



Contents lists available at ScienceDirect

## Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com



## Original article

## A pre-clinical assessment of an atlas-based automatic segmentation tool for the head and neck

Richard Sims<sup>a,c,\*</sup>, Aurelie Isambert<sup>b</sup>, Vincent Grégoire<sup>d</sup>, François Bidault<sup>b</sup>, Lydia Fresco<sup>a</sup>, John Sage<sup>a,c</sup>, John Mills<sup>a</sup>, Jean Bourhis<sup>b</sup>, Dimitri Lefkopoulos<sup>b</sup>, Olivier Commowick<sup>e,f,1</sup>, Mehdi Benkebil<sup>e</sup>, Grégoire Malandain<sup>f</sup>

<sup>a</sup> University Hospital Coventry Warwickshire (UHCW) NHS Trust, Coventry, UK

<sup>b</sup> Institut Gustave-Roussy (IGR), Villejuif, France

<sup>c</sup> University Hospitals Leicester NHS Trust, Leicester, UK

<sup>d</sup> Saint Luc University Hospital (UCL), Brussels, Belgium

<sup>e</sup> DosiSoft, Cachan, France

<sup>f</sup> Institut National de Recherche et Informatique, Sophia Antipolis, France

## ARTICLE INFO

## Article history:

Received 12 January 2009

Accepted 13 August 2009

Available online xxx

## Keywords:

Automatic segmentation  
Radiation oncology  
Qualitative evaluation study  
Head and neck cancer  
Organs at risk

## ABSTRACT

**Background and purpose:** Accurate conformal radiotherapy treatment requires manual delineation of target volumes and organs at risk (OAR) that is both time-consuming and subject to large inter-user variability. One solution is atlas-based automatic segmentation (ABAS) where *a priori* information is used to delineate various organs of interest. The aim of the present study is to establish the accuracy of one such tool for the head and neck (H&N) using two different evaluation methods.

**Materials and methods:** Two radiotherapy centres were provided with an ABAS tool that was used to outline the brainstem, parotids and mandible on several patients. The results were compared to manual delineations for the first centre (EM1) and reviewed/edited for the second centre (EM2), both of which were deemed as equally valid gold standards. The contours were compared in terms of their volume, sensitivity and specificity with the results being interpreted using the Dice similarity coefficient and a receiver operator characteristic (ROC) curve.

**Results:** Automatic segmentation took typically ~7 min for each patient on a standard PC. The results indicated that the atlas contour volume was generally within  $\pm 1SD$  of each gold standard apart from the parotids for EM1 and brainstem for EM2 that were over- and under-estimated, respectively (within  $\pm 2SD$ ). The similarity of the atlas contours with their respective gold standard was satisfactory with an average Dice coefficient for all OAR of  $0.68 \pm 0.25$  for EM1 and  $0.82 \pm 0.13$  for EM2. All data had satisfactory sensitivity and specificity resulting in a favourable position in ROC space.

**Conclusions:** These tests have shown that the ABAS tool exhibits satisfactory sensitivity and specificity for the OAR investigated. There is, however, a systematic over-segmentation of the parotids (EM1) and under-segmentation of the brainstem (EM2) that require careful review and editing in the majority of cases. Such issues have been discussed with the software manufacturer and a revised version is due for release.

Crown Copyright © 2009 Published by Elsevier Ireland Ltd. All rights reserved. Radiotherapy and Oncology xxx (2009) xxx–xxx

Recent advances in 3D conformal radiotherapy allow precise delivery of radiation dose to the tumour volume whilst maintaining sparing of normal tissue [1]. These advances demand accurate quantification of dose received to both the target volume and organs at risk (OAR) that is achieved via accurate volume delineation and dose-volume histogram (DVH) analysis from a treatment plan-

ning system (TPS). This becomes extremely important with techniques such as intensity modulated radiotherapy (IMRT) where it is possible for a large number of OAR to be “bathed” in regions of absorbed dose when striving for increased conformity. In such situations considerable clinician contouring is required that is both time-consuming and exhibits a high degree of inter-user variability [2,3, and references therein]. The development of robust automatic segmentation tools could therefore be used to accurately and reproducibly delineate all OAR to allow more time for the clinician to delineate the target volume and attend to other aspects of the patient’s treatment.

\* Corresponding author. Address: Radiotherapy Physics, Leicester Royal Infirmary, University Hospitals of Leicester, Infirmary Close, Leicester, LE1 5WW, UK.

E-mail address: richard.w.sims@uhl-tr.nhs.uk (R. Sims).

<sup>1</sup> Present address: Children’s Hospital, Boston, MA, USA.

Automatically segmenting human anatomy from medical images is a long standing problem and is where radiotherapy treatment planning has been a major driving force in the development of robust algorithms and techniques [4]. The vast majority of systems designed for use in radiotherapy treatment planning are based on semi-automatic techniques that solve the partial problem of anatomy segmentation using *a priori* information such as an anatomical atlas [4, and references therein]. Such tools will become increasingly important in the future where treatment solutions involving image-guided and adaptive radiotherapy will require very fast and robust registration algorithms to adapt patient and organ position and shape online [5,6]. The problem with medical images is that each patient's anatomy is different, and each patient's organs vary on a daily basis in size, position, shape and composition. Combined with this is the variation in image quality between successive acquisitions and different patients. Several teams of researchers have developed and/or evaluated automatic segmentation techniques for the brain [7–9] and the H&N [10,11] with the development of techniques for other anatomical sites being challenging due to increased soft-tissue deformity and respiratory motion.

This paper describes an evaluation of an atlas-based automatic segmentation (ABAS) tool (Version 3.1 ISOgray™ ABAS system by DOSIsoft) for the H&N. To the author's knowledge, only two clinical evaluations have previously been published for various H&N OAR [6,12].

## Materials and methods

### The ISOgray™ ABAS tool

The ISOgray™ ABAS tool is constructed from a database of 45 node-negative pharyngeal-laryngeal squamous cell carcinoma patients that were delineated by the Radiation Oncology Department of Saint Luc University Hospital – Catholic University of Louvain (UCL) using guidelines given by [13,14]. The construction of the ABAS tool is performed in three steps [10]. The first step involves the construction of an unbiased mean image from the database of patients by determining the transformation of each CT dataset to a reference image using a similar method to [15]. The second step involves applying the previously determined transformations to each set of manually outlined contours to transform them to the reference image before using an Expectation Maximisation (EM) algorithm to determine the mean contours. The third and final step involves symmetrising the mean image and contours to avoid discrepancies when registering with a patient. The symmetrized atlas can then be adapted onto a patient's CT data to produce a set of automatically segmented contours tailored to that patient. This process involved two steps; the first is a global affine transformation between the patient and the mean atlas image using a robust block-matching registration algorithm [16]. The remaining local deformations due to inter-patient variability are recovered using a nonlinear registration method. The resulting transformation matrix is then applied to the atlas contours to produce the new patient contours. Further information can be found in [10,17]. This ABAS tool allows automatic segmentation of brainstem, spinal cord, mandible, sub maxillary glands, lymph nodes level II to IV and parotids. In the present work, only the brainstem, mandible and parotid automatic segmentations were studied.

### Evaluation methodology

A retrospective study of 6 (3 IGR + 3 UCL) and 7 (UHCW) patients who previously underwent a radical course of conformal radiation therapy to the H&N region has been performed using

the ISOgray™ ABAS tool. All patients were CT scanned locally using standard H&N scanning protocols with slice thicknesses of between 2.5 and 3.0 mm and the patient's head immobilized using a thermoplastic shell. The CT data were subsequently imported into the ISOgray™ system and the ABAS tool was run for the mandible, brainstem and parotids. The OAR were either manually delineated by two experts (FB, VG; referred to as evaluation method 1, EM1) or the original atlas contours were reviewed and edited as necessary (LF; referred to as evaluation method 2, EM2). Both sets of clinician contours were deemed as equally valid gold standard against which the atlas results were compared. All these data were made available to the investigating physicists who performed the analysis using the methods described below.

### Analysis procedure

The performance of the ABAS tool was assessed quantitatively by comparing the results with clinician contours from the EM1 and EM2. The difference between the atlas volume  $V_{atlas}$  and the gold standard volume  $V_{gs}$  was determined for each study and OAR by evaluating:

$$\Delta V = \left( \frac{V_{atlas} - V_{gs}}{V_{gs}} \right) * 100\% \quad (1)$$

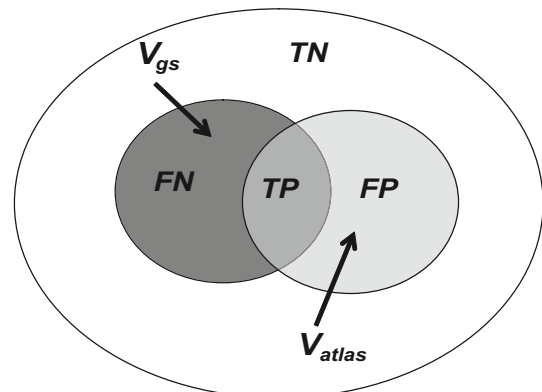
with positive or negative results indicating that the atlas contour volume has been over- or under-estimated, respectively. The sensitivity  $s_e$  and specificity  $s_p$  of the atlas contours were determined by evaluating:

$$s_e = \frac{TP}{TP + FN} \quad (2)$$

$$s_p = \frac{TN}{TN + FP} \quad (3)$$

where  $TP$ ,  $TN$ ,  $FP$  and  $FN$  are the true positive, true negative, false positive and false negative volumes of the atlas contour, respectively, shown schematically in Fig. 1. If the  $TN$  volume is simply taken as being outside the gold standard and atlas volume, but within the patient surface contour, the specificity will always tend to unity owing to the relative size of these volumes. This has been overcome by deriving a more comparable  $TN$  volume by determining the average of the atlas and the gold standard volume for each evaluation method.

If the atlas contour is an exact match in terms of size, shape and position to the gold standard the sensitivity and specificity will tend towards unity, whereas if the contours are completely differ-



**Fig. 1.** Schematic illustrating the true positive ( $TP$ ), true negative ( $TN$ ), false positive ( $FP$ ) and false negative ( $FN$ ) volumes in relation to the gold standard ( $V_{gs}$  shown in dark grey) and atlas ( $V_{atlas}$  shown in light grey) volume.

ent with no intersection the sensitivity and specificity will tend towards zero. Sensitivity or specificity, in isolation, only tells half the story and so the results have been combined in the form of a receiver operator characteristics (ROC) curve that plots the true positive rate (sensitivity) versus the false positive rate ( $1 - \text{specificity}$ ).

The similarity between the atlas and gold standard contours has been quantified on a voxel-by-voxel basis for each OAR by using the Dice similarity coefficient (DSC) [18]. The DSC has been determined by evaluating:

$$DSC = \frac{2|V_{atlas} \cap V_{gs}|}{V_{atlas} + V_{gs}} = \frac{2TP}{2TP + FN + FP} \quad (4)$$

where a weighting of two is given to intersections of the atlas and gold standard volume. It has been shown that the use of the DSC is appropriate in comparison-of-agreement studies, specifically image registration [19], with MRI studies suggesting DSC values  $>0.70$  represent good agreement [20].

## Results

The ABAS tool was used on all patients to outline the OAR of interest without any user intervention. The registration process took  $\sim 7$  min on a standard PC (2.16 GHz single core processor, 2.0 Gb RAM), which is considerably less than a typical time of approximately 1 h to manually delineate (EM1) or review and edit (EM2) the atlas contours. It is important that the results of the atlas require minimal final adjustment to ensure that the time-saving potential of the system is realised. An example illustrating all contours from both studies is shown in Fig. 2.

The atlas and gold standard volumes for all OARs and both evaluation methods are shown in Fig. 3. The results suggest that the majority of atlas volumes are within  $\pm 1SD$  of each gold standard apart from the parotids for EM1 and the brainstem for EM2 that were over- and under-estimated, respectively. The average volume difference between the atlas and each gold standard volume is shown in Table 1 for each OAR and evaluation method. The ratio of FP and FN volume to atlas volume, also shown in Table 1, illustrates how all OAR for EM1 are dominated by a high incidence of false positive volumes whereas the proportion of false negative volumes is consistent for all OAR and all evaluation methods.

The DSC results shown in Table 2 show how the result for the mandible is similar between evaluation methods whereas the DSC is smaller for all other OAR in the EM1 study. This is because the DSC is inversely proportional to the sum of both the false positive and false negative volumes (see Eq. (4)). The sensitivity and specificity results shown in Table 3 indicate how the sensitivity of all OAR is similar for both evaluation methods, whereas the specificity is significantly lower for EM1. This is consistent with the FP/Atlas and FN/Atlas results shown in Table 1 and the DSC results shown in Table 2. The sensitivity and specificity results have been combined into the ROC curve shown in Fig. 4. All the results are located within the upper-left quadrant of ROC space with the EM1 results displaced further towards the right due to the increased proportion of the atlas volume being false positive.

It is important to ensure that the variability in clinician outlining ability, and selected patients, is small. This has been determined by calculating the coefficient of variation (CV), defined as the ratio of the standard deviation to mean OAR volume, for the different clinician and patient combinations. The inter-patient

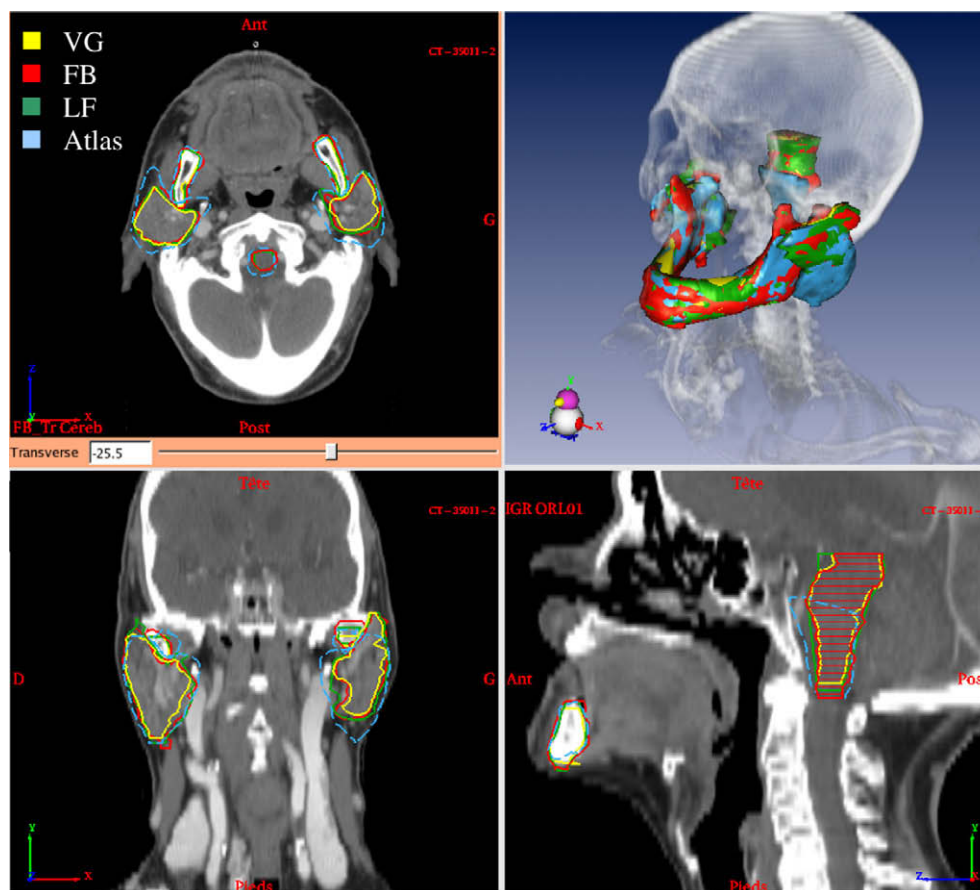
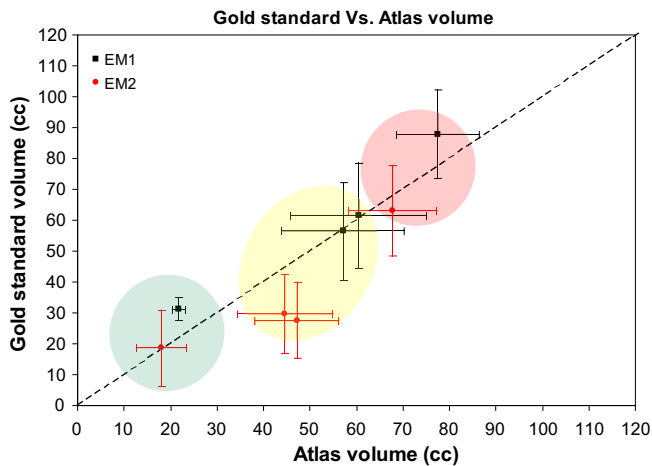


Fig. 2. Typical results for one patient used in this study. 'VG' and 'FB' correspond to the two clinicians from the EM1 study, 'LF' corresponds to the clinician from the EM2 study and 'Atlas' corresponds to the ABAS generated contours (dashed) corresponds to the ABAS generated contours.



**Fig. 3.** A comparison between the gold standard and atlas volume for the four OAR and two evaluation methods. The diagonal dashed line indicates the region where the gold standard and atlas volumes are equal, with data above or below this line indicating where the atlas volume has been under-estimated or over-estimated respectively ( $\pm 1SD$ ). Red, yellow and green shaded areas correspond to the mandible, parotids and brainstem OAR, respectively.

**Table 1**

The average percentage volume difference ( $\Delta V$ ) between the atlas and gold standard contours for all OAR and both evaluation methods. The magnitude of error associated with these measurements can be seen in Fig. 3. Also shown is the ratio  $FP/Atlas$  and  $FN/Atlas$  that helps identify large proportions of the atlas volume that are deemed as either false positive or false negative, respectively ( $\pm 1SD$ ).

OAR	EM1			EM2		
	$\Delta V$ (%)	$FP/Atlas$	$FN/Atlas$	$\Delta V$ (%)	$FP/Atlas$	$FN/Atlas$
Brainstem	-2.1	$0.43 \pm 0.18$	$0.41 \pm 0.42$	-30.2	$0.06 \pm 0.04$	$0.49 \pm 0.16$
Mandible	+7.1	$0.24 \pm 0.08$	$0.18 \pm 0.09$	-11.6	$0.13 \pm 0.04$	$0.26 \pm 0.11$
Left Parotid	+49.8	$0.42 \pm 0.13$	$0.09 \pm 0.09$	-1.6	$0.15 \pm 0.11$	$0.17 \pm 0.09$
Right Parotid	+70.7	$0.47 \pm 0.13$	$0.06 \pm 0.06$	+1.2	$0.15 \pm 0.08$	$0.13 \pm 0.11$

**Table 2**

The DSC results for the four OAR of interest ( $\pm 1SD$ ).

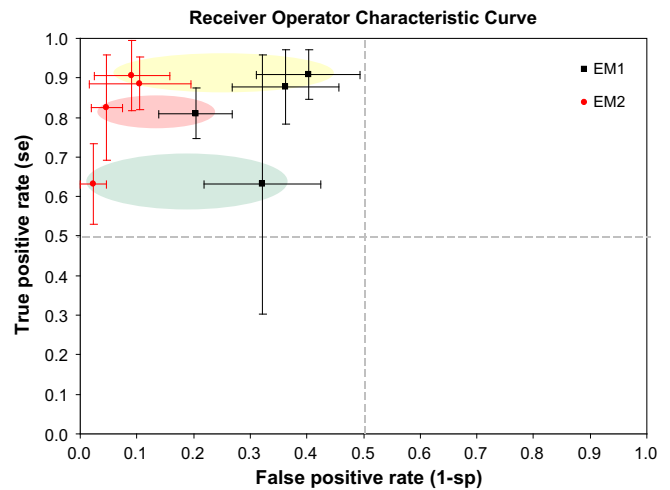
OAR	DSC	
	EM1	EM2
Brainstem	$0.58 \pm 0.20$	$0.77 \pm 0.07$
Mandible	$0.78 \pm 0.06$	$0.82 \pm 0.04$
Left Parotid	$0.69 \pm 0.09$	$0.84 \pm 0.07$
Right Parotid	$0.66 \pm 0.10$	$0.86 \pm 0.07$

**Table 3**

Sensitivity and specificity results for the four OAR and two different evaluation methods ( $\pm 1SD$ ).

OAR	Sensitivity		Specificity	
	EM1	EM2	EM1	EM2
Brainstem	$0.63 \pm 0.33$	$0.63 \pm 0.10$	$0.68 \pm 0.10$	$0.98 \pm 0.02$
Mandible	$0.81 \pm 0.06$	$0.83 \pm 0.13$	$0.80 \pm 0.07$	$0.95 \pm 0.03$
Left Parotid	$0.88 \pm 0.09$	$0.89 \pm 0.07$	$0.64 \pm 0.09$	$0.89 \pm 0.09$
Right Parotid	$0.91 \pm 0.06$	$0.91 \pm 0.09$	$0.60 \pm 0.09$	$0.91 \pm 0.07$

variability is shown in Table 4, illustrating the similarity between the two evaluation methods (well within  $\pm 1SD$ ) for all OAR apart from the brainstem where the patient variation for EM1 was significantly greater. The intra-clinician variability for EM1 and the clinician variability between the two studies (using a common patient), was smaller than the inter-patient variability and is not deemed significant.



**Fig. 4.** True positive versus false positive rate for the four OAR and two evaluation methods. All results are clustered in the top left-hand corner of the plot that corresponds to high sensitivity and specificity (all  $\pm 1SD$ ). Red, yellow and green shaded areas correspond to the mandible, parotids and brainstem OAR, respectively.

**Table 4**

Inter-patient variability, in terms of the coefficient of variation (CV) for the two different evaluation methods ( $\pm 1SD$ ).

OAR	EM1 CV	EM2 CV
Brainstem	$0.46 \pm 0.01$	0.12
Mandible	$0.18 \pm 0.07$	0.16
Left parotid	$0.30 \pm 0.07$	0.28
Right parotid	$0.31 \pm 0.08$	0.28

## Discussion

This evaluation study has presented the first results assessing the accuracy of version 3.1 of the ISOgray™ ABAS tool for the H&N. The volume of the atlas contours for the majority of OAR is within  $\pm 1SD$  of the gold standard regardless of which evaluation method was used. There are deviations from this for both parotids from the EM1 study and the brainstem from the EM2 study where the atlas volume was over- and under-estimated, respectively. The brainstem volume was under-estimated by the EM2 study due to the large proportion of the atlas contour that was deemed as false negative. Close inspection of the contours indicated that this was not due to poor conformity between the atlas and gold standard, but to a lack in superior extent of the atlas contour. This is due to the superior restriction of the CT data used to construct the brainstem contour within the atlas causing an underestimation of brainstem extent and volume. This can be seen in the sagittal view shown in Fig. 2 (see electronic version for colour reproduction). This was also the case for the EM1 study although the volume results were masked by the increased proportion of the atlas brainstem volume that was classified as false positive (see Table 1).

The parotid volume was over-estimated in the EM1 study with inspection of the contours indicating that this was a global over-estimation and not restricted to any particular direction. A similar result was identified by [10] where one patient was removed during the construction of the atlas and later the atlas was registered to this patient. A possible reason for over-segmentation was suggested to be a combination of registration discrepancies related to patient and clinician variability within the ABAS tool. It was suggested that the intra-patient variability could be reduced by creating sub-populations within the atlas that can be chosen



depending on various characteristics of a given patient. The inter-patient variability for the parotids was very similar for both studies and so this is not likely to be the cause (see Table 4). The lack of parotid over-segmentation in the EM2 study may be related to the characteristics of the patients (e.g., neck fat) being more closely matched to those that constructed the atlas, although there is also likely to be a bias towards the atlas contours because they were used as a first approximation prior to review and edit.

It is important to understand the volume discrepancies identified for the parotids in terms of the change in contour that would be required to observe such differences. [7] interpreted that the over-estimation and underestimation of volumes in terms of the margin to be applied to the automatic segmented volume to correspond to the gold standard volume. This analysis can be applied to the parotids for the EM1 study if one assumes the parotid can be approximated as a prolate ellipse. In this case the atlas contours have globally over-estimated the parotid volume by between 2.5 and 3.3 mm. This is not significant in the superior/inferior direction owing to CT slice thickness but it becomes increasingly important in the anterior/posterior and right/left directions. This global systematic over-segmentation of the parotids is due to be rectified in a future version of the software by “eroding” the atlas contour using a morphological mathematical operation (private communication A. Isambert).

The DSC results summarised in Table 2 indicate that the coefficient for the brainstem and parotids is lower for EM1 whereas the values for the mandible are very similar. Inspection of Tables 1 and 3 indicates that the sensitivity for both the EM1 and EM2 studies were consistent with one another whereas the specificity for EM1 is considerably lower due to an increased proportion of atlas contours being false positive due to the over-segmentation. The DSC value for the brainstem for the EM1 study is relatively low due to the large false positive and false negative volumes, but is also very variable due to a large inter-patient variability with a CV of  $0.46 \pm 0.01$  compared to 0.12 for the EM2 study (see Table 4). The DSC results for the parotids for EM1 can clearly be attributed to over-segmentation with an increased proportion of false positive volumes (see Tables 1 and 2). This difference in DSC may again be related to either a bias towards accepting the atlas contours for the EM2 study, or due to a better match between the EM2 patients and those that constructed the atlas. Despite these differences all data points occur in the upper-left quadrant of ROC space where the values of sensitivity and specificity are such that an acceptable sparing of OAR and coverage of a planning target volume (PTV) can be expected [7].

## Conclusions

The efficacy of the ISOgray™ H&N ABAS tool has been successfully evaluated using two different methods from two different centres for node-negative patients who had previously undergone a radical course of 3D conformal radiotherapy. The results demonstrated that all data are positioned in the favourable quadrant of ROC space for the delineation of OAR of interest in this study. The average DSC values for all organs was  $0.68 \pm 0.25$  and  $0.82 \pm 0.13$  for the EM1 and EM2 study, respectively, representing satisfactory agreement between the atlas and respective gold standard volumes. It is noted, however, that there was a tendency for the ABAS tool to over-segment the parotids and brainstem in the EM1 study that was not observed to the same extent with the EM2 study. This is largely due to the inherent bias in the EM2 study where the atlas contours were used as a first approximation prior to review and edit, although different patient characteristics between the two studies may also be a factor. The atlas-derived

brainstem contour for both studies suffered a premature termination in the superior direction that was attributed to the lack of atlas data. These issues have been discussed with the manufacturer and it is suggested that for the majority of cases careful review and editing would be required for sufficient sparing of the OAR which were studied as well as coverage of the PTV. The amount of final review and editing required would have a bearing on whether the time-saving potential of the system would prove a net benefit to clinical workload.

## Acknowledgements

This work was undertaken in the framework of the MAESTRO Integrated Project (IP CE503564) funded by the European Commission.

## References

- [1] Mackie TR, Kapatoes J, Ruchala K, et al. Image guidance for precise conformal radiotherapy. *Int J Radiat Oncol Biol Phys* 2003;56:89–105.
- [2] Geets X, Daisne J-F, Arcangeli S, et al. Inter-observer variability in the delineation of pharyngeal-laryngeal tumour, parotid glands and cervical spinal cord: comparison between CT-scan and MRI. *Radiother Oncol* 2005;77:25–31.
- [3] Fiorino C, Reni M, Bolognesi A, Cattaneo GM, Calandrino R. Intra- and inter-observer variability in contouring prostate and seminal vesicles: implications for conformal treatment planning. *Radiother Oncol* 1998;47:285–92.
- [4] Haas B, Coradi T, Scholz M, et al. Automatic segmentation of thoracic and pelvic CT images for radiotherapy planning using implicitly anatomic knowledge and organ-specific segmentation strategies. *Phys Med Biol* 2008;53:1751–71.
- [5] 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* 2005;50:2887–905.
- [6] Wang H, Garden AS, Zhang L, et al. Performance evaluation of automatic anatomy segmentation algorithm on repeat or four-dimensional computed tomography images using deformable image registration method. *Int J Radiat Oncol Biol Phys* 2008;72:210–9.
- [7] Isambert A, Dhermain F, Bidault F, et al. Evaluation of an atlas-based automatic segmentation software for the delineation of brain organs at risk in a radiation therapy clinical context. *Radiother Oncol* 2008;87:93–9.
- [8] Popple RA, Griffith HR, Sawrie SM, Fiveash JB, Brezovich IA. Implementation of talairach atlas based automated brain segmentation for radiation therapy dosimetry. *Technol Cancer Res Treat* 2006;5:15–21.
- [9] Bondiau PY, Malandain G, Chanalet S, et al. Atlas-based automatic segmentation of MR images: validation study on the brainstem in radiotherapy context. *Int J Radiat Oncol Biol Phys* 2005;61:289–98.
- [10] Commowick O, Grégoire V, Malandain G. Atlas-based delineation of lymph node levels in head and neck computed tomography images. *Radiother Oncol* 2008;87:281–9.
- [11] Street E, Hadjiiski L, Sahiner B, et al. Automated volume analysis of head and neck lesions on CT scans using 3D level set segmentation. *Med Phys* 2007;34:4399–408.
- [12] Zhang T, Yuwei C, Meldolesi E, Yan D. Automatic delineation of on-line head-and-neck computed tomography images: toward on-line adaptive radiotherapy. *Int J Radiat Oncol Biol Phys* 2007;68:522–30.
- [13] Grégoire V, Levendag P, Ang KK, et al. CT-based delineation of lymph node levels and related CTVs in the node-negative neck: DAHANCA, EORTC, GORTEC, NCI, RTOG consensus guidelines. *Radiother Oncol* 2003;69:227–36.
- [14] Grégoire V, Eisbruch A, Hamoir M, Levendag P. Proposal for the delineation of the nodal CTV in the node-positive and post-operative neck. *Radiother Oncol* 2006;79:15–20.
- [15] Guimond A, Meunier J, Thirion JP. Average brain models: a convergence study. *Comput Vis Comput Understand*. 2000;77:192–210.
- [16] Ourselin S, Roche A, Prima S, Ayache N. Block matching: a general framework to improve robustness of rigid registration of medical images. In: *Int Conf Med Image Comput Assist Interv LNCS*, vol. 1935. Springer, 2000. p.557–66.
- [17] Commowick O, Arsigny V, Isambert A, et al. An efficient locally affine framework for the smooth registration of anatomical structures. *Med Image Anal* 2008;12:427–44.
- [18] Dice LR. Measures of the amount of ecologic association between species. *Ecology* 1945;26:297–302.
- [19] Bharatha A, Hirose M, Hata N, et al. Evaluation of three-dimensional finite element-based deformable registration of pre- and intraoperative prostate imaging. *Med Phys* 2001;28:2551–60.
- [20] Zijdenbos AP, Dawant BM, Margolin RA, Palmer AC. Morphometric analysis of white matter lesions in MR images: method and validation. *IEEE Trans Med Imaging* 1994;13:716–24.