Published in final edited form as:

Int J Radiat Oncol Biol Phys. 2012 October 1; 84(2): e245–e249. doi:10.1016/j.ijrobp.2012.03.031.

Normal Tissue Anatomy for Oropharyngeal Cancer: Contouring Variability and its Impact on Optimization

Mary Feng, MD, Candan Demiroz, MD, Karen A. Vineberg, MSc, Avraham Eisbruch, MD, and James M. Balter, PhD

Department of Radiation Oncology, University of Michigan, Ann Arbor, MI, USA

Abstract

Purpose—Although variability in target delineation has been studied in head and neck cancer, variability in normal tissue delineation has not. This study evaluated the variability of organ at risk (OAR) delineation and the resulting impact on intensity-modulated radiation therapy (IMRT) treatment plan optimization.

Methods and Materials—An expert panel of three radiation oncologists jointly delineated OARs, including the parotid and submandibular glands (SM), pharyngeal constrictors (PC), larynx, and glottis (GL), in 10 patients with advanced oropharynx cancer in 3 contouring sessions, spaced at least 1 week apart. Contour variability and uncertainty, as well as their dosimetric impact on IMRT planning for each case, were assessed.

Results—The mean difference in total volume for each OAR was 1cm^3 (σ 0.5). Mean fractional overlap was 0.7 (σ 0.1), and was highest (0.8) for the larynx and bilateral SMs and parotids, and lowest (0.5) for the PC. There were considerable spatial differences in contours, with the ipsilateral parotid and PC displaying the most variability (0.9 cm), which was most prominent in cases where tumors obliterated fat planes. Both SMs and the glottis had the smallest differences (0.5 cm). The mean difference in OAR dose was 0.9 Gy (range 0.6–1.1, σ 0.1), with the smallest difference for the GL and largest for both SMs and the larynx.

Conclusions—Despite substantial difference in OAR contours, optimization was barely affected, with a 0.9 Gy mean difference between optimizations, suggesting relative insensitivity of dose distributions for IMRT of oropharynx cancer to the extent of OARs.

Keywords

organs at risk; contour variability; adaptive planning; head and neck cancer

© 2012 Elsevier Inc. All rights reserved.

Corresponding author: Mary Feng, M.D. 1500 East Medical Center Drive, UH B2C490 SPC 5010, Ann Arbor, MI 48109. Phone: 734-647-9956. Fax: 734-763-7370..

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Although variability in target delineation has been well-characterized, variability in normal tissue delineation has not. In our study, despite substantial differences in contours, these geometric differences barely affected sparing of these organs using IMRT. This demonstrates relative insensitivity of dose distributions to the exact extent of OARs. This will help guide adaptive paradigms, which must consider the actual sensitivity of the normal tissue configuration to uncertainty in extent, and its related dosimetric implications.

Presented at the 52nd Annual Meeting of the American Society of Therapeutic Radiology and Oncology, San Diego, CA, October 31–November 4, 2010.

Conflicts of Interest Notification

None

Introduction

Radiotherapy for oropharynx cancer is technically complicated, since multiple targets are in close proximity to multiple organs at risk (OARs). Typically, OARs avoided to minimize the risk of late side effects include the parotid glands and submandibular glands for xerostomia, and the pharyngeal constrictors, larynx, and glottis for dysphagia and aspiration.(1-5)When targets and OARs are geometrically close or even overlapping, treatment plan optimization and IGRT strategies involve balances between target coverage and normal tissue sparing. Inaccuracy in structure delineation thus can directly and artificially impact the planned and delivered dose to target and normal organs. Multiple studies have demonstrated variability in target delineation in head and neck cancer, as well as other malignancies.(6-9) However, no study to date has considered the effect of OAR variability in treatment planning for head and neck cancer. As adaptive paradigms are being explored, the actual sensitivity of the normal tissue configuration to uncertainty in extent, and its related dosimetric implications, must be understood. This paper begins to explore these issues from the standpoint of the initial model of normal tissue used for treatment planning.

Methods and Materials

As part of an IRB-approved prospective study investigating changes in tumors and normal tissues during treatment, 10 patients, undergoing radiation therapy for oropharyngeal cancer, consented to allow analysis of their planning CT scans, which were acquired with the patient immobilized in a 5-point thermoplastic mask, after administration of intravenous contrast. An expert panel of three radiation oncologists (M.F., C.D., and A.E.) jointly delineated OARs, including the parotids, submandibular glands (SM), pharyngeal constrictors (PC), larynx, and glottis (GL). Pharyngeal constrictors were contoured according to Feng, et al 2007,(10) and the parotid glands included both the superficial and deep lobes. For each case, 3 contouring sessions were conducted, spaced at least 1 week apart, for 30 total sets of OARs. Each session was blinded and executed with no available access to prior contouring information, for accurate assessment of contour variability.

For each OAR, contour variability was measured by the fractional overlap (FO), the intersection divided by the union of the structure from all contouring sessions, with 1 signifying perfect overlap (i.e.no variability on contour extent). Contour delineation uncertainty was assessed by expanding the intersection volume until it completely encompassed the union volume of that OAR. IMRT optimization was performed separately for each of the union (most aggressive sparing of normal tissue based on potential OAR extent), intersection (minimal sparing of normal tissue), and the first set of expert contours (representative of a random sample) for each case, for a total of 30 treatment plans, using an in-house optimization system, UMPLAN. A standard optimization schema was created, utilizing a hierarchical optimization method known as lexicographic ordering for consistent trade-offs between competing optimization goals.(11) A single cost function was applied across all patients. Prescribed doses were 70 Gy at 2 Gy per fraction to gross disease PTVs and 59-63 Gy at 1.7-1.8 Gy per fraction to low- and high-risk PTVs, respectively, delivered concomitantly. Optimization goals were as follows. Priority 1: maximum dose to spinal cord, brainstem, mandible and any part of the patient's anatomy <45 Gy, <54 Gy, <70 Gy, and <75 Gy, respectively. Priority 2: 99% of PTV70 to receive 70 Gy. Priority 3: Mean larynx dose <20 Gy. Priority 4: 100% of subclinical PTVs to receive 95% of the prescribed doses. Priority 5: Mean parotid dose <26 Gy, mean oral cavity <30 Gy, mean submandibular glands <39 Gy, mean pharyngeal constrictors <50 Gy, and mean esophagus <45 Gy. Priority 6: Maximum PTV doses <105% of prescribed doses and minimize dose to all normal tissues. The resulting mean doses to all OARs were compared between optimizations.

Results

Volume differences in OARs between contouring sessions are displayed in Table 1. The mean difference in total volume for each OAR across contouring sessions was 1 cm^3 (σ 0.5 cm^3 , range $0.1\text{--}5\text{ cm}^3$), or 6% (σ 3%, $0.1\text{--}45\%$). The pharyngeal constrictors had the largest mean absolute difference in volume (1.8 cm^3), while the glottis had the largest mean fractional difference in volume (16%). The smallest absolute difference in volume was in both submandibular glands (0.5 cm^3), while the smallest difference as a percentage of volume was in the ipsilateral parotid gland (2.1%). Mean volumes for the OARs across contouring sessions were 33.4, 29.6, 10.7, 9.2, 19.4, 17.6, and 5.7 cc for the ipsilateral parotid, contralateral parotid, ipsilateral submandibular, contralateral submandibular, pharyngeal constrictors, larynx, and glottis, respectively.

There was considerable variability in contours as assessed by fractional overlap (FO), Table 2. Mean FO was 0.7 (σ 0.1, range 0.4–0.9). Agreement was the highest for the larynx and bilateral parotid and submandibular glands (0.8) and lowest (0.5) for the pharyngeal constrictors. There were substantial spatial differences in contours, with a mean maximum distance between intersection and union volumes of 0.7 cm (σ 0.2, range 0.3–1.4 cm) (Table 3).

The ipsilateral parotid and PC displayed the most variability, with a 0.9 cm expansion of the intersection volumes needed to encompass the unions. This variability was most prominent in cases where tumors obliterated fat planes around the muscles. Both SMs and the glottis had the smallest differences, 0.5 cm. The maximum distance between the intersection and union for each OAR is displayed in Figure 1. Full coverage of the union was attained after a 0.7 cm expansion of the ipsilateral submandibular, and a 0.8 cm expansion of the contralateral submandibular and glottis. However, a 1.4 cm expansion was needed to fully cover the ipsilateral parotid and constrictor unions.

There was a 0.5 Gy difference in dose to targets across the three optimizations. The mean difference in OAR dose was 0.9 Gy (range 0.6–1.1, σ 0.1), with the smallest difference observed for the GL and largest for both SMs and the larynx (Table 4). Figure 2 illustrates two cases, first a typical case (panels A–D), with a 0.3 Gy mean difference in OAR mean doses across optimizations, despite an up to 1 cm spatial difference in individual contours, and the second (panels E–H), the case with the largest differences in OAR doses, 1.7 Gy. Mean doses for the OARs across contouring sessions were 36.7, 22.7, 65.8, 48.0, 53.7, 23.9, and 9.6 Gy for the ipsilateral parotid, contralateral parotid, ipsilateral submandibular, contralateral submandibular, pharyngeal constrictors, larynx, and glottis, respectively.

Dose distributions displayed are from optimizations based on OAR contours generated in the first session. In the second case, there was more variability in OAR contours, particularly for the constrictors, due to tumor involvement. In general, when contour differences fell within areas of homogeneous dose, the dosimetric differences were small. However, if these were within areas of rapid dose fall-off, dosimetric differences were larger.

Discussion

Several groups have demonstrated substantial inter- and intra-observer differences in target delineation, in the head and neck, as well as breast, prostate, and other sites.(6–8) These studies have demonstrated the underlying contouring variability of individual experts for target delineation, as well as the impact of training and multimodality imaging on reducing the variability in contour extent. In contrast, the variability of organ at risk (OAR) delineation has received little attention. In one of the few papers addressing this issue, Nelms, et al looked at OARs contoured on a single CT dataset for a patient with

oropharyngeal cancer.(12) 32 independent radiation treatment plans were submitted internationally for analysis of OAR delineation variation, as well as differences in calculated doses to these structures for a given IMRT dose distribution. The authors found substantial variation in both, most notably in the brainstem and parotid glands. In another paper, Faggiano and colleagues found that although there was no difference in overall volume of parotid glands when contoured by 3 different observers on kVCT, there was significant variation when MVCT was used, and the Dice similarity coefficient was also different when comparing kVCT with MVCT.(13)

The impact of uncertainty in OAR extent on treatment planning is also largely unknown to date. In the era of adaptive therapy, including reoptimization based on automated deformation of OAR and/or target volumes, it is imperative to assess the effect of OAR contouring uncertainty on not only calculated dose but also on plan modifications.

We found that despite delineating OAR contours by joint consensus of an expert panel, substantial inter-session variability was still apparent. Spatial differences in contours averaged 0.7 cm, with a maximum of 1.4 cm. A low fractional overlap, averaging 0.7, and as low as 0.4, further illustrates the lack of perfect agreement. Pharyngeal constrictors had the largest difference in contours and volume, as well as the lowest fractional overlap. This is likely due to particular uncertainty in cases with local involvement by the tumor, which obscures tissue planes and distorts anatomy, making identification of these delicate structures quite challenging. On the other hand, submandibular glands, which are relatively much more discrete, had minimal variation on contours or resulting calculated dose.

Despite spatial, volumetric, and overlap differences, dose differences were small. This can be explained by many reasons including simple geometry and the high priority we placed on target coverage in our optimizations. In the head and neck, targets and OARs are all in close proximity. This makes OAR sparing quite challenging, and the degree of sparing likely depends more on the distance from targets rather than the exact contour of that OAR. Structures in the low dose regions are less likely to see significant dose changes even with variability in their spatial extent. It is conceivable that organs at risk that are somewhat proximal to the PTV (but not overlapping), and fall in regions of intermediate to high doses, especially on dose gradients, may more likely be impacted by contouring uncertainties. Our current observations, however, have not elucidated such an instance for this specific example of oropharyngeal IMRT. The results of this study should not be used as justification to omit OARs or replace exact contours with easily or quickly schematic representations, however, as we used contours generated by consensus of a panel of radiation oncologists.

Conclusion

Despite substantial differences in OAR contours up to 1.4 cm, optimization was only minimally affected, with less than a 1 Gy difference in dose, suggesting relative insensitivity of dose distributions to the extent of OARs for IMRT of oropharynx cancer.

Acknowledgments

Supported in part by P01-CA-59827.

References

1. Eisbruch A, Kim HM, Terrell JE, et al. Xerostomia and its predictors following parotid-sparing irradiation of head-and-neck cancer. *Int J Radiat Oncol Biol Phys.* 2001; 50:695–704. [PubMed: 11395238]

2. Deasy JO, Moiseenko V, Marks L, et al. Radiotherapy dose-volume effects on salivary gland function. *Int J Radiat Oncol Biol Phys.* 2010; 76:S58–63. [PubMed: 20171519]
3. Murdoch-Kinch CA, Kim HM, Vineberg KA, et al. Dose-effect relationships for the submandibular salivary glands and implications for their sparing by intensity modulated radiotherapy. *Int J Radiat Oncol Biol Phys.* 2008; 72:373–382. [PubMed: 18337023]
4. Eisbruch A, Kim HM, Feng FY, et al. Chemo-IMRT of Oropharyngeal Cancer Aiming to Reduce Dysphagia: Swallowing Organs Late Complication Probabilities and Dosimetric Correlates. *Int J Radiat Oncol Biol Phys.* 2011
5. Feng FY, Kim HM, Lyden TH, et al. Intensity-modulated chemoradiotherapy aiming to reduce dysphagia in patients with oropharyngeal cancer: clinical and functional results. *J Clin Oncol.* 2010; 28:2732–2738. [PubMed: 20421546]
6. Riegel AC, Berson AM, Destian S, et al. Variability of gross tumor volume delineation in head-and-neck cancer using CT and PET/CT fusion. *Int J Radiat Oncol Biol Phys.* 2006; 65:726–732. [PubMed: 16626888]
7. Li XA, Tai A, Arthur DW, et al. Variability of target and normal structure delineation for breast cancer radiotherapy: an RTOG Multi-Institutional and Multiobserver Study. *Int J Radiat Oncol Biol Phys.* 2009; 73:944–951. [PubMed: 19215827]
8. Fiorino C, Reni M, Bolognesi A, et al. Intra- and inter-observer variability in contouring prostate and seminal vesicles: implications for conformal treatment planning. *Radiother Oncol.* 1998; 47:285–292. [PubMed: 9681892]
9. Steenbakkers RJ, Duppen JC, Fitton I, et al. Reduction of observer variation using matched CT-PET for lung cancer delineation: a three-dimensional analysis. *Int J Radiat Oncol Biol Phys.* 2006; 64:435–448. [PubMed: 16198064]
10. Feng FY, Kim HM, Lyden TH, et al. Intensity-modulated radiotherapy of head and neck cancer aiming to reduce dysphagia: early dose-effect relationships for the swallowing structures. *Int J Radiat Oncol Biol Phys.* 2007; 68:1289–1298. [PubMed: 17560051]
11. Jee KW, McShan DL, Fraass BA. Lexicographic ordering: intuitive multicriteria optimization for IMRT. *Phys Med Biol.* 2007; 52:1845–1861. [PubMed: 17374915]
12. Nelms BE, Tome WA, Robinson G, et al. Variations in the contouring of organs at risk: test case from a patient with oropharyngeal cancer. *Int J Radiat Oncol Biol Phys.* 2012; 82:368–378. [PubMed: 21123004]
13. Faggiano E, Fiorino C, Scalco E, et al. An automatic contour propagation method to follow parotid gland deformation during head-and-neck cancer tomotherapy. *Phys Med Biol.* 2011; 56:775–791. [PubMed: 21239848]

Figure 1.
Maximum distance between intersection and union of organ at risk volumes, across contouring sessions.

Figure 2.

Variability in contours and relationship to dose gradients. Panels A-D: typical case with small (0.3 Gy) mean difference in OAR mean doses across optimizations. Panels E-H: case with largest dose difference (1.7 Gy).

Table 1**OAR volume differences between contouring sessions**

Structure	Mean volume (cc)	Cc Difference			Percent Difference		
		Mean	Std Dev	Range	Mean	Std Dev	Range
Ipsilateral parotid	33.4	1	1.1	0.2-5.1	2.1	1.6	0.2-5.4
Contralateral parotid	29.6	1.1	1	0.1-3.3	4.9	5.2	0.1-12.7
Ipsilateral submandibular	10.7	0.5	0.4	0.1-1.2	5.3	4.4	0.7-15.9
Contralateral submandibular	9.2	0.5	0.4	0.1-1.5	5.5	3.9	0.6-13.5
Pharyngeal constrictors	19.4	1.8	1.3	0.1-4.7	4.9	4.8	0.3-21.8
Larynx	17.6	0.8	0.8	0.2-3.3	4.7	4.3	0.7-12.9
Glottis	5.7	1.5	1.7	0.1-4	16.3	12.7	1.1-45.2
Mean		1	1		6.2	5.3	
Std Dev		0.5	0.5		4.6	3.4	

Table 2
Fractional overlap differences between contouring sessions

Structure	Mean	Std Dev	Range
Ipsilateral parotid	0.8	0.07	0.6-0.8
Contralateral parotid	0.8	0.04	0.7-0.9
Ipsilateral submandibular	0.8	0.04	0.7-0.9
Contralateral submandibular	0.8	0.04	0.7-0.9
Pharyngeal constrictors	0.5	0.07	0.4-0.6
Larynx	0.8	0.1	0.6-0.9
Glottis	0.7	0.2	0.4-0.9
Cord	0.7	0.05	0.6-0.8
Esophagus	0.7	0.1	0.5-0.9
Mean	0.7	0.08	
Std Dev	0.1	0.05	

Table 3
Expansion distances of intersection volumes necessary to encompass union volumes (in cm)

Structure	Mean	Std Dev	Range
Ipsilateral parotid	0.9	0.3	0.5-1.4
Contralateral parotid	0.8	0.8	0.5-1.4
Ipsilateral submandibular	0.5	0.5	0.4-0.7
Contralateral submandibular	0.5	0.5	0.4-0.8
Pharyngeal Constrictors	0.9	0.9	0.5-1.4
Larynx	0.6	0.6	0.4-1.2
Glottis	0.5	0.5	0.3-0.8
Mean	0.7	0.6	
Std Dev	0.2	0.2	

Table 4
Dose differences between optimizations

	Mean Dose (Gy)	Mean Difference (Gy) (σ , range)		
		EXPERT	UNION	INTERSECTION
Ipsilateral parotid	36.7	0.8 (0.7, 0-3.0)	0.9 (0.8, 0-2.8)	0.9 (0.9, 0-3.2)
Contralateral parotid	22.7	0.8 (0.8, 0-3.5)	0.9 (0.9, 0-3.5)	0.9 (0.8, 0.1-3.5)
Ipsilateral submandibular	65.8	1.0 (1.4, 0-4.5)	1.1 (1.5, 0-5.3)	1.1 (1.4, 0-4.8)
Contralateral submandibular	48.0	0.9 (0.7, 0-2.8)	1.1 (1.1, 0.1-3.6)	1.1 (1.1, 0-3.4)
Pharyngeal constrictors	53.7	0.8 (0.6, 0-2.4)	0.9 (0.7, 0-2.7)	0.8 (0.6, 0-2.3)
Larynx	23.9	1.0 (0.8, 0.2-2.7)	1.1 (1.1, 0.2-3.4)	1.0 (1.0, 0.2-2.5)
Glottis	9.6	0.6 (0.6, 0-1.8)	1.0 (0.9, 0-2.9)	0.7 (0.7, 0.1-2.6)