doi:10.1016/j.ijrobp.2010.05.026

PHYSICS CONTRIBUTION

DATA-DRIVEN APPROACH TO GENERATING ACHIEVABLE DOSE–VOLUME HISTOGRAM OBJECTIVES IN INTENSITY-MODULATED RADIOTHERAPY PLANNING

BINBIN WU, PH.D.,* FRANCESCO RICCHETTI, M.D.,* GIUSEPPE SANGUINETI, M.D.,*
MICHAEL KAZHDAN, PH.D.,† PATRICIO SIMARI, PH.D.,† ROBERT JACQUES, B.S.,‡ RUSSELL TAYLOR, PH.D.,†
AND TODD MCNUTT, PH.D.*

Departments of *Radiation Oncology and Molecular Radiation Science, †Computer Science, and ‡Biomedical Engineering, Johns
Hopkins University, Baltimore, MD

Purpose: To propose a method of intensity-modulated radiotherapy (IMRT) planning that generates achievable dose–volume histogram (DVH) objectives using a database containing geometric and dosimetric information of previous patients.

Methods and Materials: The overlap volume histogram (OVH) is used to compare the spatial relationships between the organs at risk and targets of a new patient with those of previous patients in a database. From the OVH analysis, the DVH objectives of the new patient were generated from the database and used as the initial planning goals. In a retrospective OVH-assisted planning demonstration, 15 patients were randomly selected from a database containing clinical plans (CPs) of 91 previous head-and-neck patients treated by a three-level IMRT-simultaneous integrated boost technique. OVH-assisted plans (OPs) were planned in a leave-one-out manner by a planner who had no knowledge of CPs. Thus, DVH objectives of an OP were generated from a subdatabase containing the information of the other 90 patients. Those DVH objectives were then used as the initial planning goals in IMRT optimization. Planning efficiency was evaluated by the number of clicks of the "Start Optimization" button in the course of planning. Although the Pinnacle^3 treatment planning system allows planners to interactively adjust the DVH parameters during optimization, planners in our institution have never used this function in planning.

Results: The average clicks required for completing the CP and OP was 27.6 and 1.9, respectively ($p <.00001$); three OPs were finished within a single click. Ten more patient's cord + 4 mm reached the sparing goal $D_{0.1cc} <44$ Gy ($p <.0001$), where $D_{0.1cc}$ represents the dose corresponding to 0.1 cc. For planning target volume uniformity, conformity, and other organ at risk sparing, the OPs were at least comparable with the CPs. Additionally, the averages of $D_{0.1cc}$ to the cord + 4 mm decreased by 6.9 Gy ($p <.0001$); averages of $D_{0.1cc}$ to the brainstem decreased by 7.7 Gy ($p <.005$). The averages of $V(30$ Gy$)$ to the contralateral parotid decreased by 8.7% ($p <.0001$), where $V(30$ Gy$)$ represents the percentage volume corresponding to 30 Gy.

Conclusion: The method heralds the possibility of automated IMRT planning.   © 2010 Elsevier Inc.

Intensity-modulated radiotherapy, IMRT, overlap volume histogram, OVH, head-and-neck, database.

INTRODUCTION

Intensity-modulated radiotherapy (IMRT) is an inverse treatment planning process that optimizes the intensity distribution for each of a set of beams according to the dose–volume histogram (DVH) objectives chosen by planners (1). The DVH objectives guide the planning software in scoring the tradeoffs between target coverage and organ at risk (OAR) sparing. However, the DVH objectives that account for the tradeoffs for a specific patient are often unknown before planning. Currently, an IMRT plan tailored to a specific patient requires many rounds of optimization. The planner usually applies population-based DVH objectives at the beginning (2). The planner then progressively improves the plan until it becomes clinically acceptable by repeatedly adjusting the DVH objectives in each optimization round according to personal experience and clinical feedbacks. It

has been shown that IMRT plan quality varies among planners by level of experience (3). Even for the same patient, the planning results are often highly heterogeneous (4). Consequently, IMRT plan quality heavily relies on the personal experience of planners and the time they can allocate to a specific plan.

In general, the dose distributions of the OARs and targets depend on the beam characteristics (energy, modality, number of beams, beam angle, and delivery technique), prescribed dosage, and anatomic configurations of the targets and OARs. Given the consistent beam characteristics and prescribed dosages across patients, the variability of the anatomic structures between the patients, specifically, the spatial configuration between each target and OAR, dictates the achievability of a given set of DVH objectives.

On the basis of that observation, we proposed a method of IMRT planning that generates the achievable DVH objectives that account for the tradeoffs between target coverage and OAR sparing for a new patient from a database containing geometric and dosimetric information of previous patients. Before starting a new plan, the planners will search through the database and identify a group of reference patients by comparing the spatial configurations between the OARs and targets of the new patient with those of previous patients. The lowest clinically achievable OAR DVHs were retrieved from the reference group and applied as initial planning goals for the new patient's OARs.

We conducted a retrospective planning study to verify the effectiveness of our method. We applied our method to a group of IMRT head-and-neck patients with consistent beam characteristics and prescribed dosage. The purpose of the present study was to assess whether *a priori* knowledge of patient geometry-specific DVH objectives would improve both the efficiency and the consistency of IMRT planning.

METHODS AND MATERIALS

In our previous work (5,6), we introduced the concept of a shape relationship descriptor, the overlap volume histogram (OVH), to quantify the spatial configuration between an OAR and a target. In the present report, the OVH was used to compare the spatial configurations of the OARs and targets of a new patient with those of previous patients in a database. This database contained the OVH and DVH information of the patients. From the OVH analysis, the DVH objectives of the OARs for the new patient were generated from the database and applied as the initial planning goals for IMRT optimization. A brief review of the OVH and its relationship with the DVH is presented in the following section.

Definition of the OVH and its relationship to DVH

The OVH is a one-dimensional distribution associated with each OAR, measuring its Euclidean distance from the target. It describes the percentage of fractional volume, v, of an OAR that is within a specified distance, r, of a target: $v = OVH(r)$. Specifically, the value of the OVH represents the percentage of the OAR's volume that overlaps with an isotropically expanded or contracted target. In the following discussion, we use the notation r_v to represent the expansion or contraction distance that the target needs to cover a certain percentage volume v of the OAR: $r_v = OVH^{-1}(v)$.

As indicated in the reference report (5), the smaller the OVH value, the closer the OAR is to the target; the closer the OAR is to the target, the greater the dose to the OAR will more likely be. For example, for two OARs: OAR_1 and OAR_2, if $r_{v,1} \geq r_{v,2}$ for a certain percentage volume v, OAR_2 will be closer to the target at that v. We expected $D_{v,1} \leq D_{v,2}$, where D_v represents the dose of the organ at the percentage volume v. Intuitively, this was because the v percentage of the organ's volume closest to the target will largely determine the dose of the organ at that v, D_v.

Approach for generating DVH objectives

Our approach has been to use the information of previously treated patients who were determined to be at least as difficult for planning to generate the DVH objectives for the OARs of a new patient. A database containing the OVHs and DVHs of previous patients is a prerequisite for this approach. The beam characteristics and prescribed dosage to the previous patients were the same as those of the new patient.

Once the database of previous patients was created, we used it to generate the DVH objectives at a specific percentage volume v, $D_{v,n}$, for a selected OAR of a new patient n. We sought a group of previous patients with the OARs' OVH values smaller than those of the new patients and apply their minimal DVH values as the new patient's DVH objectives. Specifically, to generate $D_{v,n}$, the OVH of that selected OAR at v, $r_{v,n}$ was used to query the database to find a group of previous patients, i, whose OARs' OVH values at v, $r_{v,i}$ were smaller than those of $r_{v,n}$. Next, the minimum of D_v among the group of previous patients was chosen as the initial planning goal for $D_{v,n}$:

$$D_{v,n} = \min\{D_{v,i} | r_{v,n} \geq r_{v,i} \text{ and } V_{95,i} \geq a\%\} \qquad (1)$$

Condition, $V_{95,i} \geq a\%$, serves to confine the search results to the previous plans with a good planning target volume (PTV) coverage, where V_x represents the percentage of the PTV receiving $x\%$ of the prescription dose. The typical a value was 99 ($V_{95} \geq 99\%$ for PTV coverage).

The v value selection in Eq. 1 is OAR specific. For example, the dosimetric guideline in the *Radiation Therapy Oncology Group* 00-22 (7) for brainstem sparing was $D_0 < 54$ Gy (D_0 was the maximal dose); thus, $v = 0\%$ is chosen for it. Additionally, multiple DVH objectives can be chosen for some OARs. The Radiation Therapy Oncology Group 00-22 guideline for parotid sparing was mean dose < 26 Gy or $V(30$ Gy$) < 50\%$; thus, we set the values of v at 30%, 50%, and 70% for the parotid, where $V(x$ Gy$)$ represents the percentage OAR's volume corresponding to x Gy.

Experimental demonstration: a head-and-neck retrospective planning study

A total of 15 patients were randomly selected from an anonymized database of 91 previous head-and-neck patients for an OVH-assisted planning demonstration. For each OVH-assisted planning, DVH objectives of 13 selected OARs (brain, brainstem, cord + 4 mm, mandible, oral mucosa, left parotid, right parotid, left inner ear, right inner ear, larynx, esophagus, left brachial plexus, and right brachial plexus) were generated by Eq. 1 from the database using a leave-one-out method and directly applied as the initial planning goals in IMRT optimization.

To evaluate the effectiveness of our method, the results for three sets of plans: clinical plans (CPs), OVH-assisted plans after the first-round optimization (OP1s), and final OVH-assisted plans (OP2s) were statistically compared using the DVH.

Clinical plans

At the study (August 2009), our database included the CPs of 91 consecutive patients with head-and-neck squamous cell carcinomas treated at Johns Hopkins from June 2007. The disease subsites included oropharynx ($n = 69$, 76%), larynx ($n = 11$, 12%), nasopharynx ($n = 8$, 9%), hypopharynx ($n = 3$, 3%). All patients underwent IMRT with nine fixed co-planar 6-MV photon beams and underwent a three-level simultaneous integrated boost technique with the following treatment scheme: 70 Gy to macroscopic disease (clinical target volume [CTV]); 63 Gy to microscopic high-risk disease; 58.1 Gy to the microscopic low risk disease. The dose prescription was set to 70 Gy at 2 Gy/fraction. A 5-mm expansion margin was applied to the CTVs to obtain the corresponding PTVs, and the expansion was limited to within 4 mm from the skin surface.

The CPs were planned by 5 senior dosimetrists using the Pinnacle[3] treatment planning system (TPS, Philips Radiation Oncology Systems, Madison, WI). In the Direct Machine Parameter Optimization, the maximal number of segments was set at 120. All contours (CTVs and OARs) were consistently placed by a single observer (G.S.). Table 1 lists the in-house dose–volume points used at Johns Hopkins for head-and-neck plan evaluation. The cord + 4 mm (cord with 4-mm expansion), mandible, brainstem, and brain were the primary OARs considered in planning. The parotid, oral mucosa, inner ear, brachial plexus, esophagus, and larynx were the secondary OARs. Dosimetrists were not required to start from this table at the beginning of planning.

OVH-assisted plans

A total of 15 patients were randomly selected from the previously described database for an OVH-assisted planning demonstration after excluding those with nasopharyngeal primary tumors. The tumor characteristics for the selected patients are listed in Table 2. The OPs were generated by a single planner who did not contribute to any of the CPs, according to the following criteria:

The contours for the PTVs and OARs were unchanged from the CPs.

The 9 fixed co-planar 6-MV photon beams and treatment machine used in the CP were applied to the corresponding OP; the planner had no knowledge of the CPs beyond the above information.

Table 1. In-house dosimetric guidelines used at Johns Hopkins for head-and-neck plan evaluation

Target	Endpoint	Goal	Minor
PTV[58.1]	V_{95}	99%	95%
PTV[63]	V_{95}	99%	95%
PTV[70]	V_{95}	99%	95%
OAR			
Cord + 4 mm	$D_{0.1cc}$	44 Gy	46.2 Gy
Mandible	$D_{0.1cc}$	73.5 Gy	77 Gy
Brainstem	$D_{0.1cc}$	54 Gy	60 Gy
Brain	D_{1cc}	60 Gy	63 Gy
Brachial plexus	$D_{0.1cc}$	60 Gy	66 Gy
Esophagus	D_{1cc}	45 Gy	55 Gy
Parotid*	$V(30\ Gy)$	50%	60%
Larynx	$V(50\ Gy)$	25%	30%
Inner ear	D_{mean}	50 Gy	52.5 Gy
Oral mucosa	$V_{cc}(66.5\ Gy)$	64 cm^3	70 cm^3

Abbreviations: V_x = percentage volume receiving x% of prescription dose; $D_{x\ cc}$ = dose corresponding to x cm^3; $V(x\ Gy)$ = percentage volume corresponding to x Gy; D_{mean} = mean dose; $V_{cc}(x\ Gy)$ = volume corresponding to x Gy; cc = cm^3

* At least one parotid per patient.

Table 2. Tumor characteristics for the 15 randomly selected head-and-neck patients.

Pt. No.	Subsite	PTV[70] volume (cm^3)	PTV[63] volume (cm^3)	PTV[58.1] volume (cm^3)
1	Oropharynx	230	471	932
2	Larynx	285	486	861
3	Oropharynx	364	644	1,160
4	Oropharynx	112	282	943
5	Oropharynx	95	267	923
6	Oropharynx	43	192	755
7	Oropharynx	154	407	1,244
8	Oropharynx	19	145	730
9	Oropharynx	117	293	1,075
10	Oropharynx	65	224	693
11	Oropharynx	89	203	926
12	Oropharynx	141	277	856
13	Oropharynx	64	336	949
14	Hypopharynx	365	743	1,400
15	Oropharynx	45	166	548
Mean		145.8	342.4	933

Abbreviations: Pt. No. = patient number.

Direct Machine Parameter Optimization was applied to generate OPs. The maximal number of segments was set at 120 that were the same in CPs.

A leave-one-out method was applied to generate the DVH objectives for the 13 selected OARs. For each OP, a subdatabase containing the information of the other 90 patients was constructed. Each OAR's DVH objectives were then generated from that subdatabase.

In the first round of optimization, the planner would directly apply the database-generated DVH objectives to the Pinnacle[3] TPS as the initial planning goals (OP1). The final plan (OP2) could be achieved through more than one round of optimization if necessary. Before each round of optimization after the first, the planner could adjust the DVH parameters (objectives and weights) according to their personal judgment but was not allowed to make such adjustments during the process of any optimization. The planner would finish the plan when satisfied with the results. No physician interaction occurred during OP planning.

Criteria for plan comparison

The three sets of plans (CP, OP1, and OP2) were compared in terms of target coverage, homogeneity, conformity, OAR sparing, and efficiency. The Wilcoxon rank-sum p test and chi-square p test were used for the statistical comparisons, as appropriate. The data were considered statistically significant at $p < .05$.

For the comparison of the target coverage and OAR sparing, we followed the in-house dosimetric guidelines summarized in Table 1. In addition to V_{95}, the PTV coverage was evaluated by V_{98} and V_{100}. The homogeneity of the PTV was evaluated by $D_5–D_{95}$, where D_x represented the dose corresponding to x% of volume. The conformity of the PTV was evaluated by the conformity index (CI), defined as the volume of the isosurface of the prescription dose divided by the volume of the corresponding PTV. For example, the CI for the PTV[58.1], CI[58.1], was defined as Vol(58.1 Gy isosurface)/ Vol(PTV[58.1]).

The planning efficiency was evaluated by the number of clicks of the "Start Optimization" button in the course of planning. For comparison, the number was extracted from the transcript files of the Pinnacle[3] TPS. An optimization round where referenced represented the event of a single click. One click triggers one optimization round. Although the Pinnacle[3] TPS allows planners to interactively

adjust the DVH parameters during optimization, planners in our institution never used this function in planning.

RESULTS

Planning efficiency comparison

Figure 1 illustrates the distributions of the number of optimization rounds required for a complete CP and its corresponding OP. The average number of optimization rounds per OP was 1.9 (SD 0.6); that number for the CP was 27.6 (SD 10.4; $p <.00001$). Three OPs were completed in a single optimization round.

Plan comparison: PTV coverage, homogeneity and conformity

Table 3 lists the number of individual plans satisfying the goal and minor criteria for PTV coverage (Table 1). All plans met the minor criteria. We could not detect any difference between the OP1s and OP2s. Additionally, the results were very similar between the CPs and OPs.

The summary of averages of the dosimetric results for the PTVs of the three sets of plans are listed in Table 4. In both OPs, averages of V_{95}, V_{98}, and V_{100} were desirably greater and average of D_5–D_{95} was desirably lower; the averages of CI were comparable between the CPs and OPs. No statistically significant differences were observed.

Plan comparison: OAR sparing

The number of individual plans satisfying the goal and minor criteria for the various selected OARs is listed in Table 5. The contributions of the OP2s to the OP1s can be summarized as follows: one more cord + 4 mm, one more brainstem, and one more patient's parotid in the OP2s reached the goal of sparing; for the minor, we could not detect any difference.

Table 5 also lists that both OPs were at least as effective as the CPs in meeting the goals for all OARs, except for one brain, two ipsilateral brachial plexuses, and two contralateral brachial plexuses. However, for those OARs, the differences were not statistically significant. Taking into account the mi-
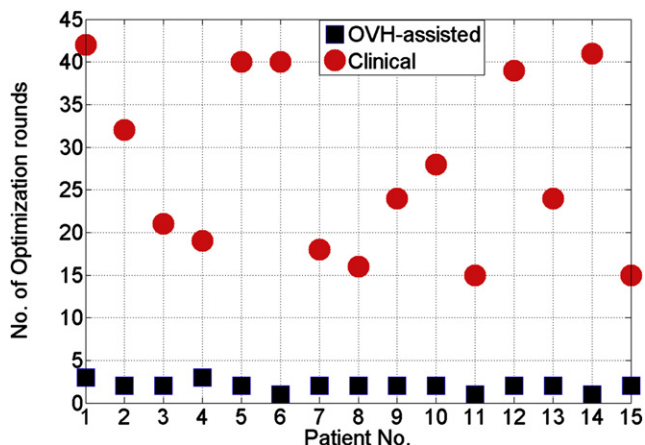
nor violations, all OARs in the OPs were not inferior to the CPs, except for one contralateral brachial plexus. In contrast, we observed a statistically significant advantage for both OPs compared with the CPs in terms of cord sparing ($p <.0001$): 9 and 10 more cord + 4 mm reached the goal in the OP1s and OP2s, respectively. However, it should be noted that $D_{0.1cc}$ to the cord + 4 mm for the 4 CPs that did not meet the minor was always <50 Gy, and the maximal dose to the cord itself (without expansion) was always <45 Gy.

A summary of averages of the dosimetric results for the OARs of the three sets of plans are listed in Table 6. For the cord + 4 mm, the averages of $D_{0.1cc}$ to the OP1s and OP2s decreased by 6.1 Gy and 6.9 Gy ($p <.0001$) respectively. For the brainstem, the averages of $D_{0.1cc}$ to the OP1s and OP2s decreased by 7.3 Gy and 7.7 Gy, respectively ($p <.005$). For the contralateral parotid, $V(30\ Gy)$ to the OP1s and OP2s decreased by 7% and 8.7% ($p <.0001$),

Table 3. Number of individual plans satisfying goal and minor criteria for PTV coverage

Target	Endpoint	CP	OP1	OP2
PTV[58.1]	Goal $V_{95} \geq 99\%$	8	8	8
	Minor $V_{95} \geq 95\%$	15	15	15
PTV[63]	Goal $V_{95} \geq 99\%$	14	15	15
	Minor $V_{95} \geq 95\%$	15	15	15
PTV[70]	Goal $V_{95} \geq 99\%$	15	15	15
	Minor $V_{95} \geq 95\%$	15	15	15

Abbreviations: CP = clinical plan; OP1 = first-round OVH-assisted plan; OP2 = final OVH-assisted plan; other abbreviations as in Table 1.
No statistically significant differences were observed.

Table 4. Summary of the dosimetric results for the PTVs in the three sets of plans.

Variable	CP Average	OP1 Average	OP2 Average	CP vs OP1	CP vs OP2	OP1 vs OP2
PTV[58.1]						
V_{100} (%)	94.1	94.3	94.5	.56	.23	.85
V_{98} (%)	97.1	97.9	98	.3	.24	.6
V_{95} (%)	98.9	99	99	.8	.71	.6
D_5–D_{95} (Gy)	16	13.9	13.7	.2	.24	.85
CI[58.1]	1.2	1.2	1.2	.55	.76	.95
PTV[63]						
V_{100} (%)	98.7	99.1	99	.08	.15	.9
V_{98} (%)	99.2	99.6	99.6	.12	.23	.55
V_{95} (%)	99.7	99.8	99.9	.34	.77	.43
D_5–D_{95} (Gy)	9	8	8.1	.1	.28	.67
CI[63]	1.3	1.3	1.3	.6	.45	.65
PTV[70]						
V_{100}(%)	95.1	95.4	95.3	.5	.32	.9
V_{98}(%)	98.6	98.8	99	.4	.21	.9
V_{95}(%)	99.8	99.9	99.9	.3	.2	.93
D_5–D_{95} (Gy)	3.7	3	3.2	.6	.97	.7
CI[70]	1.2	1.3	1.3	.6	.42	.88

Abbreviations as in Tables 1 and 3.
No statistically significant differences were observed.

Fig. 1. Distribution of number of optimization rounds required for complete plan. OVH = overlap volume histogram.

Table 5. Number of individual plans satisfying goal and minor criteria for OAR sparing

OAR	Endpoint	CP	OP1	OP2
Cord + 4 mm*	Goal $D_{0.1\,cc} \leq 44$ Gy	5	14	15
	Minor $D_{0.1\,cc} \leq 46.2$ Gy	11	15	15
Brainstem	Goal $D_{0.1\,cc} \leq 54$ Gy	13	14	15
	Minor $D_{0.1\,cc} \leq 60$ Gy	15	15	15
Brain	Goal $D_{0.1\,cc} \leq 60$ Gy	15	14	14
	Minor $D_{0.1\,cc} \leq 63$ Gy	15	15	15
Mandible	Goal $D_{0.1\,cc} \leq 73.5$ Gy	15	15	15
	Minor $D_{0.1\,cc} \leq 77$ Gy	15	15	15
Parotid†	Goal $V(30$ Gy$) \leq 50\%$	12	13	14
	Minor $V(30$ Gy$) \leq 60\%$	14	14	14
Larynx	Goal $V(50$ Gy$) \leq 25\%$	1	2	3
	Minor $V(50$ Gy$) \leq 30\%$	4	5	5
Ipsilateral brachial plexus	Goal $D_{0.1\,cc} \leq 60$ Gy	5	3	3
	Minor $D_{0.1\,cc} \leq 66$ Gy	9	10	10
Contralateral brachial plexus	Goal $D_{0.1\,cc} \leq 60$ Gy	11	9	9
	Minor $D_{0.1\,cc} \leq 66$ Gy	15	14	14
Esophagus	Goal $D_{1\,cc} \leq 45$ Gy	2	2	2
	Minor $D_{1\,cc} \leq 55$ Gy	5	5	5
Ipsilateral inner ear	Goal $D_{mean} \leq 50$ Gy	15	15	15
	Minor $D_{mean} \leq 52.5$ Gy	15	15	15
Contralateral inner ear	Goal $D_{mean} \leq 50$ Gy	15	15	15
	Minor $D_{mean} \leq 52.5$ Gy	15	15	15
Oral mucosa	Goal V_{cc} (66.5 Gy) ≤ 64 cm^3	12	12	12
	Minor V_{cc} (66.5 Gy) ≤ 70 cm^3	13	13	13

Abbreviations as in Tables 1 and 3.

* Statistical significance was observed for goal: $p_{cp,op1} < .0001$ and $p_{cp,op2} < .0001$.

† At least one parotid per patient.

respectively. For the ipsilateral parotid, although $V(30$ Gy$)$ to the OP1s and OP2s decreased by 8% and 6.5%, respectively, no statistically significant differences were observed. For the rest of the OARs, the OPs were at least comparable to the CPs, on average, and no statistically significant differences were observed. In addition, averages of the evaluation points between the OP1s and OP2s were similar, and we could not detect any statistically significant difference.

DISCUSSION

To demonstrate the effectiveness of our method, a retrospective planning study was performed to compare the dosimetric results among the CPs, OP1s, and OP2s of the 15 randomly selected head-and-neck patients. We found that neither OP was inferior to the CPs, and the OP1s and OP2s were remarkably similar. Before discussing the findings in detail, we would like to clarify some of the methodologic aspects.

Our method is dependent on the availability of a database of previous patients with the same prescribed dosage and beam characteristics. All contours and plans in our database were done under the supervision of a single physician (G.S.). How the method would perform for a set of patients from another institution or treated by multiple physicians has not yet been determined.

Our method is not intended to determine the optimal plan for a specific patient, in general, which depends on other physical and clinical factors. Rather, it aims to achieve the results of the best plan according to the knowledge of the treatment plans of previous patients. Specifically, our method aimed to find the lowest (favorable) clinically achievable OAR doses among those related previous patients identified by the OVH analysis and apply them as initial planning goals to new patient's OARs.

Because the knowledge of the previous plans is used to generate the DVH objectives, the performance of the method depends on the quality of the previous plans and the size of the database.

Our method is patient-geometry specific. Specifically, it is based on the spatial relationship between an OAR and a target characterized by the OVH. It implicitly assumes ''OAR-independence'': the dosimetric influence of other OARs on each other is assumed to be a second-order effect. From a practical point of view, we found that this OAR-independence assumption is applicable to the 13 OARs in head-and-neck planning; however, this might not be the case for OARs in

Table 6. Summary of dosimetric results for OARs in three sets of plans

OAR	Endpoint	CP Average	OP1 Average	OP2 Average	Wilcoxon p CP vs OP1	CP vs OP2	OP1 vs OP2
Cord + 4 mm	$D_{0.1cc}$	45.6	39.5	38.7	<.0001*	<.0001*	.7
Mandible	$D_{0.1cc}$	67.4	67.3	67.8	.79	1	.91
Brainstem	$D_{0.1cc}$	47.7	40.4	40	<.005*	<.005*	.85
Brain	D_{1cc}	50.8	50	49.6	.5	.38	.88
Ipsilateral parotid	$V(30$ Gy$)$	65	57	58.5	.21	.3	.8
Contralateral parotid	$V(30$ Gy$)$	52	45	43.3	<.0001*	<.0001*	.56
Larynx	$V(50$ Gy$)$	55.4	53.3	50.1	.66	.57	.91
Esophagus	D_{1cc}	53.9	54.1	54	1	.9	.95
Ipsilateral brachial plexus	$D_{0.1cc}$	62.2	62.7	62	.97	.93	.9
Contralateral brachial plexus	$D_{0.1cc}$	58.4	59.44	59.53	.79	.84	.86
Oral mucosa	$V_{cc}(66.5$ Gy$)$	37.6	39.5	40	.6	.74	.93
Ipsilateral inner ear	D_{mean}	31	25.7	26	.32	.47	1
Contralateral inner ear	D_{mean}	25	19.5	21	.2	.43	1

Abbreviations as in Tables 1 and 3.

* Statistically significant.

close proximity. Moreover, patients with cancer arising in the nasopharynx were excluded, because their planning involves other OARs (*i.e.,* the optic pathways) than those considered in the present planning exercise. Whether the method can be extended to other OARs and disease sites required additional investigation.

In our method, the CPs were considered as the reference plans for a given new patient. With the exception of a few OARs, the CPs achieved satisfactory (goal + minor variation) sparing for most of patients and target coverage for all patients. In the CPs, the greater than desirable doses to the larynx and esophagus were both justified by the location of some tumors in the lower part of the oropharynx or in the larynx/hypopharynx and the use of whole field IMRT (8). Similarly, for some patients, it was impossible to meet the dose criteria for the brachial plexus owing to its overlap/proximity to the tumors. We noted that the maximal dose to the cord + 4 mm was suboptimal in some CPs. Because the dose limit of 45 Gy to the nonexpanded cord was always met in the CPs, we believe this represents an oversight intrinsic to the nonscientific method that clinical plan evaluation was previously done at our institution. It is noteworthy that since we introduced a tabular method to summarize the individual dosimetric data in approximately October 2009, all the CPs met at least the minor criterion for the expanded cord.

In the present study, the dosimetric results (Tables 3–6) generated by OVH-assisted planning were not inferior to the CPs. That both OPs resulted in significantly lower doses to the cord + 4 mm, brainstem, and contralateral parotid was certainly remarkable, but it should not be overemphasized for the following reasons. First, the CPs were not generated intentionally to reach the lowest possible dose to the OARs but to meet the goal and minor criteria in Table 1. Second, and probably more importantly, additional OAR sparing could possibly be achieved at the cost of PTV coverage and homogeneity. Owing to the large size of the PTV[58.1], PTV underdosing might not be easily detectable by considering only the values of V_{95}, V_{98}, or V_{100}. However, it is still remarkable that the doses to those three OARs were significantly reduced without degrading the PTV or other OARs using the DVH. On average, the PTV coverage and homogeneity were slightly improved in both OPs.

The OVH can help to direct planners' efforts toward an "achievable" (because previously clinically achieved in similar patients), individualized (because patient geometry-specific instead of the standard DVH objectives) and often more favorable (because it continues to ameliorate by aiming at the lowest clinically achieved OAR doses in the database) DVH objectives than initially thought. However, whether the OVH-assisted planning advances the plan quality over the traditional planning would need prospective validation within a head-to-head comparison between simultaneously generated CPs and OPs where several other details of planning are controlled and equally stressed. We believe that physician validation and slice-by-slice review of the plan is still a fundamental requisite, not only for plan acceptance, but also for "fine tuning" of the plan.

With this in mind, the limited improvement observed using the dosimetric criteria for the OP2s compared with the OP1s further highlights the "goodness" of the OP1s as an efficient and reliable method to obtain a baseline plan for physician evaluation.

Our study used the number of optimization rounds required for a complete plan to evaluate planning efficiency. In terms of the present report, it is an objective measure for the efficiency of planning because planners in our institution never adjusted the DVH parameters during optimization. When adjustments were needed during optimization, the planners always first terminated the optimization process by clicking the "Stop Optimization" button, made adjustments, and then ran a new optimization round by clicking the "Start Optimization" button. An alternative measure of evaluation would be the time dedicated to planning. However, the planning time is not readily quantifiable, because a part of it is the machine computing time that can take place in the background while planners are working on other plans.

In our experience, the average number of optimization rounds required for a complete CP was 27.6, roughly 14 times more than that for an OP ($p < .00001$). The number of optimization rounds varies with the experience of the planners, complications of the disease sites, approach to planning, and quality of plans. One might argue that 27.6 is an excessive number for an experienced planner; however, using an "educated" (driven by experience) estimate on the DVH objectives, there is little doubt that the planner can potentially reduce number of optimization rounds needed to achieve a given plan. However, to the best of our knowledge, no data are available from published reports for this specific setting for comparison. Moreover, because of the retrospective nature of the present study, we could not provide information on the amount of improvements and reasons associated with each optimization round in the CPs. However, it remains a remarkable finding that, in an uncontrolled setting, the OPs were finalized in less than two optimization rounds, on average.

Our proposal generates IMRT plans by using dose–volume-based optimization in the Pinnacle[3] TPS. A different approach uses the generalized equivalent uniform dose (EUD) formalism (9,10). Because the OVH is used to predicate clinical achievable OAR's DVH and EUD is a function of DVH, it is possible to use the OVH to predicate clinical achievable EUD, which can be used in generalized EUD-based IMRT optimization. However, this potential application was beyond the scope of the present report.

The quality of the previous plans in the database will determine the performance of our method. It could be suggested that the quality of the database can be improved through iteration by running through each entry with the same leave-one-out strategy used in our retrospective study. If a new OP is better than the corresponding CP, this OP will replace the CP in the database; the process continues until no additional improvement can be made. However, all plans currently in our database were approved by a physician (G.S.) and actually delivered to the patients. The dosimetric results in these

plans reflected interaction of planners with that specific physician in previous planning sessions. By replacing the CP with the corresponding OP, the plans in the new database were exclusively determined by the anatomy structures of the patients. As a result, we would likely lose the physicians' input. The implication of excluding the physician's opinion in the database is unknown. A database of previous plans from other institutions (or other physicians) is needed to explore the issue.

The OPs were generated in a consistent method using an objective criterion: the DVH objectives of OARs were generated by Eq. 1. In contrast, the CPs were planned by 5 dosimetrists. As mentioned in the "Introduction," there is little doubt that current IMRT planning is associated with a high degree of heterogeneity among institutions and even within the same institutions (3,4). Our method thus provides a consistent and systematic method for generating DVH objectives in IMRT planning.

A final aspect worth noting is that our approach heralds the possibility of automated IMRT planning. Because the database-generated DVH objectives of OARs were patient geometry-specific and the tradeoff between target coverage and OAR sparing was considered in previous planning sessions, they are very likely achievable for the new patient. Thus, the weights used in IMRT planning would no longer be important. In planning the OPs, the weights of the OARs and targets were set to be the same.

CONCLUSION

In the present report, we introduced an efficient method of generating achievable DVH objectives that account for the tradeoffs between the target coverage and OAR sparing using geometric and dosimetric information retrieved from a database of previous plans. This offers a method of predicting clinically achievable doses ahead of planning. A head-and-neck retrospective planning study was performed to demonstrate the effectiveness of the method. The results illustrated that with the combined use of OVH and a database of previous treatment plans, a planner will be able to achieve the quality of the most favorable plan in the database in fewer than two optimization rounds. It makes IMRT planning no longer a trial-and-error process. The method heralds the possibility of automated IMRT planning.

REFERENCES

1. Wu Q, Mohan R. Algorithms and functionality of an intensity modulated radiotherapy optimization system. *Med Phys* 2000; 27:701–711.
2. Wu Q, Chi Y, Chen PY, *et al.* Adaptive replanning strategies accounting for shrinkage in head and neck IMRT. *Int J Radiat Oncol Biol Phys* 2009;75:924–932.
3. Chung HT, Lee B, Park E, *et al.* 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* 2008;71:1167–1174.
4. Williams MJ, Bailey M, Forstner D, *et al.* Multicentre quality assurance of intensity-modulated radiation therapy plans: A precursor to clinical trials. *Australas Radiol* 2007;51:472–479.
5. Wu B, Ricchetti F, Sanguineti G, *et al.* Patient geometry-driven information retrieval for IMRT treatment plan quality control. *Med Phys* 2009;36:5497–5505.
6. Kazhdan M, Simari P, McNutt T, *et al.* A shape relationship descriptor for radiation therapy planning. *Lecture Notes Computer Sci* 2009;5762:100–108.
7. Eisbruch A, Harris J, Garden AS, *et al.* Multi-institutional trial of accelerated hypofractionated intensity-modulated radiation therapy for early-stage oropharyngeal cancer (RTOG 00-22). *Int J Radiat Oncol Biol Phys* 2010;76:1333–1338.
8. Dabaja B, Salehpour MR, Rosen I, *et al.* Intensity-modulated radiation therapy (IMRT) of cancers of the head and neck: Comparison of split-field and whole-field techniques. *Int J Radiat Oncol Biol Phys* 2005;63:1000–1005.
9. Wu Q, Mohan R, Niemierko A, *et al.* Optimization of intensity-modulated radiotherapy plans based on the equivalent uniform dose. *Int J Radiat Oncol Biol Phys* 2002;52:224–235.
10. Thieke C, Bortfeld T, Niemierko A, *et al.* From physical dose constraints to equivalent uniform dose constraints in inverse radiotherapy planning. *Med Phys* 2003;30:2332–2339.