

A Statistical Approach for Achievable Dose Querying in IMRT Planning

Patricio Simari¹, Binbin Wu², Robert Jacques³, Alex King¹, Todd McNutt²,
Russell Taylor¹, and Michael Kazhdan¹

¹ Department of Computer Science, Johns Hopkins University, USA

² Department of Radiation Oncology and Molecular Radiation Science, Johns
Hopkins University, USA

³ Department of Biomedical Engineering, Johns Hopkins University, USA*

Abstract. The task of IMRT planning, particularly in head-and-neck cancer, is a difficult one, often requiring days of work from a trained dosimetrist. One of the main challenges is the prescription of achievable target doses that will be used to optimize a treatment plan. This work explores a data-driven approach in which effort spent on past plans is used to assist in the planning of new patients. Using a database of treated patients, we identify the features of patient geometry that are correlated with received dose and use these to prescribe target dose levels for new patients. We incorporate our approach in a quality-control system, identifying patients with organs that received a dose significantly higher than the one recommended by our method. For all these patients, we have found that a replan using our predicted dose results in noticeable sparing of the organ without compromising dose to other treatment volumes.

Keywords: Data-driven IMRT planning, achievable dose querying, Overlap Volume Histogram.

1 Introduction

We explore a data-driven approach for achievable dose querying in intensity-modulated radiation therapy (IMRT) planning. Our work is motivated by the time-intensive nature of the planning process, in which a dosimetrist specifies target dosage for the different organs and an optimization process is used to find the treatment plan that best meets these objectives. In general, a dosimetrist will seek doses that minimize the radiation received by organs at risk (OARs) while providing sufficient coverage of the primary treatment volume (PTV).

The pitfalls of specifying dosage objectives are two-fold: If the dosimetrist specifies too low a target dose to the OAR, the treatment plan may not be realizable and the treatment objectives will not be met. Conversely, if the dosimetrist specifies too high a dose to the OAR, the treatment plan will be realized at the

* This research was supported in part by the generosity of Paul Maritz, Philips Radiation Oncology Systems (Madison, WI), and by Johns Hopkins University internal funds.

cost of excessive irradiation of essential organs. As a result, IMRT planning often requires numerous trial-and-error iterations of computationally intensive dose simulation in order to finalize on a treatment plan.

Our goal is to facilitate the planning process by using information from previously treated patients in designing treatment plans for new patients. To this end, we use a database of previously treated patients to query for the lowest *achievable* dose to a new patient's OARs. We do this in two steps. First, for each OAR, we identify a subset of *relevant* database patients – patients that have OARs which we expect to be at least as difficult to spare as the new patient's. Then, using the assumption that a dose achieved by these more challenging patients should be achievable by the new patient, we set the target dose for the new patient's OAR to the *minimum* dose achieved, over the subset of relevant patients.

Contribution. To prescribe an achievable dose for a query patient's OAR, we need to identify the database patients that are harder to treat than the query. As with previous approaches (e.g. [4,1,7,3,6]), we use a shape descriptor that characterizes the geometry of the treatment volumes to retrieve patients from the database. However, while previous approaches have used shape descriptors primarily for measuring similarity, our research must address the more challenging task of using the geometric information to *rank* achievable dose.

We address this challenge by using the patient database to learn the correlation between features of the shape descriptors and achievable doses. Specifically, we use the recently developed Overlap Volume Histogram (OVH) [6], which characterizes the distribution of distances of points within the OAR from the surface of the PTV. For each OVH, we extract a set of features (e.g. the mean of the distribution, the standard deviation, *etc.*) and compute the rank correlation between the features of the OVH and the dose the associated OAR received.

Then, we predict that a database patient is more difficult to treat if a comparison of the database and query patients' features indicate that the database patient is more difficult to plan, for *every* significantly correlated feature. That is, a database patient's organs are said to be more difficult to spare than the query's if they have larger feature values for every feature that is positively correlated with dose and smaller feature values for every feature that is negatively correlated.

We deployed our dose querying as a means for quality-control in a database of 91 head-and-neck patients. Using our approach we identified database patients whose OARs received a dose exceeding the dose recommended by the RTOG treatment protocol [2] but for which our method returned a dose that would meet the requirement. These patients were flagged for re-planning, with the retrieved dose used to set the target dose for optimization. For each of these patients, the re-plan lowered the dose to the OAR without compromising dose to other treatment volumes and resulted in treatments whose achieved dose were never larger than one Gy of the dose prescribed by our method.

Related Work. In the works most similar to ours, Hunt *et al.* [5] and Wu *et al.* [9] seek to address a similar problem by explicitly using the properties of the OVH to predict dose. For example, Hunt *et al.* use the overlap of the PTV with

the OAR (equivalent to the zero-crossing of the OVH) to predict dose, while Wu *et al.* use the OVH values to predict Dose Volume Histogram (DVH) values.

The use of rank correlation in our method distinguishes our approach in two ways. First, by using rank correlation to define relevant OAR instances, we are able to query for achievable dose without explicitly fitting a model that relates OVH values to dose. Second, our use of rank correlation allows us to identify predictive features on a per-OAR basis. Thus while we may identify the zero-crossing as an important feature for organs close to the PTV (where it is an indication of the proximity of the organ to the tumor, and hence a good predictor of the radiation the organ will receive), we may disregard it for organs distant from the PTV (where we expect it to always be zero, and hence have no predictive value).

Outline. The individual steps of our implementation are described in detail in the following sections.

- **Feature Definition:** We define a set of candidate OVH features and compute their correlations with OAR dose in Sec. 2.
- **Patient Relevance and Dose querying:** We formalize our notion of patient *relevance* based on feature sets and their correlations to dose in Sec. 3.
- **Feature Selection:** We select a maximally predictive subset of features based on our notion of relevance in Sec 4.
- **Dose Prediction:** We apply our achievable dose querying framework for plan quality control and present results in Sec. 5.

2 Overlap Volume Histogram Features

Recently, Kazhdan *et al* [6] introduced the Overlap Volume Histogram (OVH) as a means to characterize the spatial relationship between an OAR and a PTV. For a PTV T and OAR O , the value of the OVH of O with respect to T at distance t is defined as the fraction of volume of O that is at most a distance of t from T :

$$\text{OVH}_{O,T}(t) = \frac{|\{p \in O | d(p, T) \leq t\}|}{|O|},$$

where $d(p, T)$ is the signed distance of p from T 's boundary and $|\cdot|$ represents the volume of a set.

In this paper we will consider the OVH distributions in their normalized differential form (dOVH). Seen in this way, they can be interpreted as probability density functions. Given an OAR O , a target volume T , and a distance t , $\text{dOVH}_{O,T}(t)$ represents the relative likelihood that a uniformly randomly selected point in O will be at a distance of t from T .

Fig. 1 illustrates examples of these distributions for this paper's particular organs of interest with respect to PTV⁷⁰ for three patients in our database. We note that the different organs' dOVH curves have characteristically different shapes. Also of interest is the fact that in all cases except the mandible, the zero crossing value is zero and thus provides no information.

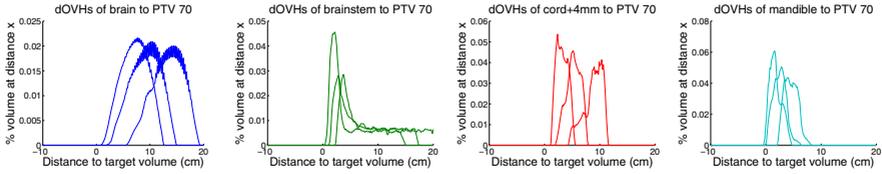


Fig. 1. dOVHs for the brain, brainstem, cord+4mm and and mandible with respect to PTV 70 for three patients in our database. While exhibiting variability within a same organ, each organ’s dOVH has a characteristically different shape.

The characteristic shapes of the dOVH curves for the different organs motivate us to develop a dose prediction model in which the different features are correlated with dose in different ways and the correlation is determined in an organ-dependent manner. In particular, we consider the following statistical features (and denote them with the parenthesized abbreviation): mean (*mu*), standard deviation (*sig*), skewness (*skw*), median (*med*), mode (*mod*), minimum (*min*), maximum (*max*), and the zero-crossing (*zc*) of the cumulative OVH. Additionally, we compute each of these for PTVs at 58.1, 63 and 70 Gy prescription values, and denote each feature with the corresponding number. For example, for a given patient and OAR, $mu70$ denotes the mean of the dOVH distribution of this organ with respect to PTV⁷⁰. This results in a total of 24 features (the 8 statistical features mentioned above times the 3 PTVs).

We begin by analyzing the rank correlation of the above-mentioned features in relation to the maximum dose delivered to each of our database cases. This provides a better alternative to standard linear correlation since we wish to make minimal assumptions as to the nature of the functional relationship between the feature and the dose value. Furthermore, while some features are expressed in linear distance units, (e.g. *mu*), others are cubic (e.g. *zc*), while yet others are unitless (e.g. *skw*). As a measure of rank correlation, we use the Kendall tau coefficient, which is defined as $\tau = (n_c - n_d) / (\frac{1}{2}n(n-1))$, where n_c is the number of concordant pairs (i.e. pairs in which the ordering of the feature values agrees with that of dose), n_d is the number of discordant pairs (pairs in which the ordering is reversed), and the denominator expresses the total number of pairs in the data set of n entries. The value of τ ranges from -1 to 1 and its sign indicates if the correlation is positive or negative.

3 Achievable Dose Querying

To query for achievable dose to an organ, we define the notion of a *relevant organ*. The intuition is that a database OAR will be relevant to a query OAR if it is at least as difficult to spare when irradiating the PTV. In doing so, we query for the dose to each organ independently, implicitly assuming that organ dose is primarily dictated by its proximity to the PTV, as in [6,9], and not by the dose received by other organs.

Let us assume for the moment that we have selected a subset of the features discussed above $F = \{f_1, f_2, \dots, f_n\}$. We formalize the notion of relevance as follows. Given treatment plans for two OARs of the same type with corresponding dOVHs o_1 and o_2 , we say that o_2 is relevant to o_1 , and denote it $o_1 \prec_F o_2$, if for all i it holds that $f_i(o_1) < f_i(o_2)$ if $\tau(f_i) \geq 0$ and $f_i(o_1) > f_i(o_2)$ if $\tau(f_i) < 0$.

Using our database, we query for the achievable dose to a query OAR as the minimum of the doses delivered to all relevant database cases; that is:

$$\text{achievable}(o) = \min\{\text{maxdose}(o_i) : o \prec_F o_i\}$$

where $\text{maxdose}(o_i)$ denotes the maximum dose delivered to the organ associated to o_i . Notice that this formulation generalizes the approach of Wu *et al.* [9] reducing to it under the feature set $F = \{\text{min}58, \text{min}63, \text{min}70\}$.

4 Feature Selection

The above framework for achievable dose querying requires that we determine a set of features to use. In principal, we could use all features. However, the \prec operator induces a partial order and, as such, not all pairs of feature vectors are comparable. Specifically, given dOVHs o_1 and o_2 it can easily be the case that neither $o_1 \prec o_2$ nor $o_2 \prec o_1$ is true. As we increase the number of features, the tendency will be for fewer pairs to be comparable, exponentially fewer in the general case, thus dramatically reducing the size of instances relevant to a query. This observation provides a strong incentive to choose a small set of features under which the \prec operator is most predictive of dose relationship while still allowing for many instances to be comparable.

We can measure the predictiveness of a candidate set of features F by using a generalization of rank correlation to this multi-feature case. For all dOVH pairs o_i and o_j of a given organ type with corresponding maximum dose values d_i and d_j , we count the number of times in which the \prec relation agrees with the dose relation, that is $o_i \prec_F o_j$ and $d_i < d_j$, denoted n_c . We also count the total number of comparable pairs n_{\prec} , that is, pairs such that $o_i \prec_F o_j$ is true. We define our measure of predictiveness as n_c/n_{\prec} and our measure of comparability as $n_c/(\frac{1}{2}n(n-1))$.

To increase robustness to small perturbations of dose, we only examine pairs of organs for which $|d_i - d_j| > \frac{1}{2}\sigma$, where σ is the standard deviation of all observed dose values to the organ in question. Note that we do not count discordant pairs as in τ since the directionality of the correlation has been taken into account in the definition of \prec .

Our goal is to select the subset of features that maximizes predictiveness while retaining a minimum measure of comparability. Given that the space of all subsets is too large to explore exhaustively, we formulate a greedy approach. Beginning with an empty set $F_0 = \emptyset$, at each iteration k we select the feature f_i that maximizes the the predictiveness of $F_k \cup \{f_i\}$ and thus define F_{k+1} . We terminate the iteration when the comparability of F_k drops below 50%. This is formalized in Algorithm 1.

Algorithm 1. Select Features (*dOVH list O* , *max dose list D*)

Require: the i -th member of D is the max dose received by the organ instance whose dOVH is the i -th instance of O .

Feature set $F \leftarrow \emptyset$

repeat

$f \leftarrow \operatorname{argmax}_f \text{predictiveness}(F \cup \{f\}, O, D)$

$c \leftarrow \text{comparability}(F \cup \{f\}, O)$

if $c \geq .5$ **then**

$F \leftarrow F \cup \{f\}$

end if

until $c < .5$

return F

5 Experimental Results

We ran our feature selection algorithm on a clinical database consisting of 91 patients, their OAR contours, and associated DVHs for their clinical plans. We applied our method to four high priority organs: brain, brainstem, cord+4mm expansion and mandible and extracted the following features as maximally predictive of achievable maximum dose while maintaining at least 50% comparability.

Organ	F	Pred (%)	Comp (%)
Brain	$zc58, max58, mu58$	98.38	50.03
Brainstem	$min58, skw58$	75.12	64.64
Cord+4mm	$mod70, mod58, min58, min70$	68.88	53.17
Mandible	$zc70, max63, min58, min63, max70$	99.11	51.01

The way these numbers are interpreted is, for instance in the case of the mandible, of all pairs of organ instances in our database, 51.01% are comparable according to the relevance relationship. Of those instances, in 99.11% of the cases the relevance relationship was in agreement with the dose relationship. We note that in the case of the brain and brainstem, the most predictive features come from the relationship to the PTV^{58.1}. Clinicians take great care to avoid contamination from the higher dose PTV⁶³ and PTV⁷⁰ onto these organs. In the case of the other two, the higher dose PTVs come into play, especially in the mandible where overlap is frequent and the overlap volume with the PTV⁷⁰ (given by $zc70$) is selected first.

To evaluate our method, we applied it to the database. For each of the four organs we found an instance where the RTOG protocol max dose threshold [2] had been exceeded. In our database of 91 patients, there were 31 instances where the dose exceeded the recommended value for the brain, 35 for the brainstem, 66 for the cord+4mm and 12 for the mandible. We chose those instances where the dose was highest and our method predicted the threshold could be met. We re-planned each of these instances using the clinical plan as a starting point, lowering the objective of the organ in question as indicated by our query. In each instance, our query results were very nearly met or exceeded without sacrificing coverage or dose to other OARs. These results are summarized in Fig. 2.

Patient #	Organ Dose Point	Clinical Plan (Gy)	Est. Achievable (Gy)	Re-plan (Gy)
<i>a</i>	brain <i>D0</i>	60.24	52.02	52.91
<i>b</i>	brainstem <i>D0</i>	60.04	52.75	53.70
<i>c</i>	cord+4mm <i>D0</i>	50.26	46.86	42.91
<i>d</i>	mandible <i>D0</i>	77.80	71.33	71.10

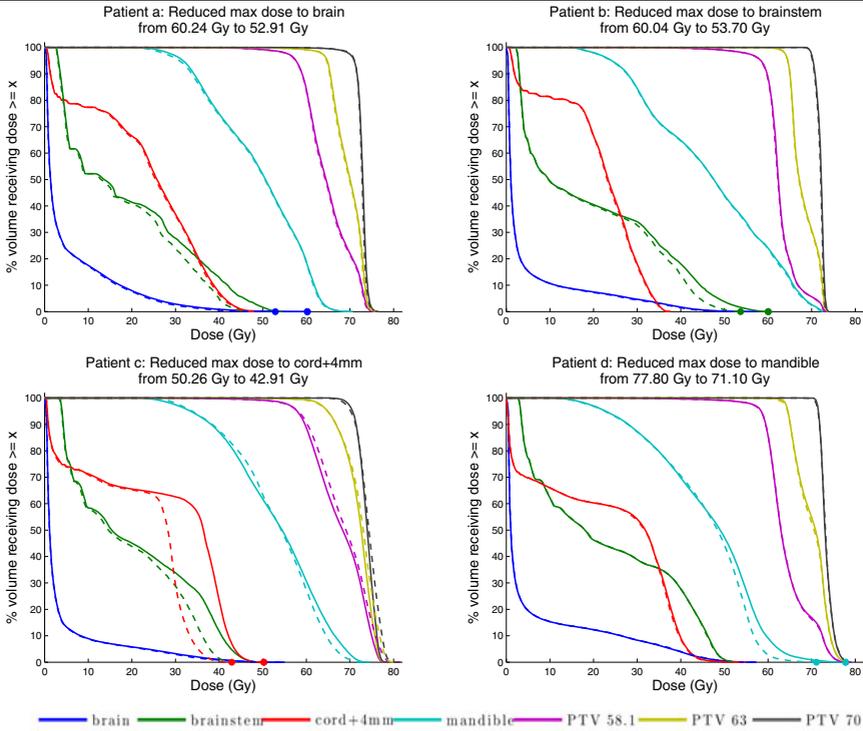


Fig. 2. DVHs resulting from re-planning patients based on our achievable dose query results. In each case, the recommended max dose threshold for the organ had been violated in the clinical plan (solid line) and was met in a re-plan (dashed line), which closely approximated our query results without sacrificing coverage or dose to other OARs.

6 Discussion and Future Work

We have presented a statistical approach for the querying of achievable dose to OARs in IMRT planning of head-and-neck cancer patients. In this presentation, our description has focused on the use of the OVH in prescribing treatment for head-and-neck patients. In future work, we will consider generalizing our approach in two ways. First, we will consider the use of other shape descriptors for defining feature sets for dose prediction. Second, we will explore the use of other OARs and other treatment sites. In concurrent work, we have used a simpler, heuristic OVH feature set to automatically generate all necessary planning constraints [8]. In future work, we will attempt to use the feature selection presented here in the same manner.

With regards to limitations, it should be noted that the quality of our query results will naturally be determined by the quality and case coverage of the database being queried. Furthermore, our approach relies on an assumption that the relative configuration of anatomic structures is static and cannot be purposefully altered by the clinician to facilitate treatment. This may not hold for other sites such as thorax and prostate respectively.

References

1. Ankerst, M., Kastenmüller, G., Kriegel, H., Seidl, T.: 3D shape histograms for similarity search and classification in spatial databases. In: Güting, R.H., Papadias, D., Lochovsky, F.H. (eds.) *SSD 1999. LNCS*, vol. 1651, pp. 207–228. Springer, Heidelberg (1999)
2. Eisbruch, A., Chao, K.C., Garden, A.: Phase I/II study of conformal and intensity modulated irradiation for oropharyngeal cancer (RTOG 0022). Radiation Therapy Oncology Group of the American College of Radiology (2004)
3. Funkhouser, T., Min, P., Kazhdan, M., Chen, J., Halderman, A., Dobkin, D., Jacobs, D.: A search engine for 3D models. *ACM Transactions on Graphics* 22(1), 83–105 (2003)
4. Horn, B.: Extended Gaussian images. *Proceedings of the IEEE* 72, 1656–1678 (1984)
5. Hunt, M.A., Jackson, A., Narayana, A., Lee, N.: Geometric factors influencing dosimetric sparing of the parotid glands using IMRT. *International journal of radiation oncology, biology, physics* 66, 296–304 (2006)
6. Kazhdan, M., Simari, P., McNutt, T., Wu, B., Jacques, R., Chuang, M., Taylor, R.: A shape relationship descriptor for radiation therapy planning. *Medical Image Computing and Computer-Assisted Intervention* 5762/2009(12), 100–108 (2009)
7. Osada, R., Funkhouser, T., Chazelle, B., Dobkin, D.: Matching 3D models with shape distributions. In: *IEEE International Conference on Shape Modeling and Applications*, pp. 154–166 (2001)
8. Wu, B., Ricchetti, F., Sanguineti, G., Kazhdan, M., Simari, P., Jacques, R., Taylor, R., McNutt, T.: A data-driven approach to generating achievable dvh objectives in imrt treatment planning. *International Journal of Radiation Oncology, Biology, Physics* (2010) (to appear)
9. Wu, B., Ricchetti, F., Sanguineti, G., Kazhdan, M., Simari, P., Chuang, M., Taylor, R., Jacques, R., McNutt, T.: Patient geometry-driven information retrieval for IMRT treatment plan quality control. *Medical Physics* 36(12), 5497–5505 (2009)