Radiotherapy and Oncology xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Original article

Prospective randomized double-blind study of atlas-based organ-at-risk autosegmentation-assisted radiation planning in head and neck cancer ☆

Gary V. Walker [a,b], Musaddiq Awan [a], Randa Tao [a], Eugene J. Koay [a], Nicholas S. Boehling [a], Jonathan D. Grant [a], Dean F. Sittig [b], Gary Brandon Gunn [a], Adam S. Garden [a], Jack Phan [a], William H. Morrison [a], David I. Rosenthal [a], Abdallah Sherif Radwan Mohamed [a], Clifton David Fuller [a,c,*]

[a] Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center; [b] School of Biomedical Informatics, University of Texas Health Science Center; and [c] University of Texas Graduate School of Biomedical Sciences, Houston, USA

ARTICLE INFO

Article history:
Received 13 December 2013
Received in revised form 7 August 2014
Accepted 12 August 2014
Available online xxxx

Keywords:
Atlas-based autosegmentation
Normal tissue
Autocontouring
Head and neck
Automatic segmentation
Organs-at-risk

ABSTRACT

Background and purpose: Target volumes and organs-at-risk (OARs) for radiotherapy (RT) planning are manually defined, which is a tedious and inaccurate process. We sought to assess the feasibility, time reduction, and acceptability of an atlas-based autosegmentation (AS) compared to manual segmentation (MS) of OARs.

Materials and methods: A commercial platform generated 16 OARs. Resident physicians were randomly assigned to modify AS OAR (AS + R) or to draw MS OAR followed by attending physician correction. Dice similarity coefficient (DSC) was used to measure overlap between groups compared with attending approved OARs (DSC = 1 means perfect overlap). 40 cases were segmented.

Results: Mean ± SD segmentation time in the AS + R group was 19.7 ± 8.0 min, compared to 28.5 ± 8.0 min in the MS cohort, amounting to a 30.9% time reduction (Wilcoxon $p < 0.01$). For each OAR, AS DSC was statistically different from both AS + R and MS ROIs (all Steel–Dwass $p < 0.01$) except the spinal cord and the mandible, suggesting oversight of AS/MS processes is required; AS + R and MS DSCs were non-different. AS compared to attending approved OAR DSCs varied considerably, with a chiasm mean ± SD DSC of 0.37 ± 0.32 and brainstem of 0.97 ± 0.03.

Conclusions: Autosegmentation provides a time savings in head and neck regions of interest generation. However, attending physician approval remains vital.

The advent of conformal radiotherapy (RT), as well as the developing paradigm of image-guided RT, affords delivery of tumoricidal radiation doses to user-defined target volumes while minimizing dose to organs-at-risk (OARs). However, since the optimization software responsible for dose-reduction to OARs requires accurate region of interest (ROI) segmentation for all subsequent dose calculations, accurate ROI definition is recognized as of paramount importance for RT planning, representing perhaps contributing to the most uncertainty in RT planning.

Target volumes and OARs for RT planning are typically manually defined by human users, which is a tedious and time consuming process [1]. Manual segmentation (MS), which uses a digital paintbrush controlled by the user to outline the ROI, remains highly variable, particularly for novices [2]. Inexperienced trainees or those who treat particular oncologic sites infrequently may under or over-contour particular OAR regions-of-interest (ROIs), especially those that are difficult to visualize on computed tomography (CT). Atlas-based autosegmentation (AS) algorithms have been shown to accurately delineate OARs ROIs for a variety of disease sites and early results show a time saving advantage [3–10].

Head and neck cancers present a unique set of challenges in terms of target delineation for conformal RT [11,12]. With data from Fuller et al. suggesting that head and neck cancer target volumes [13] are comparatively difficult to contour, time-consuming, and have higher inter-observer variability than other anatomic sites, we sought to ascertain whether time-savings might be achieved using OAR AS as part of our clinical workflow. To this end we constructed a randomized, blinded prospective *in silico* study with the following specific aims:

☆ This work was presented in abstract form at the 55th annual meeting of the American Society for Radiation Oncology on September 23, 2013 in Atlanta, GA.

* Corresponding author at: Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd., Unit 97, Houston, TX 77030, USA.

E-mail address: cdfuller@mdanderson.org (C.D. Fuller).

http://dx.doi.org/10.1016/j.radonc.2014.08.028
0167-8140/

(1) Determine the real-time workflow feasibility and capacity for delineation time-reduction using AS software for head and neck OARs in a representative clinical population.
(2) Evaluate individual OAR acceptability of AS/AS-assisted definition of head and neck ROIs using resident and expert physician comparators.

Materials and methods

This prospective, randomized, double-blind study was approved by the Institutional Review Board. Patients being simulated for definitive RT ± chemo for head and neck malignancies were included. Patients were excluded for the following reasons: (1) under the age of 18 years, (2) prior history of RT or surgery to the head and neck, and (3) cutaneous malignancies. These patients were excluded so as to reduce bias as these patients could have anatomical variations. Patients were simulated supine with chins extended utilizing a custom thermoplastic mask for immobilization. Non-contrast axial CT slices were obtained using 2.5–3.0 mm slices. All patients were treated using intensity modulated RT (IMRT) with either the whole-field IMRT technique or the half-beam technique with the upper IMRT field matched to a static low neck field as previously described [14].

Study design

Following CT acquisition, all DICOM files were processed using a commercial, proprietary AS platform (Pinnacle 9.4, SPICE AS algorithm, Philips Healthcare, Andover, MA, USA). This software package performs an initial registration, dense deformable registration and then probabilistic refinement using an automated platform in the background while the user performs other tasks. Residents were randomly assigned to manually modify AS OARs ROIs (AS + R) or to draw MS OARs ROIs *de novo* using a pairwise randomization technique so that each resident did equal number of AS + R and MS cases. These 16 OARs included the spinal cord, brainstem, optic chiasm, mandible, oral cavity, soft palate, larynx, pharyngeal constrictors as well as bilateral optic nerves, parotid glands, submandibular glands and cochlea. For the purposes of this study we did not evaluate gross or clinical tumor volumes as these display significant variation between patients based on clinical presentation necessitating MS followed by multi-physician examination and real-time ROI quality assurance, using a methodology described previously [12,15]. Total resident segmentation/correction time was recorded. Residents rotate through the head and neck service during the first, second and in either their third or fourth year of training. Attending physicians subsequently reviewed all OARs and manually corrected them as necessary, blinded to AS or MS ROIs priors. All OARs underwent established processes including approval by the attending physician and the head and neck service quality assurance team consisting of multiple attending physicians to minimize inter-observer variation in OAR delineation [12,15]. ROIs were never used for treatment planning without attending approval.

Endpoints and analysis

Specified outcome variables included overall segmentation time (including AS processing time) and segmentation accuracy (using attending physician approved contours as a reference). For this study, nonparametric statistical comparison was implemented using the paired Wilcoxon rank sum test for between-group comparison of ordinal and scalar variables (Supplement 1). Based on these calculations, a maximal enrollment of 106 patients (53 per arm) was approved by the institutional review board, with an interim analysis conducted when enrollment met 40 patients. As the Wilcoxon rank-sum test reached statistical significance, enrollment was stopped early as per protocol specifications.

For ROI accuracy assessment, Dice similarity coefficient (DSC) was used to measure overlap between the AS, AS with resident correction (AS + R), and MS compared with attending approved ROIs, where DSC = 1 means perfect ROI overlap, and DSC = 0 is totally discordant. An informal *a priori* threshold DSC of $\geqslant 0.85$ was utilized as an acceptability threshold (e.g., ROIs with 85% DSC agreement were considered to require clinically acceptable modification, whereas >15% disagreement was considered clinically unacceptable within our current workflow). A paired Kruskal–Wallis test was performed using the Steel–Dwass methodology (the non-parametric analog comparable to Tukey's range test) was used to compare the total process time between the groups stratified by OAR ROI. Univariate descriptive statistics among the groups were compared using a Pearson χ^2 test. A non-Bonferroni-corrected alpha of 0.05 was specified for statistical significance for all analyses. Data analysis was performed using Stata/SE 12.0 (Stata Corporation, College Station, TX, USA) and JMP v10/SAS v9.2 (both by SAS Institute, Cary, NC, USA) statistical software.

Results

Baseline demographic and tumor characteristics were similar between the 40 cases undergoing MS and AS (Supplemental Table 1). The median age was 59.8 with a mean body mass index (BMI) of 29.2. The oropharynx was the most common site of disease (65.0%), followed by larynx (12.5%) with a range of TNM stages (Table 1).

All 40 cases were segmented by one of 8 residents and approved by one of 7 head and neck section expert attending physicians. Mean ± standard deviation (SD) resident segmentation time in the AS + R group was 19.7 ± 8.0 min, compared to 28.5 ± 8.0 min in the MS cohort, amounting to a 30.9% time reduction (Wilcoxon $p = 0.002$) (Table 1). There was no difference in the number of cases done by residents in each year of training (p = n.s.). Average time savings per resident ranged from 2.3–15.7 min.

Fig. 1 displays an example of the AS and attending physician approved OAR ROIs. The mean ± SD DSC for the uncorrected AS, AS + R, and MS ROIs compared to final attending approved ROIs was 0.74 ± 0.29, 0.96 ± 0.10, and 0.95 ± 0.16, respectively (Table 2, Fig. 2). For each OAR, AS DSC was statistically different from both AS + R and MS ROIs (all Steel–Dwass $p < 0.01$) except the spinal cord ($p = 0.4$) and the mandible (0.4); AS + R and MS DSCs were non-different (all p = n.s.) suggesting significant oversight of AS/MS processes is still required.

AS accuracy compared to attending approved OAR DSCs varied considerably with the following structures showing the most overlap (i.e., most accurate): mandible (DSC 0.98 ± 0.2), brainstem (0.97 ± 0.03), pharyngeal constrictors (0.93 ± 0.09), spinal cord (0.90 ± 0.14), parotid glands (0.89 ± 0.11), and soft palate (0.86 ± 0.15). However, specific structures showed reduced DSC values, representative of poor agreement with expert ROIs. Notably, the optic chiasm (0.37 ± 0.32), oral cavity (0.53 ± 0.24), cochlea (0.56 ± 0.38), larynx (0.65 ± 0.26), optic nerves (0.71 ± 0.26), and submandibular glands (0.73 ± 0.25) were notable for comparatively suboptimal segmentation, with a statistically different performance across OAR ROIs (Steel–Dwass $p < 0.01$).

Discussion

These data represent, to our knowledge, the largest prospective randomized controlled study evaluating the ROI accuracy and time-efficiency gains of OAR AS software in head and neck RT.

Please cite this article in press as: Walker GV et al. Prospective randomized double-blind study of atlas-based organ-at-risk autosegmentation-assisted radiation planning in head and neck cancer. Radiother Oncol (2014), http://dx.doi.org/10.1016/j.radonc.2014.08.028

Table 1
Average segmentation time by resident stratified by randomization to atlas-based autosegmentation or manual segmentation ($n = 40$).

	Number of cases	Resident year	AS group Average (min)	MS group Average (min)	Difference (min)	p
Resident 1	6	1	32.7	35.0	2.3	
Resident 2	6	1	19.7	35.3	15.7	
Resident 3	12	2	15.4	25.3	9.9	
Resident 4	4	2	22.0	37.3	15.3	
Resident 5	6	3 or 4	20.3	22.5	2.2	
Resident 6	2	3 or 4	16.1	22.0	5.9	
Resident 7	2	3 or 4	15.4	18.6	3.2	
Resident 8	2	3 or 4	8.0	20.0	12.0	
Average	40		19.7	28.5	8.8	0.002
SD			8.0	8.0		

Abbreviations: AS, atlas-based autosegmentation; MS, manual segmentation; CI, confidence interval.

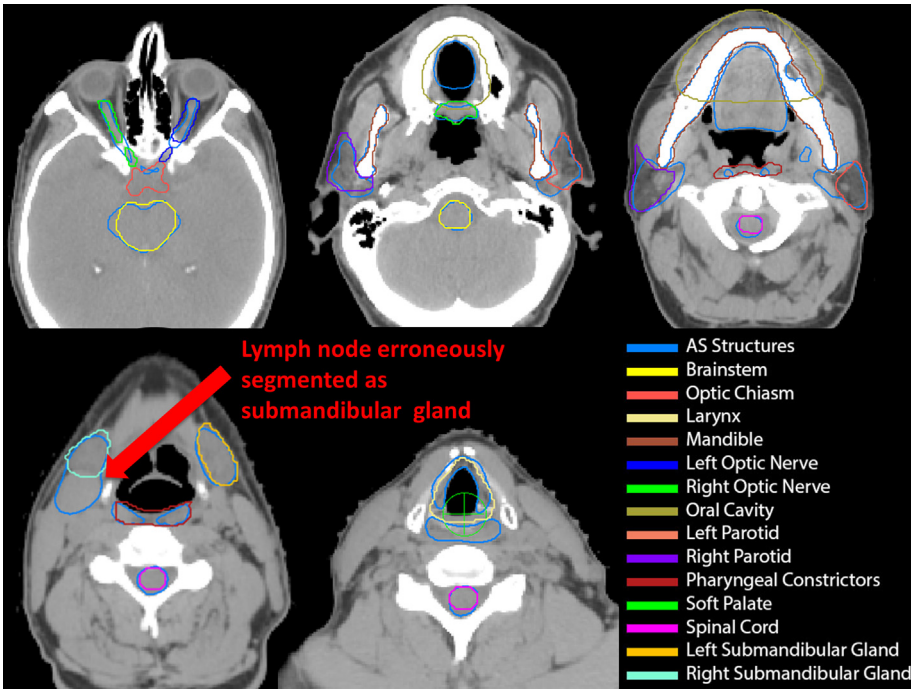

Fig. 1. Illustrative example of the AS and attending physician approved OAR ROIs.

Table 2
Mean dice similarity coefficient (DSC) values by organ-at risk compared to attending approved structures.

	DSC ± standard deviation*		
	AS	AS + R	MS
Brainstem	0.97 ± 0.03	0.99 ± 0.01	0.95 ± 0.14
Spinal cord	0.90 ± 0.14	0.97 ± 0.06	0.90 ± 0.18
Optic nerves	0.71 ± 0.26	0.77 ± 0.24	0.77 ± 0.24
Optic chiasm	0.37 ± 0.32	0.93 ± 0.17	0.96 ± 0.14
Mandible	0.98 ± 0.2	0.99 ± 0.01	0.98 ± 0.05
Parotid glands	0.89 ± 0.11	0.90 ± 0.11	0.90 ± 0.11
Submandibular glands	0.73 ± 0.25	0.76 ± 0.26	0.77 ± 0.26
Oral cavity	0.53 ± 0.24	0.87 ± 0.24	0.92 ± 0.18
Soft palate	0.86 ± 0.15	0.96 ± 0.08	0.93 ± 0.14
Cochleas	0.56 ± 0.38	0.61 ± 0.39	0.60 ± 0.40
Larynx	0.65 ± 0.26	0.94 ± 0.13	0.94 ± 0.15
Pharyngeal constrictors	0.93 ± 0.09	0.98 ± 0.04	0.91 ± 0.17

Abbreviations: AS = uncorrected autosegmented (AS) regions of interest (ROIs), AS + R = resident corrected AS ROIs, MS = manually segmented (ROIs).
* Dice similarity coefficient (DSC) compared to attending approved.

We found that AS priors resulted in a 30.9% savings in head and neck over manual OARs delineation. AS algorithm-generated structures were similar to resident MS for several structures. However, attending physician approval for all OARs remains vital in order to assure proper normal tissue avoidance during the treatment planning process, as AS performed poorly for other OAR ROIs.

The goals of AS are to reduce segmentation time, enhance the accuracy of segmentation, as well as reduce inter- and intra-user variability [4,11]. These advantages will be essential as the field moves toward increased adoption of adaptive RT, which allows radiation plans to be adapted during the course of treatment based on clinical and radiographic tumor response and may result in better normal tissue sparing over the course of treatment and less toxicity [16]. Initial studies have validated the utility AS for OAR identification in adaptive planning of head and neck [17] and prostate cases [18]. Additionally, with the widespread adoption of IMRT, clinicians have to consider toxicities to structures not previously appreciated with 3D-conformal techniques [19]. Accurately segmenting these OAR ROIs is critical in order to adequately spare them. There has also been shown to be significant variability between clinicians in contouring target volumes and OARs, which makes RT planning using uniform normal tissue dose constraints more challenging [20].

Please cite this article in press as: Walker GV et al. Prospective randomized double-blind study of atlas-based organ-at-risk autosegmentation-assisted radiation planning in head and neck cancer. Radiother Oncol (2014), http://dx.doi.org/10.1016/j.radonc.2014.08.028

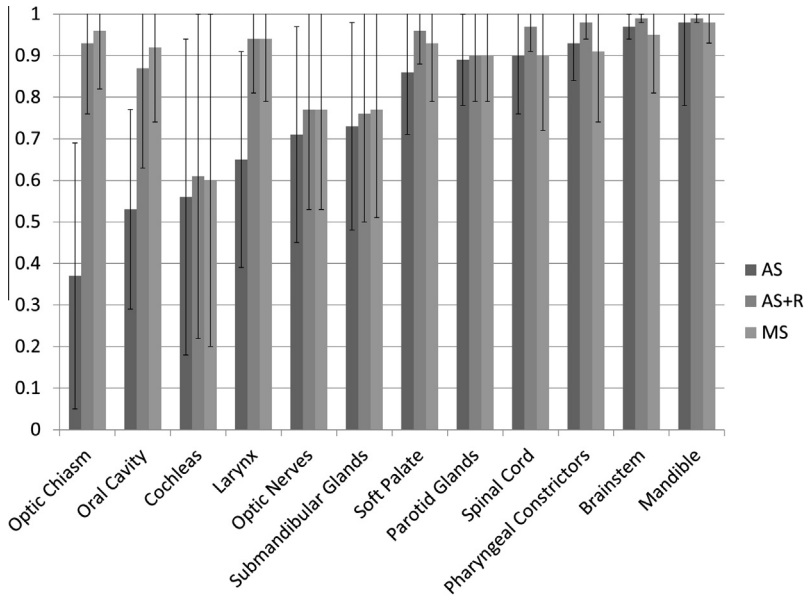

Fig. 2. Graphical presentation of Dice similarity coefficient (DSC) values by organ-at risk compared to attending approved structures.

Before implementing commercial AS tools, potential relative time savings and software accuracy need to be rigorously investigated, as plan optimization (and thus OAR sparing and target coverage) is predicated on accurate ROI segmentation [21,22]. AS algorithms have been shown to be quite reliable in appropriately delineating OARs [5]. Teguh et al. compared MS OARs and target volumes to AS structures using a commercial atlas-based AS software [8]. An expert panel subjectively scored all OARs in the 12 patients as, "minor deviation, editable" or better with excellent agreement based on DSC. They found AS decreased the total segmentation time of 20 OARs and neck levels from 180 min to 73 min in head and neck cases. Similarly, La Macchia et al. evaluated three commercially available AS tools to determine their clinical utility in delineating OARs in the head and neck, pelvis and thorax. They concluded that there was significant time savings with AS (approximately 60 min for head and neck cases) [3]. They found better DSC using MIM 5.1.1, (MIMVista corp, Cleveland, Ohio) compared to ABAS 2.0 (CMS-Elekta, Stockholm, Sweden) or VelocityAI 2.6.2 (Velocity Medical Systems, Atlanta, Georgia).

While AS algorithms are promising for workflow efficiency improvement, attending physician oversight remains critical. As Voet et al. conducted an analysis using AS to define nodal target volumes as well as OARs which were then used unedited for planning [23], resulting in under dosage of target volumes, we are as yet conceptually unwilling to accept unmodified AS-defined GTV/CTV ROIs, including elective neck contours generated as AS-contoured nodal basins. Our data suggest clinically unacceptable AS segmentation for several critical OAR structures (e.g., chiasm, cochlea, and larynx), inadvertent overdosage of which might result in blindness [24], hearing loss [25], or aspiration/dysphagia [26].

Additionally, it must be carefully stressed that the criticality of ROI segmentation remains, at its most fundamental, the primary driver of subsequent planning. For instance, in Fig. 1, a pathologic lymph node in level II was segmented erroneously as part of the adjacent submandibular gland. Had this not been detected and addressed, planning constraints might have been compromised. For this reason, even as AS serves to augment physician efficacy in a useful and demonstrable manner, continued clinical vigilance in CTV and OAR delineation remains imperative in order to ensure patient safety.

While randomization accounts for much variance, there remain notable limitations to the present data. The conclusions drawn from this study apply only to the specific version of the AS algorithm used in a head and neck clinic, and must be replicated for other software systems/versions in other anatomic sites. In addition, these results only apply to patients being treated with definitive intent, and not patients being treated with post-operative RT. Competing commercial or research algorithms could have markedly different results in regard to time-savings and/or acceptable ROI segmentation accuracy. However, we anticipate that these results would be similar using other algorithms and in other clinical settings. In many departments, OAR ROI generation is done by technicians, dosimetrists or trainees, which are then reviewed by attending physicians. We would anticipate similar time savings in these clinical situations. While one resident segmented a large number of cases, excluding this individual did not alter the significant time savings ($p = 0.02$). In addition, observer bias could alter efficiency estimates, as, while attending physicians were blinded to AS vs. MS priors, residents were necessarily alerted to the presence of existing AS structure sets. We chose the DICE index as it assigns double value to the overlap area, thus reducing the influence of one outlier value [7]. It is important to note that this can provide a false impression of high agreement.

In summary, atlas-based AS priors provided a detectable physician time savings in head and neck OAR ROI generation over MS, performing comparably to resident MS for many structures. However, attending physician approval for all OARs remains vital, as several key structures including the optic chiasm and optic nerves were poorly segmented by the AS algorithm. Future studies are warranted to compare various AS platforms head-to-head.

Conflicts of Interest Notification

Dr. Walker is supported by a training fellowship from the Keck Center for Interdisciplinary Bioscience Training of the Gulf Coast Consortia (Grant No. T15 LM007093) and from a 2013 Concur Cancer Foundation of the American Society of Clinical Oncology Merit Award. Dr. Fuller received/receives grant/funding support from the SWOG/Hope Foundation Dr. Charles A. Coltman, Jr., Fellowship in Clinical Trials, the National Institutes of Health Clinician Scientist Loan Repayment Program (L30 CA136381-02), Elekta AB (Stockholm, SE), the Center for Radiation Oncology Research at MD Anderson Cancer Center, and the MD Anderson Institutional Research Grant Program.

In-kind support was provided as pre-release version of Pinnacle3 V9.4 by Koninklijke Philips N.V. (Eindhoven, NL), under a master research agreement (MRA) with the University of Texas MD Anderson Cancer Center.

These listed funders/supporters played no role in the study design, collection, analysis, interpretation of data, manuscript writing, or decision to submit the report for publication.

The authors wish to gratefully thank Michelle M. Kim, M.D., Anna Likhacheva, M.D. M.P.H., Shervin M. Shirvani, M.D., M.P.H. for their segmentation assistance, John A. Garcia, Matthew K. Holmes, Kristen M. Houtman, Kelli K. McSpadden, Manuel A. Oyervides, Shirly J. Kuruvila and Jenny Nuanjing for their dosimetry support, as well as Michael E. Kantor and Jared D. Ohrt for their technical expertise.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.radonc.2014.08.028.

References

[1] Austin-Seymour M, Chen GT, Rosenman J, Michalski J, Lindsley K, Goitein M. Tumor and target delineation: current research and future challenges. Int J Radiat Oncol Biol Phys 1995;33:1041–52.
[2] Bekelman JE, Wolden S, Lee N. Head-and-neck target delineation among radiation oncology residents after a teaching intervention: a prospective, blinded pilot study. Int J Radiat Oncol Biol Phys 2009;73:416–23.
[3] La Macchia M, Fellin F, Amichetti M, Cianchetti M, Gianolini S, Paola V, 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 2012;7:160.
[4] Chaney EL, Pizer SM. Autosegmentation of images in radiation oncology. J Am Coll Radiol 2009;6:455–8.
[5] Sims R, Isambert A, Gregoire V, Bidault F, Fresco L, Sage J, et al. A pre-clinical assessment of an atlas-based automatic segmentation tool for the head and neck. Radiother Oncol 2009;93:474–8.
[6] Anders LC, Stieler F, Siebenlist K, Schafer J, Lohr F, Wenz F. Performance of an atlas-based autosegmentation software for delineation of target volumes for radiotherapy of breast and anorectal cancer. Radiother Oncol 2012;102:68–73.
[7] Fotina I, Lutgendorf-Caucig C, Stock M, Potter R, Georg D. Critical discussion of evaluation parameters for inter-observer variability in target definition for radiation therapy. Strahlenther Onkol 2012;188:160–7.
[8] Teguh DN, Levendag PC, Voet PW, Al-Mamgani A, Han X, Wolf TK, et al. Clinical validation of atlas-based auto-segmentation of multiple target volumes and normal tissue (swallowing/mastication) structures in the head and neck. Int J Radiat Oncol Biol Phys 2011;81:950–7.
[9] Simmat I, Georg P, Georg D, Birkfellner W, Goldner G, Stock M. Assessment of accuracy and efficiency of atlas-based autosegmentation for prostate radiotherapy in a variety of clinical conditions. Strahlenther Onkol 2012;188:807–15.
[10] Daisne JF, Blumhofer A. Atlas-based automatic segmentation of head and neck organs at risk and nodal target volumes: a clinical validation. Radiat Oncol 2013;8:154.
[11] Rasch C, Steenbakkers R, van Herk M. Target definition in prostate, head, and neck. Semin Radiat Oncol 2005;15:136–45.
[12] Tao R, Fuller CD, Gunn GB, Beadle BM, Phan J, Frank SJ, et al. Real-time peer review quality assurance conferences incorporating physical examination for head-and-neck cancer radiation therapy result in clinically meaningful target volume alteration: results of a prospective volumetric analysis. Int J Radiat Oncol Biol Phys 2012;84:S151.
[13] Multi-Institutional Target Delineation in Oncology Group. Human–computer interaction in radiotherapy target volume delineation: a prospective, multi-institutional comparison of user input devices. J Digital Imag 2011;24:794–803.
[14] Dabaja B, Salehpour MR, Rosen I, Tung S, Morrison WH, Ang KK, 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;63:1000–5.
[15] Rosenthal DI, Asper JA, Barker Jr JL, Garden AS, Chao KS, Morrison WH, et al. Importance of patient examination to clinical quality assurance in head and neck radiation oncology. Head Neck 2006;28:967–73.
[16] Gregoire V, Jeraj R, Lee JA, O'Sullivan B. Radiotherapy for head and neck tumours in 2012 and beyond: conformal, tailored, and adaptive? Lancet Oncol 2012;13:e292–300.
[17] Hardcastle N, Tome WA, Cannon DM, Brouwer CL, Wittendorp PW, Dogan N, et al. A multi-institution evaluation of deformable image registration algorithms for automatic organ delineation in adaptive head and neck radiotherapy. Radiat Oncol 2012;7:90.
[18] Wang H, Dong L, Lii MF, Lee AL, de Crevoisier R, Mohan R, et al. Implementation and validation of a three-dimensional deformable registration algorithm for targeted prostate cancer radiotherapy. Int J Radiat Oncol Biol Phys 2005;61:725–35.
[19] Rosenthal DI, Chambers MS, Fuller CD, Rebueno NC, Garcia J, Kies MS, et al. Beam path toxicities to non-target structures during intensity-modulated radiation therapy for head and neck cancer. Int J Radiat Oncol Biol Phys 2008;72:747–55.
[20] Mukesh M, Benson R, Jena R, Hoole A, Roques T, Scrase C, et al. Interobserver variation in clinical target volume and organs at risk segmentation in post-parotidectomy radiotherapy: can segmentation protocols help? Br J Radiol 2012;85:e530–6.
[21] Weiss E, Hess CF. The impact of gross tumor volume (GTV) and clinical target volume (CTV) definition on the total accuracy in radiotherapy theoretical aspects and practical experiences. Strahlenther Onkol 2003;179:21–30.
[22] Jeanneret-Sozzi W, Moeckli R, Valley JF, Zouhair A, Ozsahin EM, Mirimanoff RO, et al. The reasons for discrepancies in target volume delineation: a SASRO study on head-and-neck and prostate cancers. Strahlenther Onkol 2006;182:450–7.
[23] Voet PW, Dirkx ML, Teguh DN, Hoogeman MS, Levendag PC, Heijmen BJ. Does atlas-based autosegmentation of neck levels require subsequent manual contour editing to avoid risk of severe target underdosage? A dosimetric analysis. Radiother Oncol 2011;98:373–7.
[24] Mayo C, Martel MK, Marks LB, Flickinger J, Nam J, Kirkpatrick J. Radiation dose-volume effects of optic nerves and chiasm. Int J Radiat Oncol Biol Phys 2010;76:S28–35.
[25] Bhandare N, Jackson A, Eisbruch A, Pan CC, Flickinger JC, Antonelli P, et al. Radiation therapy and hearing loss. Int J Radiat Oncol Biol Phys 2010;76:S50–7.
[26] Rancati T, Schwarz M, Allen AM, Feng F, Popovtzer A, Mittal B, et al. Radiation dose–volume effects in the larynx and pharynx. Int J Radiat Oncol Biol Phys 2010;76:S64–9.

Please cite this article in press as: Walker GV et al. Prospective randomized double-blind study of atlas-based organ-at-risk autosegmentation-assisted radiation planning in head and neck cancer. Radiother Oncol (2014), http://dx.doi.org/10.1016/j.radonc.2014.08.028