Table I. The training set for an OAR DVH model did not include the cases in which there is no physician prescription for such OAR or the cases in which the OAR dose sparing did not meet the physician's prescription.

The primary tumor sites of two patient groups were similar including oropharynx, base of tongue, tonsil, laryngopharynx, and nasopharynx. The modeled OARs included the parotid, brain stem, spinal cord, larynx, and submandibular gland (SMG). The model prediction and actual results of the parotid, brainstem, and larynx were presented in detail. The institutional review boards of two institutions approved this patient study involving de-identified retrospective data.

2.B. Anatomical features and correlation model

An array of anatomical features was analyzed to study their contributions to OAR dose sparing. The OARs and their anatomical features analyzed for this study are listed in Table II. The distance to target histogram (DTH) encodes the spatial relationship between the OARs and the PTV.[10] When the Euclidean form of the distance function is used, DTH is equivalent to overlap volume histogram (OVH) defined by Wu *et al.*[5] In the Euclidean space, DTH at a distance d is calculated as the fraction of OAR volume with its maximum distance to the PTV surface (S_{PTV}) less than d. The distance function r from an OAR voxel v^i_{OAR} to the PTV surface is defined as $r(v^i_{OAR}, \mathrm{PTV})$,

$$r\left(v^i_{OAR}, \mathrm{PTV}\right) = \min_k \left\{ \left\| v^i_{OAR} - v^k_{PTV} \right\| \; \middle| \; v^k_{PTV} \in S_{PTV} \right\}.$$

In the actual implementation, two modifications to the Euclidean distance are made in order to account for the increased dose fall-off rate for the voxels outside the primary treatment fields (out-of-field) and to account for the slower dose fall-off in the regions far away from the PTV. Principal component analysis (PCA) tool is widely used to identify the features hidden in high dimension data.[10,19] For

TABLE II. Anatomical and dosimetric features considered in the Tomotherapy models.

Anatomical and dosimetric features
Distance to primary and boost target histogram (DTH)
OAR volumes
Primary and boost PTV volume
Fraction of OAR volume overlapping with PTVs (overlap volume)
Fraction of OAR volume outside the treatment fields (out-of-field volume)
Dose prescription to the primary and boost PTVs
Primary PTV D95
Boost PTV dose homogeneity (D2-D99)

instance the first order of PC is related to the mean of the histogram and the second order is related to the dominant gradient within a range. Similar to our previous study,[10] we used the first three principal components to quantitatively characterize the DVH and DTH curves with greatly reduced dimensions. The correlation between the principal components of DVH and patient anatomical features was obtained by a stepwise multiple regression method. The stepwise regression method includes multiple related anatomical factors in the beginning and then eliminates the least significant one at each step of regression during the process of modeling. The significance factors were identified with p-value < 0.05. The details of anatomical factors and modeling of FG-IMRT planning were described in previous studies.[10,12]

2.C. Model validation

The validation of models includes three parts. (1) *Intrapatient group*: The model was trained using all the patients in the database of the same group except the test case (leave-one-out). The OAR DVH of the validation case was predicted by the model and compared with the actual plan DVH. The dosimetric parameters such as the median dose were computed from the modeled DVHs and then compared with the DVHs of the actual plan. (2) *Intertechnique*: At this step, we studied the effect of treatment modality by fixing the clinical objective first. The parotid was used as an example OAR since a similar constraint was adopted by two institutions in their respective FG-IMRT and Tomotherapy planning. The parotid DVH models trained with the FG-IMRT plans were applied to the Tomotherapy cases. The predicted DVHs by the FG-IMRT models were compared with the actual curves of Tomotherapy plans in a leave-one-out fashion. (3) *Interinstitutional planning protocol comparison*: After the previous step of intertechnique study, we changed the clinical objective using the larynx (a parallel structure) and brainstem (a serial structure) as the examples. Two institutions have very different dosimetric requirements of the brainstem and larynx and consequently may result in different planning efforts as well as dose sparing. The FG-IMRT model predicted and actual Tomotherapy plan dosimetric endpoints of these two structures were compared. In all the comparisons for model validation, a 10% of target prescription dose value (7 Gy) is defined as the error bound.

3. RESULTS

3.A. Factors of interpatient variation in the training dataset

We have modeled the OAR DVHs for the parotid, larynx, spinal cord, brainstem, and SMG in the Tomotherapy plans. An array of anatomical features was screened for individual OARs, respectively. The combination of these factors had strong correlation with the DVH PCS1 (1st of the three most significant principle components) of all OARs. The multiple determination coefficients ranged from 0.63 to 0.88. DVH PCS2 (2nd of the three most significant principle components) of OARs except cord and larynx also showed high combined coefficients with the significant anatomical factors (0.60–0.78). We expanded our previous modeling work of FG-IMRT to include more OARs of the HN region.[12] Similar patterns of correlation were found in the FG-IMRT model and the Tomotherapy model.

3.B. Accuracy of model prediction within patient group

The DVHs of the parotid, larynx, spinal cord, brainstem, and SMG for Tomotherapy plans were calculated by the multiple regression model in a leave-one-out fashion. The model-predicted DVHs were compared to their corresponding DVHs in the actual plans to assess the effectiveness of the model. The comparison of the parotid DVH for a sample case is shown in Fig. 1(a). The Tomotherapy-based-model predicted DVH (solid red curve with maximum dose 55 Gy) was a good

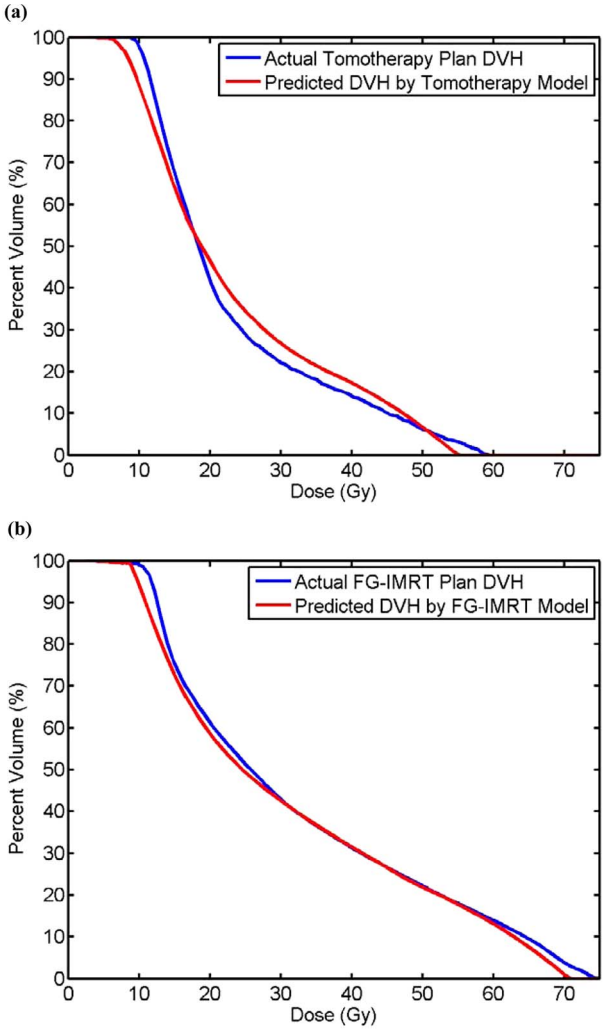

FIG. 1. Comparison of model predicted DVH and actual DVH of the parotid of a sample clinical case. (a) Tomotherapy model (solid red curve with maximum dose 55 Gy) and actual DVH of a Tomotherapy plan. (solid blue curve with maximum dose 59 Gy) (b) FG-IMRT model (solid red curve with maximum dose 70 Gy) and actual DVH of a FG-IMRT plan (solid blue curve with maximum dose 73 Gy).

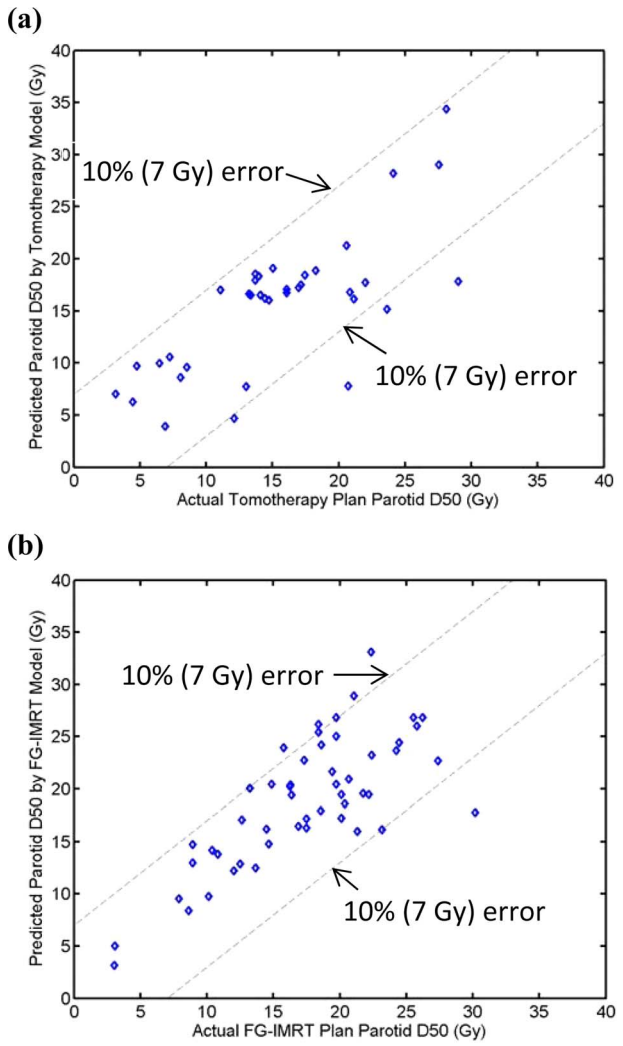

FIG. 2. Correlation of model predicted parotid median dose with actual clinical values. (a) Tomotherapy model vs clinical values for Tomotherapy cases. (b) FG-IMRT model vs clinical values for FG-IMRT cases.

estimate of actual DVH of Tomotherapy plan (solid blue curve with maximum dose 59 Gy). Figure 2(a) shows the correlation of model predicted parotid median dose with actual clinical values for the parotids in all the Tomotherapy cases. The differences between the modeled median dose values and the clinical values were within 10% error bound in 33 of the 37 parotids. The difference of the median dose between the predicted and actual plans was 1.5% on average, with the maximum positive deviation of 8.5% and negative deviation of 18.4%.

For the FG-IMRT dataset, the predicted OAR DVHs generated by the FG-IMRT model were also compared with the actual plan DVHs in a leave-one-out fashion. Consistent with our finding in Tomotherapy study, in most of the plans the predicted and actual plan DVHs were well matched. The modeled and the actual parotid DVH for a sample case are shown in Fig. 1(b). The differences between the modeled median dose values and the clinical values for 46 out of 52 parotids were within 10% error bound [Fig. 2(b)]. The difference of the median dose between the predicted and actual plans was 0.9% on average, with the maximum positive deviation of 15.3% and negative deviation of 17.7%.

3.C. Accuracy of model prediction across treatment modalities

The parotid was spared in a similar way by both institutions. When the parotid overlapped the PTV, the coverage of target often took the priority. The hot spot was usually excluded in the overlapped area. The mean dose and volume of low dose area of parotid were optimized as low as possible. In this way, the overall curve of parotid DVH was driven similarly in plans designed by two institutions, regardless of IMRT techniques. Using the model built for the FG-IMRT patients, we predicted the DVHs of the parotid of Tomotherapy patients. In Fig. 3, we showed the results from one sample patient (a) to a group of patients (b) and statistical results (c). Figure 3(a) shows FG-IMRT model predicted DVHs (dashed red curve with maximum dose 50 Gy), Tomotherapy model predicted DVHs (solid red curve with maximum dose 55 Gy) and actual clinical plan DVHs (solid blue curve with maximum dose 59 Gy) of a sample Tomotherapy patient. The model based on FG-IMRT plans was found to predict the median dose of the parotid of Tomo-IMRT plans very well, with a mean error of 2.6%, close to the mean error of 1.5% for intratomotherapy prediction. The maximum deviation ranged from −13.8% to 10.7%. Compared with the clinical values, the median dose predicted for the Tomotherapy cases by the FG-IMRT models for 34 out of 37 parotids were within 10% error bound of [Fig. 3(b)].The distribution of parotid median dose difference between: (1) IMRT model prediction and Tomotherapy plan value, (2) Tomotherapy model prediction and Tomotherapy plan value, and (3) IMRT model prediction and IMRT plan value are compared in Fig. 3(c) using box plots. All mean differences were small within about 3%. The p-value for a Kruskal Wallis test is 0.75, indicating the differences of the median values of the distributions were not significant. This suggests that the parotid model trained on a patient group treated with one IMRT technique can predict the median dose (and possibly the entire DVH) in another patient group treated with a different IMRT modality. We confirmed this finding with another critical structure, the spinal cord, which was spared similarly between two institutions. The mean difference of D2 (dose of 2% volume) of the spinal cord is −2.0% between the IMRT model prediction and Tomotherapy plan value, 1.3% between the Tomotherapy model prediction and Tomotherapy plan value, and −1.7% between the IMRT model prediction and IMRT plan value. The results of the spinal cord agree with those of the parotid.

3.D. Accuracy of model prediction across planning protocols

Besides the technical differences, in order to further study whether a model built at one institution can be used as a reference at another institution with different protocols, we applied the FG-IMRT model to predict the dose of the larynx and

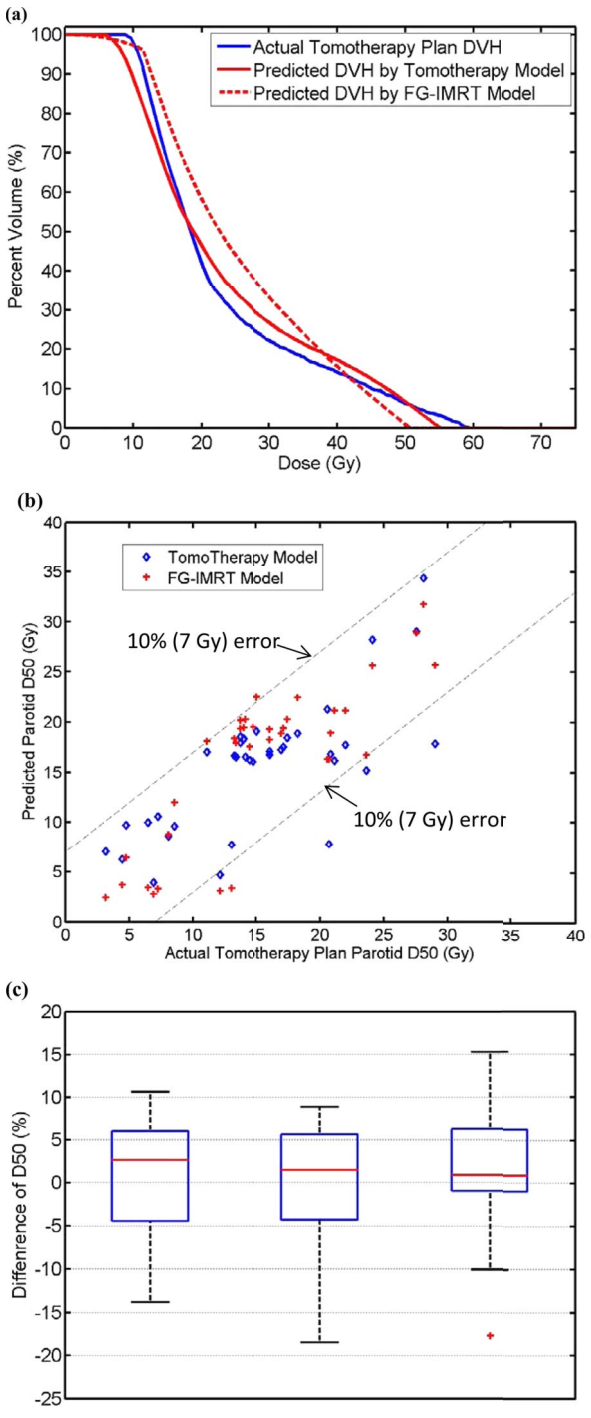

FIG. 3. Comparison of dosimetry of the parotid of actual Tomotherapy plan and predicted values by two models. (a) Comparison of DVH of an actual Tomotherapy plan (solid blue curve with maximum dose 59 Gy) and predicted DVHs by Tomotherapy model (solid red curve with maximum dose 55 Gy) and FG-IMRT model (dashed red curve with maximum dose 50 Gy). (b) Correlation of the parotid median dose of actual Tomotherapy plans in the study group vs Tomotherapy model and FG-IMRT model. (c) Deviations of the predicted parotid median dose by the FG-IMRT model and Tomotherapy model from the actual values of Tomotherapy or FG-IMRT plan. The hormonal red line inside box is the mean value. The blue box represents the interquartile range (IQR) from 25% quartile to 75% quartile. The red cross indicates the outlier which is defined as the points more than 1.5 times IQR away from the box edge. The black dashed lines connect the box with the most extreme values which are not outliers.

brainstem of Tomotherapy plans. During the Tomotherapy planning, the larynx mean dose (≤41 Gy) and V60 (≤24%) were two main objectives of dose restrictions. Since the larynx often had overlap with the primary PTV, the DVH points penalized were often chosen in the low dose region for maintaining target coverage. With FG-IMRT planning, the optimizer was in general pushed harder in sparing the larynx. Although there was no general template used for FG-IMRT, the expected median dose value based on experience was 15–30 Gy, which was much lower than the default objective in the Tomotherapy planning template. The median dose of model prediction and actual plan is compared in Fig. 4(a). When the respective model was used, there was a small difference between the predicted and actual plan value of the median dose (Tomo model 2.1%; FG-IMRT model 1.9%). However, when the FG-IMRT model was used to estimate the median dose of the larynx of Tomotherapy patients, the projected result was 10.5% lower than the actual value of the clinical plan. The FG-IMRT model suggests possible better sparing of the larynx of Tomotherapy plans, which is consistent with the stricter constraint used in the FG-IMRT optimization.

The brainstem was planned differently in two institutions as well. The major constraint used in Tomotherapy planning was to limit the maximum point dose to less than 54 Gy. FG-IMRT planning varies the goal according to the individual patient with the typical maximum dose set to 25–35 Gy. In addition, low dose volume was pushed to be as small as possible such as to spare the overall volume. Two dosimetric endpoints, D2 and median dose were selected to assess the difference of model prediction and actual value. Similar to the parotid and larynx reported previously, the model established on the same patient group in general predicts well the actual D2 [Fig. 4(b)] and median dose [Fig. 4(c)]. If the FG-IMRT model was used to predict the D2 of Tomotherapy patients, the difference (FG-IMRT Model – Tomotherapy Plan) was −20.4% [Fig. 4(b)]. The median dose predicted by the FG-IMRT model was 12.8% lower than the actual Tomotherapy plan value [Fig. 4(c)]. The difference found in the dosimetry index agrees largely with the different degree of strictness of constraint goals used in planning.

To further illustrate the possibility that models built at one institution can accurately portray protocol differences at another institution, we re-generated Tomotherapy plans for three patients with the maximum deviation of larynx dosimetry and three patients with the maximum deviation of brainstem dosimetry between the original Tomotherapy plan and FG-IMRT model prediction (Fig. 5). (The Tomo model prediction and actual value are similar.) The FG-IMRT model calculated DVHs were used as the objective to improve the dosimetry of the larynx and brainstem during replanning, while the target coverage and the sparing of other OARs were kept similarly. With the FG-IMRT model predicted value as the reference, we compared dosimetric endpoints of larynx median dose, brainstem D2 and brainstem median dose in the original and re-optimized Tomotherapy plans. From these extreme cases, replanning lowered the larynx median dose from on average 33.8 to 20.6 Gy, while the FG-IMRT model prediction is

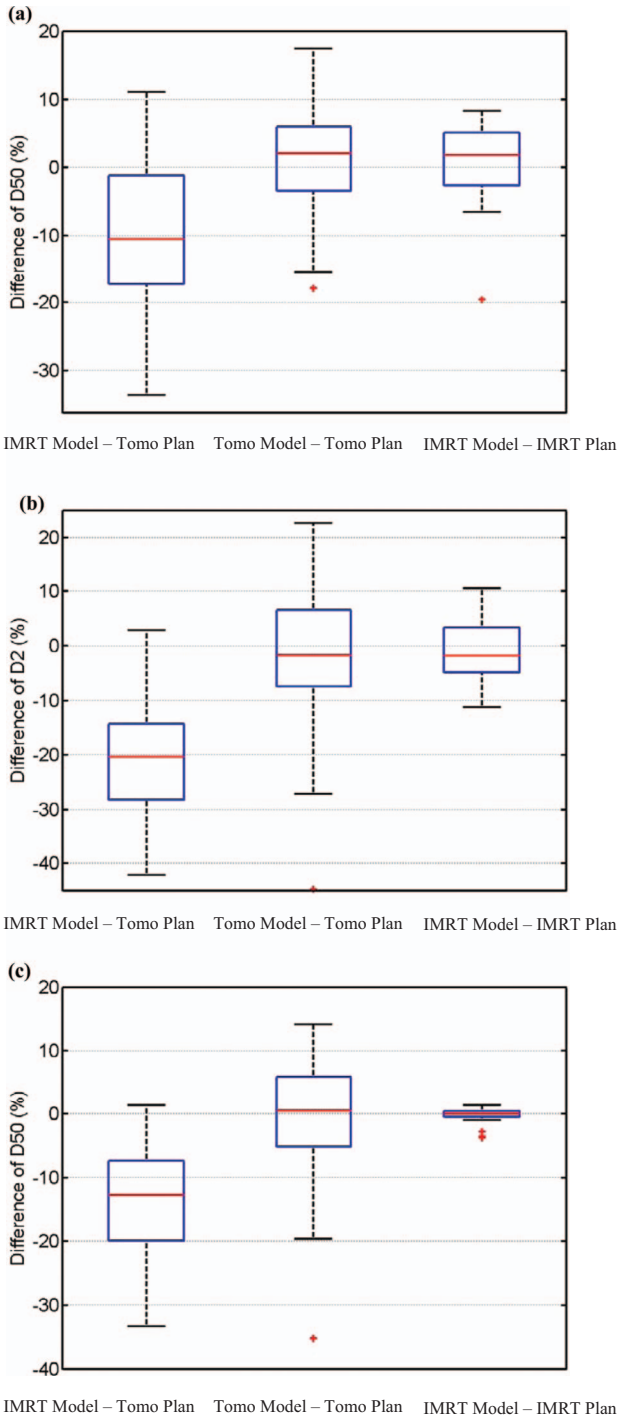

FIG. 4. Deviations of the predicted dosimetric endpoints by the FG-IMRT model and Tomotherapy model from the actual values of Tomotherapy or FG-IMRT plan. (a) larynx median dose (b) brainstem D2 and (c) brainstem median dose. The hormonal red line inside box is the mean value. The blue box represents the interquartile range (IQR) from 25% quartile to 75% quartile. The red crosses indicate the outlier which is defined as the points more than 1.5 times IQR away from the box edge. The black dashed lines connect the box with the most extreme values which are not outliers.

21.7 Gy on average. Similarly, brainstem D2 was reduced from 35.9 to 22.7 Gy with model prediction value 22.0 Gy. The brainstem median dose was also lowered from 23.5 to 10.7 Gy which is even lower than the model predicted value

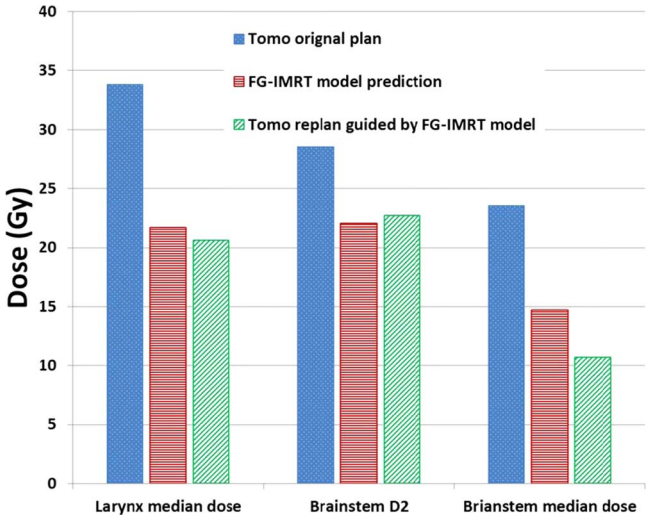

FIG. 5. Dosimetric endpoints of six Tomotherapy extreme cases replanned with the guide of the FG-IMRT model.

14.7 Gy. Please note in the original TomoTherapy planning, these dosimetric end points were in general not optimized. Replanning was found to have no noticeable effect on the dose of other OARs. The compromise of PTV coverage is very slight, from 96.9% to 96.3%. Although the FG-IMRT model was established based on its own patient group, planning protocol, and treatment technique, it proves useful in guiding the Tomotherapy planning of a completely different patient group.

4. DISCUSSION

From practice we know radiotherapy treatment planning involves two sets of information. One set is the anatomy including volume, shape and location of structures. The other set is the dosimetry and it is determined by multiple factors such as the anatomy, constraints and experience of the planner. Our modeling correlates these two sets of variables quantitatively. The experimental results of Tomotherapy and FG-IMRT cases showed the actual planed values could be predicted by its respective model. Therefore we believe two models are accurate and both describe the relations between the key features of anatomy and key properties of dosimetry of real clinical plans. Furthermore, the FG-IMRT model was able to estimate the dosimetry of the parotid planned in a similar way but treated with Tomotherapy. This implies a model built upon one IMRT technique may be generalized to some different techniques with similar theoretical underpinning. This will benefit the quality control of the implementation of a new IMRT technique because the model built on a traditional technique can provide reference dosimetry for the new technique as long as the constraint is similar. When the FG-IMRT model was applied to study the larynx and brainstem of Tomotherapy patients, model estimated the dosimetric values (such as median dose and D2) of clinical plans could be further reduced by 10%–20%. This is largely in agreement with the different constraints used in two types of treatment

planning. The replanning of six Tomotherapy cases with extreme deviations confirmed the prediction of the FG-IMRT model was useful in guiding the effort of further sparing the larynx and brainstem in spite of different patient groups and IMRT techniques used.

The application of this work may help the implementation of automated IMRT planning and qualitative quality control of plan dosimetry. The knowledge driven autoplanning approach allows the prediction of dosimetry of a new plan by an anatomy-dosimetry correlation model built upon previously executed treatment plans[5,9,10] The model presents the planner the anticipated DVH of a new case based on the previous experience prior to the treatment planning. This can be a very useful tool in helping the design of dosimetric optimizer parameters and tells planner whether it is an appropriate time to stop the iterations of optimization. However more effort is planned in our future work to make the optimization process automatic. It has been found from our clinical experience that the optimized result could be different for some cases when setting up the optimizer once or gradually adjusting the parameters, although the final parameters of the two methods are the same. Field width, pitch value, and modulation factor are another set of Tomotherapy planning parameters which may be modeled from a library of previous experience of planning results in our future work. We also want to stress that the use of model prediction does not necessarily generate better dosimetry than using a general dosimetry goal sheet. We can imagine when a highly experienced planner has a great amount of extra time to improve a plan even after all constraints of the goal sheet have been met, this plan may exceed the average plans made within the time limit of normal clinical workflow. The typical clinical plans were used in our modeling work such that model prediction reflects the realistic prior clinical experience and suggests best practice goals based on available evidence.

Similar to the IMRT on conventional Linac, current QA of Tomotherapy usually includes two components: (1) chart check to verify the agreement between physician directive and plan reported values and (2) phantom irradiation to compare the calculated and delivered dose. The checking of the quality of dosimetry itself always counts on the experience of an individual and often turns out to be inconclusive. For patients with complicated anatomy, it is almost infeasible for a human being to estimate the optimal dosimetry. The model we built for Tomotherapy cannot only improve planning efficiency, but can also serve as a quality control tool for dosimetric evaluation of clinical plans. This will help the establishment of a more comprehensive, quantitative and objective QA procedure for complex Tomotherapy planning.

There have been numerous publications to compare the dosimetry of Tomotherapy with other IMRT techniques.[20–23] The aim of this paper is not to compare the dosimetry of Tomotherapy and FG-IMRT. However the models we built for these modalities from a library of high quality plans can in fact provide more detailed dosimetric information like DVH of each structure. Therefore, models can be helpful in choosing an appropriate treatment modality for certain disease sites. The current clinical trial protocol of RTOG or other organizations normally simplifies the dosimetric requirement with a few dosimetric parameters. While meeting all of these loosely sampled DVH points, the overall shape of DVH may be very different among plans. Therefore, the accurate control of dosimetry quality of patients under the protocol can be challenging and even impossible. With the method demonstrated in this work, our model can produce the anticipated DVHs of an individual patient case. In this way, the institutions that participate in the trial can have a clear and comprehensive dosimetric goal to follow. Consequently the treatment plans generated by different medical centers may have more uniform dosimetric quality, which may lead to more robust data analysis of clinical trials.

5. CONCLUSIONS

We established a mathematical model to correlate the anatomical features and dosimetric indexes of OARs of HN patients in Tomo-IMRT and FG-IMRT plans. The model may be used for the setup of patient-specific OAR dose sparing goals and quality control of planning results. The institutional clinical experience was incorporated into the model which allows the model from one institution to generate a reference plan for another institution or another IMRT technique.

ACKNOWLEDGMENTS

L.Y., Y.G., F.Y., and J.W. are partially supported by NIH/NCI under Grant No. R21CA161389 and a master research grant from Varian Medical System. The authors appreciate Mrs. Rebecca Green for carefully reviewing and editing the paper and planning groups of two institutions generating high quality treatment plans.

[a]Electronic mail: jun_lian@med.unc.edu
[b]Electronic mail: lulin.yuan@duke.edu
[c]Author to whom correspondence should be addressed. Electronic mail: jackie.wu@duke.edu

[1]S. Webb, "The physical basis of IMRT and inverse planning," Br. J. Radiol. **76**, 678–689 (2003).
[2]Q. Wu and R. Mohan, "Algorithms and functionality of an intensity modulated radiotherapy optimization system," Med. Phys. **27**, 701–711 (2000).
[3]T. R. Mackie, T. W. Holmes, P. J. Reckwerdt, and J. Yang, "Tomotherapy: Optimized planning and delivery of radiation therapy," Int. J. Imaging Syst. Technol. **6**, 43–55 (1995).
[4]X. Zhang, X. Li, E. M. Quan, X. Pan, and Y. Li, "A methodology for automatic intensity-modulated radiation treatment planning for lung cancer," Phys. Med. Biol. **56**, 3873–3893 (2011).
[5]B. Wu, F. Ricchetti, G. Sanguineti, M. Kazhdan, P. Simari, M. Chuang, R. Taylor, R. Jacques, and T. McNutt, "Patient geometry-driven information retrieval for IMRT treatment plan quality control," Med. Phys. **36**, 5497–5505 (2009).
[6]B. Mathayomchan, Multiobjective Approach to Morphological Based Radiation Treatment Planning (PhD Thesis), Case Western Reserve University, Cleveland, OH, 2006.
[7]M. A. Hunt, A. Jackson, A. Narayana, and N. Lee, "Geometric factors influencing dosimetric sparing of the parotid glands using IMRT," Int. J. Radiat. Oncol., Biol., Phys. **66**, 296–304 (2006).
[8]N. Reddy, D. Nori, H. Chang, C. S. Lange, and A. Ravi, "Prostate and seminal vesicle volume based consideration of prostate cancer patients for treatment with 3D-conformal or intensity-modulated radiation therapy," Med. Phys. **37**, 3791–3801 (2010).

[9]K. L. Moore, R. S. Brame, D. A. Low, and S. Mutic, "Experience-based quality control of clinical intensity-modulated radiotherapy planning," Int. J. Radiat. Oncol., Biol., Phys. **81**, 545–551 (2011).
[10]X. Zhu, Y. Ge, T. Li, D. Thongphiew, F. F. Yin, and Q. J. Wu, "A planning quality evaluation tool for prostate adaptive IMRT based on machine learning," Med. Phys. **38**, 719–726 (2011).
[11]B. Wu, F. Ricchetti, G. Sanguineti, M. Kazhdan, P. Simari, R. Jacques, R. Taylor, and T. McNutt, "Data-driven approach to generating achievable dose-volume histogram objectives in intensity-modulated radiotherapy planning," Int. J. Radiat. Oncol., Biol., Phys. **79**, 1241–1247 (2011).
[12]L. Yuan, Y. Ge, W. R. Lee, F. F. Yin, J. P. Kirkpatrick, and Q. J. Wu, "Quantitative analysis of the factors which affect the interpatient organ-at-risk dose sparing variation in IMRT plans," Med. Phys. **39**, 6868–6878 (2012).
[13]L. M. Appenzoller, J. M. Michalski, W. L. Thorstad, S. Mutic, and K. L. Moore, "Predicting dose-volume histograms for organs-at-risk in IMRT planning," Med. Phys. **39**, 7446–7461 (2012).
[14]T. Bortfeld and S. Webb, "Single-Arc IMRT?," Phys. Med. Biol. **54**, N9–N20 (2009).
[15]H. T. Chung, B. Lee, E. Park, J. J. Lu, and P. Xia, "Can all centers plan intensity-modulated radiotherapy (IMRT) effectively? An external audit of dosimetric comparisons between three-dimensional conformal radiotherapy and IMRT for adjuvant chemoradiation for gastric cancer," Int. J. Radiat. Oncol., Biol., Phys. **71**, 1167–1174 (2008).
[16]I. J. Das, C. W. Cheng, K. L. Chopra, R. K. Mitra, S. P. Srivastava, and E. Glatstein, "Intensity-modulated radiation therapy dose prescription, recording, and delivery: patterns of variability among institutions and treatment planning systems," J. Natl. Cancer Inst. **100**, 300–307 (2008).
[17]B. E. Nelms, W. A. Tome, G. Robinson, and J. Wheeler, "Variations in the contouring of organs at risk: Test case from a patient with oropharyngeal cancer," Int. J. Radiat. Oncol., Biol., Phys. **82**, 368–378 (2012).
[18]J. O. Deasy, V. Moiseenko, L. Marks, K. S. C. Chao, J. Nam, and A. Eisbruch, "Radiotherapy dose–volume effects on salivary gland function," Int. J. Radiat. Oncol., Biol., Phys. **76**, S58–S63 (2010).
[19]M. Sohn, M. Alber, and D. Yan, "Principal component analysis-based pattern analysis of dose-volume histograms and influence on rectal toxicity," Int. J. Radiat. Oncol., Biol., Phys. **69**, 230–239 (2007).
[20]C. Fiorino, I. Dell'Oca, A. Pierelli, S. Broggi, E. De Martin, N. Di Muzio, B. Longobardi, F. Fazio, and R. Calandrino, "Significant improvement in normal tissue sparing and target coverage for head and neck cancer by means of helical tomotherapy," Radiother. Oncol. **78**, 276–282 (2006).
[21]P. Mavroidis, C. Shi, G. A. Plataniotis, M. G. Delichas, B. C. Ferreira, S. Rodriguez, B. K. Lind, and N. Papanikolaou, "Comparison of the helical tomotherapy against the multileaf collimator-based intensity-modulated radiotherapy and 3D conformal radiation modalities in lung cancer radiotherapy," Br. J. Radiol. **84**, 161–172 (2011).
[22]J. Pardo-Montero and J. D. Fenwick, "Tomotherapy-like versus VMAT-like treatments: A multicriteria comparison for a prostate geometry," Med. Phys. **39**, 7418–7429 (2012).
[23]R. Taylor, K. Opfermann, B. D. Jones, L. E. Terwilliger, D. G. McDonald, M. S. Ashenafi, E. Garrett-Meyer, and D. T. Marshall, "Comparison of radiation treatment delivery for pancreatic cancer: Linac intensity-modulated radiotherapy versus helical tomotherapy," J. Med. Imaging Radiat. Oncol. **56**, 332–337 (2012).