

Retrospective Validation and Clinical Implementation of Automated Contouring of Organs at Risk in the Head and Neck: A Step Toward Automated Radiation Treatment Planning for Low- and Middle-Income Countries

Rachel E. McCarroll
 Beth M. Beadle
 Peter A. Balter
 Hester Burger
 Carlos E. Cardenas
 Sameera Dalvie
 David S. Followill
 Kelly D. Kisling
 Michael Mejia
 Komeela Naidoo
 Chris L. Nelson
 Christine B. Peterson
 Karin Vorster
 Julie Wetter
 Lifei Zhang
 Laurence E. Court
 Jinzhong Yang

Author affiliations and support information (if applicable) appear at the end of this article.

Corresponding author: Beth M. Beadle, MD, PhD, Department of Radiation Oncology, Stanford University, 875 Blake Wilbur Dr, MC-5847, Stanford, CA 94305; Twitter:@TheCourtLab; e-mail: bbeadle@stanford.edu.

abstract

Purpose We assessed automated contouring of normal structures for patients with head-and-neck cancer (HNC) using a multiatlas deformable-image-registration algorithm to better provide a fully automated radiation treatment planning solution for low- and middle-income countries, provide quantitative analysis, and determine acceptability worldwide.

Methods Autocontours of eight normal structures (brain, brainstem, cochleae, eyes, lungs, mandible, parotid glands, and spinal cord) from 128 patients with HNC were retrospectively scored by a dedicated HNC radiation oncologist. Contours from a 10-patient subset were evaluated by five additional radiation oncologists from international partner institutions, and interphysician variability was assessed. Quantitative agreement of autocontours with independently physician-drawn structures was assessed using the Dice similarity coefficient and mean surface and Hausdorff distances. Automated contouring was then implemented clinically and has been used for 166 patients, and contours were quantitatively compared with the physician-edited autocontours using the same metrics.

Results Retrospectively, 87% of normal structure contours were rated as acceptable for use in dose-volume-histogram–based planning without edit. Upon clinical implementation, 50% of contours were not edited for use in treatment planning. The mean (\pm standard deviation) Dice similarity coefficient of autocontours compared with physician-edited autocontours for parotid glands (0.92 ± 0.10), brainstem (0.95 ± 0.09), and spinal cord (0.92 ± 0.12) indicate that only minor edits were performed. The average mean surface and Hausdorff distances for all structures were less than 0.15 mm and 1.8 mm, respectively.

Conclusion Automated contouring of normal structures generates reliable contours that require only minimal editing, as judged by retrospective ratings from multiple international centers and clinical integration. Autocontours are acceptable for treatment planning with no or, at most, minor edits, suggesting that automated contouring is feasible for clinical use and in the ongoing development of automated radiation treatment planning algorithms.

© 2018 by American Society of Clinical Oncology Licensed under the Creative Commons Attribution 4.0 License

INTRODUCTION

When considering the need for radiation therapy (RT) in low-resource settings around the world, many challenges need to be addressed. The availability of RT resources remains the primary challenge for many regions^{1,2}; however, some find that the lack of trained staff and other workflow bottlenecks, including the contouring of normal

structures, are primary issues.³ Improving availability of RT worldwide depends not only on the deployment of equipment, but more importantly, on training staff and developing more efficient, streamlined, and practical workflows.

To address the availability of RT and simultaneous integration of high-quality techniques, there is considerable interest in the commercial development

of fully automated state-of-the-art technology at an affordable cost, including intensity-modulated RT and volumetric modulated arc therapy, which offer many dosimetric advantages compared with 3D conformal and older 2D planning techniques.^{4,5} These techniques, especially critical for the treatment of head-and-neck cancer (HNC), have been shown to improve outcomes compared with 2D or 3D techniques.^{6,7} However, these techniques require significant training for optimal target and normal tissue identification, treatment planning parameters, and treatment delivery. Contouring is known to be time consuming, has considerable interphysician variability,⁸⁻¹² and includes the component of planning, which introduces the most error.^{13,14} For physicians, the delineation of up to three prescription volumes for as many as 25 normal structures¹⁵ may prove daunting or simply be impossible. Opportunities to automate these steps are crucial to allowing global access to high-quality RT for HNC.

Studies into the use of automated contouring, followed by manual editing and also as a stand-alone contouring method, have found that for HNC treatment plans, automatic segmentation of normal structures can significantly reduce interobserver variability and contouring time.¹⁶⁻¹⁸ Authors have reported on the limited implementation of automated contouring for small structure sets (eg, brachial plexus¹⁹ or heart chambers²⁰) and for other anatomic sites (eg, the prostate²¹). However, these uses of automated contouring require a deviation in workflow for select structures in the treatment plan. Herein, we report on the clinical implementation of a large set of automatically contoured normal structures in the head and neck.

Our objective was to retrospectively validate and assess the clinical implementation of automated contouring of eight normal structures in the head and neck as a crucial step in the development of a fully automated RT planning system for low- and middle-income countries. This included retrospective validation and the clinical implementation of normal structure autocontours generated by using an in-house autocontouring solution. To our knowledge, this is the first comprehensive assessment of automated contouring for head-and-neck structures and of the role automated contouring may play in the workflow of a typical radiation oncology clinic. Collaboration with global radiation oncologists brings to the

forefront the role automated contouring will have in the practical implementation of advanced RT technologies in resource-constrained settings.

METHODS

Contouring Algorithm

The previously developed^{22,23} automated contouring algorithm, termed multiatlas contouring systems,^{19,20} consists of three distinct steps. First, rigid registration is performed between the test patient's simulation computed tomography (CT) scan and the CT scans of each atlas patient using 2D sagittal and coronal projections. Second, the test patient is deformably registered to each test patient using dual-force Demons deformable registration.²⁴ By using the resultant deformation vector fields, the contours from each atlas patient are mapped to the test patient,²⁵ resulting in a number of individual contours equal to the number of atlas patients. Finally, the STAPLE algorithm with a built-in tissue appearance model²⁶ is used to combine the individual segmentations, which generates a fusion contour approximating a true segmentation.

Central to the algorithm is an atlas of patients who are representative of the patients for which the algorithm will be used. In this work, 12 patients recently treated for HNC were selected. The atlas of normal structures included brain, brainstem, cochleae, eyes, lungs, mandible, parotid glands, and spinal cord. These structures were chosen after a small retrospective study that included other normal structures such as the esophagus, lens, optic chiasm, optic nerves, and submandibular glands. The results of that study indicated that these eight normal structures were best suited for autocontouring on the basis of comparison with independently drawn contours. The contours used for atlas building were either extracted from patient treatment plans (reviewed before treatment by a head-and-neck quality assurance peer review clinic²⁷) or created by using thresholding tools. All contours were reviewed by a medical dosimetrist with 13 years of experience and a head-and-neck radiation oncologist with 8 years of experience.

Retrospective Validation of the Automated Contouring Algorithm

Physician ratings. Under an approved institutional review board protocol, the latest 128

patients stored in the database at The University of Texas MD Anderson Cancer Center and with physician-approved contours and treatment plans, were used as the study cohort. Autocontours of the eight normal structures were created and rated by an experienced head-and-neck radiation oncologist as needing no edits, needing minor edits, or needing major edits for use in dose-volume-histogram (DVH)-based planning.

To assess possible rater bias, contours from a subset of 10 patients were reviewed by five additional radiation oncologists currently treating patients with 3D conformal or volumetric modulated arc therapy techniques from four international institutions. Physician agreement was assessed by grouping each pair of ratings (one rating from the primary physician and one from an outside physician) into one of three categories. Category I agreement occurs when the physicians agreed on the degree of editing needed, and Category II agreement indicates that both physicians agreed that no more than minor edits were needed or that major edits were required. Finally, category III agreement indicates that the physicians disagreed on the acceptability of the contour, with one physician indicating that major edits were required and the other indicating that no or minor edits were required.

Quantitative assessment. Autocontours were compared with independently physician-drawn contours from the patient's original clinical plan. The Dice similarity coefficient (DSC), mean surface distance (MSD), and Hausdorff distance (HD) were measured to assess contour accuracy. DSC measures the volume overlap of the contours as a ratio to their total volume, with a minimum value of 0 and a maximum value of 1, which indicates perfect agreement. The symmetric 3D MSD between the two contours has a minimum value of 0, indicating perfect agreement and no maximum value. HD, the maximum Euclidean distance from each point in the physician contour to the nearest point in the autocontour has a minimum value of 0, indicating perfect agreement and no maximum value.

For two of the eight normal structures (lungs and spinal cord), we performed an additional quantitative analysis considering only CT slices contoured by the automated contouring algorithm. This analysis better represents the contouring accuracy of the algorithm (compared with whole-structure quantitative analysis) because it

eliminates errors that arise owing to differences in CT scan extent between the test and atlas patients and among the atlas patients.

Clinical Implementation

Workflow and assessment. After retrospective validation, we began a limited introduction into our head-and-neck clinic at The University of Texas MD Anderson Cancer Center under an approved institutional review board protocol. During 10 months of clinical implementation, autocontours were generated for 166 patients. Typically, the clinical workflow of automated contouring involves initializing the algorithm by a dosimetrist who uses a script in the treatment planning system, importing the structures into the treatment plan, and reviewing which includes any needed editing of the contours by the radiation oncologist. All final contours were reviewed and edited by the attending physician in the same way that initial resident contours would be reviewed; the final contours reflected approval by the physician, with or without editing as deemed appropriate.

Quantitative assessment of contour edit. To assess clinical autocontour edits, the original autocontours were compared with the contours edited for treatment planning by the physician using the DSC, MSD, and HD. In addition to analyzing the eight clinically implemented normal structures, we have also quantitatively compared edits of modified lung and spinal cord autocontours as described in the Quantitative assessment section. This provides a better estimate of the edits independent of scan and atlas extent. Generally, this is reasonable because the autocontoured structure extends inferiorly enough that it is contoured fully within the dose calculation region.

RESULTS

Automated Contouring Algorithm

Of the 12 patients in the contouring atlas, nine were male, 11 had primary oropharynx disease, and one patient had unknown primary cancer. The mean age was 72 years. Ten had clinical stage IVa, one had stage IVb, and one had stage III disease according to the American Joint Committee on Cancer 7th edition. All were treated with curative intent. The patient population for clinical validation primarily featured patients with oropharynx cancer (31%), oral cavity cancer

(13%), and skin cancer (10%). Other primary cancer sites included hypopharynx, larynx, nasopharynx, salivary gland, and thyroid along with leukemia, lymphoma, myeloma, sarcoma, and sinonasal cancer. Ten percent of the 166 patients were younger than age 40 years, 55% were between age 40 and 65 years, and 35% were older than age 65 years. Seventy-two percent of the patients were male.

Retrospective Validation of the Automated Contouring Algorithm

Creation of the normal structures took a mean \pm standard deviation (SD) of 13.2 ± 4.1 minutes when run on a Windows 2008-based personal computer with a 6-core Xeon E5680 3.33-GHz central processing unit and 24 GB of memory. Multithread computing was enabled, which allowed two registration tasks to be run simultaneously.

Physician ratings. Of the eight normal structures, on average, six were retrospectively rated as clinically acceptable for use in DVH-based planning without edits. The remaining two structures had ratings that indicated the need for minor edits for use in DVH-based planning. For all normal structures, 87% received ratings that indicated no need for edits for use in treatment planning. Furthermore, 97% of normal structures received a rating that indicated a need for only minor editing if the structure was near a target volume. Details of the score distribution by structure are provided in [Figure 1](#).

Interobserver variability. For a subset of 10 randomly selected patients, five additional radiation oncologists from four international institutions evaluated the normal-structure autocontours. The radiation oncologists (according to a self-reported questionnaire) had an average of 8.25 years of experience (range, 3.0 to 12.5 years) and contoured and/or reviewed an average of seven patients per week (range, two to 15 patients), spending an average of 95 minutes per patient on contouring (range, 45 to 180 minutes).

For all structures, 45% (245 of 547) of the ratings by the outside physicians matched those of the primary physician (Category I agreements). Considering Category II agreements, the physicians assigned an additional 48% (262 of 547) of the contours to the same group, either as

needing no or minor edits for use (48%) or as needing major edits for use (0%). Finally, only 7% of contours received Category III agreement, indicating disagreement between the two ratings. The percentage of each contour classified into each of the three categories is provided in [Table 1](#).

Retrospective assessment of quantitative contour accuracy. Quantitatively, normal structure autocontours were compared with structures drawn independently by a radiation oncologist using the DSC, MSD, and HD. The DSC was greater than 0.75 for all structures except the cochleae, lungs, and spinal cord, and the MSD was less than 3 mm for all structures except the lungs and spinal cord. In these two structures, the largest difference between the autocontour and physician-drawn contour occurred at the inferior extent and was often contributable to a difference in contouring below the target volume and out of dose calculation range; considering the modified structures, the DSC and MSD improved. [Table 2](#) lists the mean \pm SD for DSC, MSD, and HD for all structures, including the two modified structures. No correlation between physician score and quantitative agreement metrics was found.

Clinical Implementation

Since clinical implementation of the automated contouring software, 22 radiation oncologists have used it to generate normal structure contours for 166 patients. The seven attending physicians who used the tool the most accounted for 23%, 15%, 14%, 9%, 7%, 7%, and 5% of the total use. The mean \pm SD time required for generation of the autocontours was 11.5 ± 3.1 minutes when run on a Windows 2012-based personal computer with an 8-core Xeon E5-2697 v3 2.6-GHz central processing unit and 16 GB of memory. Multithread computing was again implemented. No oversight is required and thus it can occur simultaneously with other required treatment planning tasks. [Figure 2](#) shows the distribution of autocontour edits, as measured by the DSC, MSD, and HD. Only contours remaining in the treatment plan at the time of treatment with the same naming convention were considered.

Notably, radiation oncologists did not edit 49.8% of the contours for treatment planning. As shown in [Figure 2](#), 31% of automatically contoured

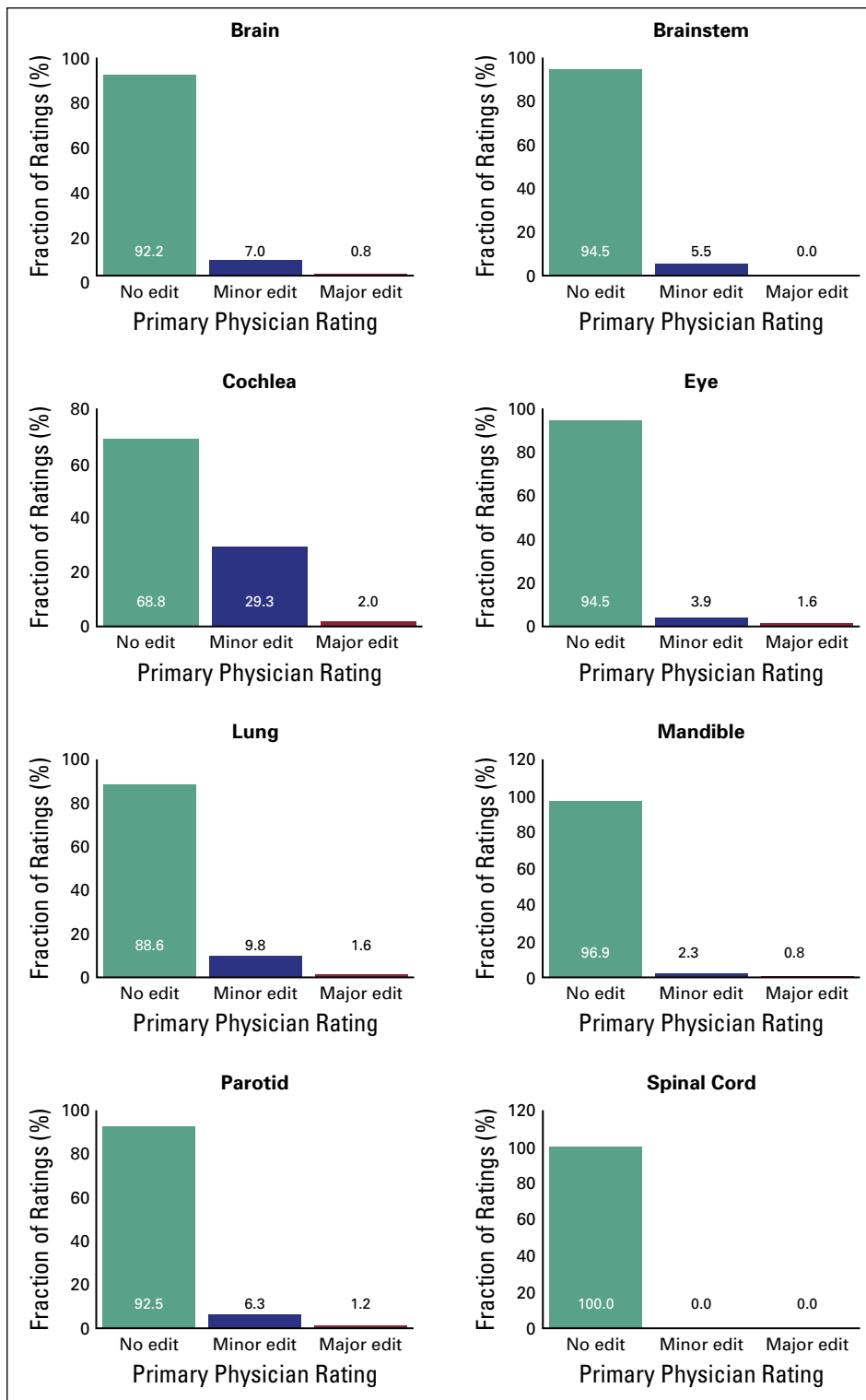


Fig 1. Distribution of the primary physician ratings of the eight automatically contoured normal structures (128 patients). One patient had a surgically removed parotid; for five patients, the lungs were not visible in the patient computed tomography scan, thus no rating was recorded for these structures. The mean physician ratings are displayed in the graphs.

Table 1. Percentage of Agreements in Categories I, II, and III for Various Structures

Structure	Agreements in Category (%)		
	I	II	III
Brain	22	76	2
Brainstem	48	32	20
Cochleae	46	49	5
Eyes	47	42	11
Lungs	60	36	4
Mandible	40	58	2
Parotid gland	57	38	5
Spinal cord	22	68	10
Total	45	48	7

NOTE. Category I indicates the scores by the two physicians matched, and Category II indicates that the ratings did not match but the contours were rated into the same group (either as needing no or minor edits or as needing major edits) by both physicians. Category III represents disagreement between two reviewing physicians.

brainstems, 40% of parotids, and 48% of modified spinal cords were not edited for clinical use. The structures edited least often were the eyes (69%) and modified lungs (74%). The maximum MSD was seen for the unmodified lung (3.29 mm) and unmodified spinal cord (2.65 mm). As before, considering the modified lung avoidance and spinal cord structures, the MSDs were decreased nearly 10-fold to 0.33 mm and 0.20 mm, respectively.

DISCUSSION

This study demonstrates the clinical feasibility of high-quality automated contouring of normal structures relevant for HNC radiation treatment

planning and points toward its applicability in low-resource settings by saving time and removing a major hurdle when transitioning to more advanced RT techniques. This study represents, to the best of our knowledge, the first implementation of an automated contouring process for a large structure set in the head and neck. Retrospective quantitative comparison of autocontours with clinically used contours, ratings from multiple international radiation oncologists and, most importantly, the clinical implementation of automated contouring in a robust clinical practice shows the promise of autocontouring techniques.

In this study, we focused on organs at risk with safety implications (eg, spinal cord and brainstem) or broadly acknowledged toxicity benefit (eg, parotid glands). Other normal tissue structures can be relevant in certain settings, especially for toxicity avoidance (eg, early-stage cancers), including the submandibular glands, optic structures, esophagus, and larynx. These will be the subject of further improvement and be integrated into planning in future studies.

Before the clinical implementation of the software, the algorithm was validated internally and externally. Contours on a subset of patients were rated by both the internal primary physician and by multiple international physicians. Generally, we observed agreement between physicians, but consistent with previous findings,^{10,11} we also observed some interobserver variability. Ninety-three percent received scores from both

Table 2. Quantitative Retrospective Comparison of Autocontours Compared With Physician-Drawn Structures for 128 Patients

Structure	Dice Similarity Coefficient		Mean Surface Distance (mm)		Hausdorff Distance (mm)	
	Mean	SD	Mean	SD	Mean	SD
Brain	0.98	0.01	1.03	0.31	14.01	9.05
Brainstem	0.81	0.12	2.31	1.50	9.10	5.61
Cochlea	0.47	0.16	1.75	0.71	4.87	2.47
Eye	0.85	0.06	1.33	0.45	4.62	1.42
Lungs	0.48	0.12	21.88	10.65	58.02	54.87
Lungs (modified)*	0.93	0.04	2.05	1.24	19.59	6.94
Mandible	0.84	0.07	1.89	1.55	18.63	14.90
Parotid gland	0.78	0.07	2.39	0.80	14.24	6.65
Spinal cord	0.71	0.13	4.77	6.13	40.07	36.05
Spinal cord (modified)*	0.81	0.06	1.16	0.40	4.58	1.22

Abbreviation: SD, standard deviation

*Comparison only for slices contoured by the automated contouring algorithm.

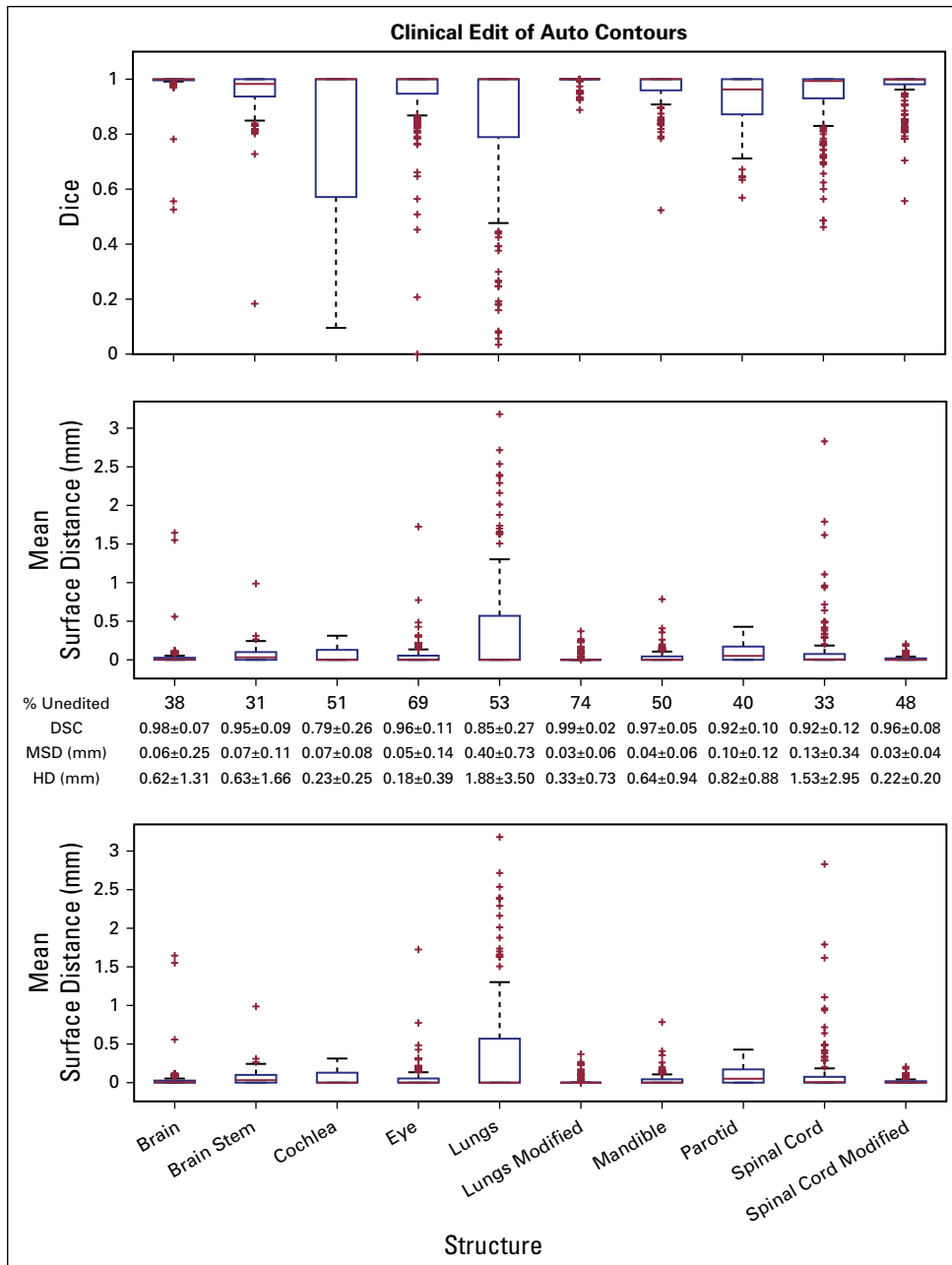


Fig 2. Distribution of the clinical edits to the autocontours. The red line within the box plot represents the median, and box edges represent the 25th and 75th percentile. Outliers are indicated by red crosses and are values outside the 25th or 75th percentile by more than 1.5 times the interquartile range. Between the box plots are the percent of unedited contours and the means (\pm standard deviation) of the Dice similarity coefficients and the mean surface distances for each for the automatically contoured or modified structures. DSC, Dice similarity coefficient; HD, Hausdorff distance; MSD, mean surface distance.

physicians who grouped the contour into needing either no or minor edits, which indicated that each physician may have slightly different criteria for determining when minor edits are needed. Furthermore, the discrepancy between ratings may have arisen because the atlas contours were based on contours drawn or approved by the primary physician. Considering the large variability

in physicians' contouring practices, these results are not surprising.

Upon clinical implementation, 95% of contours had a maximum distance to edit (HD) of 2 mm, indicating that for the vast majority of patients, only minor edits were required for use in treatment planning. Discussions with the attending physicians revealed a perceived time benefit

and, although the time needed to edit the autocontours was not an aim of this study, the study has established that the time required for editing an autocontour is less than that needed to generate the contour by hand.²⁸⁻³⁰ Since its clinical implementation, 22 attending physicians have used the automated contouring software, and nearly half the autocontours were not edited for use in the treatment plan. An investigation of the potential dosimetric impact of these clinical edits is ongoing.³¹ Used with or without edits, automated contouring solutions promise to be a pivotal part of the efforts to improve efficiency of RT planning in resource-constrained settings.

There are limitations for this analysis and more broadly for the potential utility of this approach. First, the contours used in the atlas were delineated by a single physician (although they were reviewed by a team of dedicated head-and-neck radiation oncologists), and the delineation of normal structures on the atlas patients will be propagated to test patients. Therefore, systematic differences in contouring practices between physicians may lead to edits. Although a single contouring atlas was used for this study, future work will determine whether additional center- or region-specific atlases may be required. Second, some physicians modify the normal structures away from base anatomy because of the need to cover adjacent areas with target dose. It is our opinion that true anatomy should be contoured accurately, with potential substructures created for planning (for instance, parotid planning target volume) and not to adjust the normal structure itself; however, some of our discrepancies are likely related to this. Data were not collected to determine the motivation for any clinical edit of autocontours.

Finally, the implementation of such a tool must be accompanied by proper training, cost assessment, and safety procedures.³² All autocontours must be thoroughly reviewed and approved by

the physician before planning treatment. Quality assurance of the treatment and accessory equipment, of the treatment planning process, of the delivery, and of the quality assurance process itself, must not be overlooked. Many guiding documents are available that discuss accepted quality assurance processes,³³⁻³⁵ and any change of practice must be accompanied by a reassessment of these guiding principles. Automatic contouring requires the watchful eye of a trained clinician and vigilance to ensure automatic contouring is used safely and effectively. This article describes our initial work at developing, validating, and integrating the automated contouring of normal structures for HNC cases into clinical practice. This approach promises to improve efficiency in resource-constrained settings, in which transition to intensity-modulated RT planning may be prohibitive because of the time constraints of contouring multiple normal structures. More importantly, it serves as a crucial part of the larger attempt to automate RT planning for low- and middle-income countries to improve availability, efficiency, and quality of care worldwide. The next step in this work is the development of automated contouring of nodal targets, which would require the physician to delineate only the gross tumor volume and select which nodal levels are at risk. This has the potential to reduce the extensive variability in the delineation of target volumes, as noted in many previous investigations.³⁶ Our results suggest that this approach is feasible, reproducible, and acceptable to an international group of physicians. Automation, including that of the contouring process, continues to be a promising avenue to relieve part of the staffing burden, time, and cost required for RT deployment in low-resource settings.

DOI: <https://doi.org/10.1200/JGO.18.00055>

Published online on jgo.org on August 15, 2018.

AUTHOR CONTRIBUTIONS

Conception and design: Rachel E. McCarroll, Beth M. Beadle, Peter A. Balter, David S. Followill, Kelly D. Kisling, Michael Mejia, Christopher L. Nelson, Lifei Zhang, Laurence E. Court, Jinzhong Yang

Financial support: Laurence E. Court

Administrative support: Laurence E. Court

Provision of study materials or patients: Peter A. Balter, Sameera Dalvie, Julie Wetter

Collection and assembly of data: Rachel E. McCarroll, Beth M. Beadle, Hester Burger, Komeela Naidoo, Karin Vorster, Julie Wetter

Data analysis and interpretation: Rachel E. McCarroll, Beth M. Beadle, Carlos E. Cardenas, Sameera Dalvie, Kelly D. Kisling, Michael Mejia, Christine B. Peterson, Laurence E. Court

Manuscript writing: All authors

Final approval of manuscript: All authors

Accountable for all aspects of the work: All authors

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/jco/site/ifc.

Rachel E. McCarroll

No relationship to disclose

Beth M. Beadle

No relationship to disclose

Peter A. Balter

Speakers' Bureau: Varian Medical Systems

Research Funding: Varian Medical Systems, Sun Nuclear (Inst), RaySearch Laboratories

Travel, Accommodations, Expenses: Varian Medical Systems, Sun Nuclear

Hester Burger

No relationship to disclose

Carlos E. Cardenas

No relationship to disclose

Sameera Dalvie

No relationship to disclose

David S. Followill

Honoraria: Varian Medical Systems

Travel, Accommodations, Expenses: Varian Medical Systems

Kelly D. Kisling

No relationship to disclose

Michael Mejia

Honoraria: Innogene Kalbiotech

Komeela Naidoo

No relationship to disclose

Christopher L. Nelson

No relationship to disclose

Christine B. Peterson

No relationship to disclose

Julie Wetter

Travel, Accommodations, Expenses: Genentech

Lifei Zhang

No relationship to disclose

Laurence E. Court

No relationship to disclose

Jinzhong Yang

Research Funding: Elekta

Patents, Royalties, Other Intellectual Property: Royalties from Varian Medical Systems for licensing of the deformable image registration technique

ACKNOWLEDGMENT

We thank Vikash Sewram for his contributions to this project.

Affiliations

Rachel E. McCarroll, Peter A. Balter, Carlos E. Cardenas, David S. Followill, Kelly D. Kisling, Christopher L. Nelson, Christine B. Peterson, Lifei Zhang, Laurence E. Court, and Jinzhong Yang, The University of Texas MD Anderson Cancer Center, Houston, TX; **Beth M. Beadle,** Stanford University, Stanford, CA; **Hester Burger, Sameera Dalvie, and Julie Wetter,** Groote Schuur Hospital and University of Cape Town; **Komeela Naidoo,** Stellenbosch University and Tygerberg Hospital, Cape Town; **Karin Vorster,** University of the Free State, Bloemfontein, South Africa; and **Michael Mejia,** University of Santo Tomas Hospital, Benavides Cancer Institute, Manila, Philippines.

Support

Supported by Grant No. UH2-CA202665 from the National Cancer Institute, by Varian Medical Systems, and by Mobius Medical Systems.

REFERENCES

1. Datta NR, Samiei M, Bodis S: Radiation therapy infrastructure and human resources in low- and middle-income countries: Present status and projections for 2020. *Int J Radiat Oncol Biol Phys* 89:448-457
2. Zubizarreta EH, Fidarova E, Healy B, et al: Need for radiotherapy in low and middle income countries: The silent crisis continues. *Clin Oncol (R Coll Radiol)* 27:107-114, 2015
3. Efstathiou JA, Heunis M, Karumekayi T, et al: Establishing and delivering quality radiation therapy in resource-constrained settings: The story of Botswana. *J Clin Oncol* 34:27-35, 2016
4. Pirzkall A, Carol M, Lohr F, et al: Comparison of intensity-modulated radiotherapy with conventional conformal radiotherapy for complex-shaped tumors. *Int J Radiat Oncol Biol Phys* 48:1371-1380
5. Gupta T, Agarwal J, Jain S, et al: Three-dimensional conformal radiotherapy (3D-CRT) versus intensity modulated radiation therapy (IMRT) in squamous cell carcinoma of the head and neck: A randomized controlled trial. *Radiother Oncol* 104:343-348, 2012

6. Nutting CM, Morden JP, Harrington KJ, et al: Parotid-sparing intensity modulated versus conventional radiotherapy in head and neck cancer (PARSPORT): A phase 3 multicentre randomised controlled trial. *Lancet Oncol* 12:127-136, 2011
7. Kam MK, Leung SF, Zee B, et al: Prospective randomized study of intensity-modulated radiotherapy on salivary gland function in early-stage nasopharyngeal carcinoma patients. *J Clin Oncol* 25:4873-4879, 2007
8. Horan G, Roques TW, Curtin J, et al: "Two are better than one": A pilot study of how radiologist and oncologists can collaborate in target volume definition. *Cancer Imaging* 6:16-19, 2006
9. Jeanneret-Sozzi W, Moeckli R, Valley JF, et al: The reasons for discrepancies in target volume delineation: A SASRO study on head-and-neck and prostate cancers. *Strahlenther Onkol* 182:450-457, 2006
10. Brouwer CL, Steenbakkers RJ, van den Heuvel E, et al: 3D variation in delineation of head and neck organs at risk. *Radiat Oncol* 7:32, 2012
11. Nelms BE, Tomé 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* 82:368-378, 2012
12. Bortfeld T, Jeraj R: The physical basis and future of radiation therapy. *Br J Radiol* 84:485-498, 2011
13. Jameson MG, Holloway LC, Vial PJ, et al: A review of methods of analysis in contouring studies for radiation oncology. *J Med Imaging Radiat Oncol* 54:401-410, 2010
14. Njeh CF: Tumor delineation: The weakest link in the search for accuracy in radiotherapy. *J Med Phys* 33:136-140, 2008
15. Brouwer CL, Steenbakkers RJ, Bourhis J, et al: CT-based delineation of organs at risk in the head and neck region: DAHANCA, EORTC, GORTEC, HKNPCSG, NCIC CTG, NCRI, NRG Oncology and TROG consensus guidelines. *Radiother Oncol* 117:83-90, 2015
16. Hu K, Lin A, Young A, et al: Timesavings for contour generation in head and neck IMRT: Multi-institutional experience with an atlas-based segmentation method. *Int J Radiat Oncol Biol Phys* 72:S391, 2008
17. Stapleford LJ, Lawson JD, Perkins C, et al: Evaluation of automatic atlas-based lymph node segmentation for head-and-neck cancer. *Int J Radiat Oncol Biol Phys* 77:959-966
18. Gambacorta MA, Valentini C, Dinapoli N, et al: Clinical validation of atlas-based auto-segmentation of pelvic volumes and normal tissue in rectal tumors using auto-segmentation computed system. *Acta Oncol* 52:1676-1681, 2013
19. Yang J, Amini A, Williamson R, et al: Automatic contouring of brachial plexus using a multi-atlas approach for lung cancer radiotherapy. *Pract Radiat Oncol* 3(4), 2013
20. Zhou R, Liao Z, Pan T, et al: Cardiac atlas development and validation for automatic segmentation of cardiac substructures. *Radiother Oncol* 122:66-71, 2017
21. Gambacorta MA, Boldrini L, Valentini C, et al: Automatic segmentation software in locally advanced rectal cancer: READY (REsearch program in Auto Delineation sYstem)-RECTAL 02—Prospective study. *Oncotarget* 7:42579-42584, 2016
22. Wang H, Dong L, O'Daniel J, et al: Validation of an accelerated 'demons' algorithm for deformable image registration in radiation therapy. *Phys Med Biol* 50:2887-2905, 2005
23. Warfield SK, Zou KH, Wells WM: Simultaneous truth and performance level estimation (STAPLE): An algorithm for the validation of image segmentation. *IEEE Trans Med Imaging* 23:903-921, 2004
24. Wang H, Dong L, Lii MF, et al: Implementation and validation of a three-dimensional deformable registration algorithm for targeted prostate cancer radiotherapy. *Int J Radiat Oncol Biol Phys* 61:725-735, 2005
25. Yang J, Zhang Y, et al: .Medical image analysis for the clinic: A grand challenge, in van Ginneken, Murphy K, Heimann T, et al (eds): *Workshop Proceedings from the 13th International Conference on Medical Image Computing and Computer Assisted Intervention*, Beijing, China, on September 20-24, 2010. pp 273-323

26. Yang J, Garden A, Zhang Y, et al: WE-E-213CD-09: Multi-atlas fusion using a tissue appearance model. *Med Phys* 39:3961, 2012
27. Cardenas CE, Mohamed ASR, Tao R, et al: Prospective qualitative and quantitative analysis of real-time peer review quality assurance rounds incorporating direct physical examination for head and neck cancer radiation therapy. *Int J Radiat Oncol Biol Phys* 98:532-540, 2017
28. La Macchia M, Fellin F, Amichetti M, et al: Systematic evaluation of three different commercial software solutions for automatic segmentation for adaptive therapy in head-and-neck, prostate and pleural cancer. *Radiat Oncol* 7:160, 2012
29. Daisne JF, Blumhofer A: Atlas-based automatic segmentation of head and neck organs at risk and nodal target volumes: A clinical validation. *Radiat Oncol* 8:154, 2013
30. Thomson D, Boylan C, Liptrot T, et al: Evaluation of an automatic segmentation algorithm for definition of head and neck organs at risk. *Radiat Oncol* 9:173, 2014
31. Delaney AR, Dahele M, Slotman BJ, et al: Is accurate contouring of salivary and swallowing structures necessary to spare them in head and neck VMAT plans? *Radiother Oncol* 127:190-196, 2018
32. Van Dyk J, Zubizarreta E, Lievens Y: Cost evaluation to optimise radiation therapy implementation in different income settings: A time-driven activity-based analysis. *Radiother Oncol* 125:178-185, 2017
33. Podgorsak EB (ed): *Radiation Oncology Physics: A Handbook for Teachers and Students*. Vienna, Austria, International Atomic Energy Agency, 2005, pp 407-
34. Kutcher GJ, Coia L, Gillin M, et al: Comprehensive QA for radiation oncology: Report of AAPM Radiation Therapy Committee Task Group 40. *Med Phys* 21:581-618, 1994
35. Thwaites D, Scalliet P, Leer JW, et al: Quality assurance in radiotherapy: European Society for Therapeutic Radiology and Oncology Advisory Report to the Commission of the European Union for the 'Europe Against Cancer Programme'. *Radiother Oncol* 35:61-73, 1995
36. Rasch C, Steenbakkers R, van Herk M: Target definition in prostate, head, and neck. *Semin Radiat Oncol* 15:136-145, 2005