Influence of Contouring Style on Accuracy of Atlas-Based Auto-Contouring

G.N Stevenson¹, L. Durrant², C. Eccles³, K-Y Chu², T. Kadir¹ and M. Gooding¹
¹Mirada Medical, Oxford, United Kingdom, ² Gray Institute of Radiation Oncology and Biology, University of Oxford, Oxford, United Kingdom

Objective
This study aims to investigate the impact of atlas-based auto contouring upon the accuracy of contouring for operators with varying training and familiarity with a particular contouring convention.

Materials and Method
A set of ten head and neck CT datasets were contoured by Contourer 1 and validated by another expert clinician. These data were used as the ground truth when comparing conformity. A single dataset was selected to serve as an atlas to auto-contour the remaining nine cases using commercial deformable image registration software (Mirada RTx, Mirada Medical, UK).

For each of these nine cases, the brainstem, spinal cord, left and right parotid and left and right submandibular glands were independently contoured by Contourers 1, 2 and 3 by hand from scratch to produce a manual segmentation. Subsequently, each contourer then edited the auto-contours provided by the atlas-based contouring system to produce a set of edited auto-contours for each subject.

Contourer 2 was familiar with the contouring convention used by Contourer 1 whilst Contourer 3 was not.

Dice similarity coefficient (DSC) was used to compare manual contours and the edited auto-contours against the original contours of Contourer 1 for six different regions of interest.

Results
Table 1 shows DSC of manual contours drawn from scratch against the ground truth. Table 2 shows DSC for each contourer’s edited auto-contours against the ground truth.

Table 1 – Average DSC for Manual Contours vs. Ground Truth

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Contourer 1</td>
<td>0.90</td>
<td>0.86</td>
<td>0.87</td>
<td>0.90</td>
<td>0.88</td>
<td>0.90</td>
</tr>
<tr>
<td>Contourer 2</td>
<td>0.88</td>
<td>0.83</td>
<td>0.82</td>
<td>0.89</td>
<td>0.81</td>
<td>0.86</td>
</tr>
<tr>
<td>Contourer 3</td>
<td>0.85</td>
<td>0.82</td>
<td>0.55</td>
<td>0.76</td>
<td>0.49</td>
<td>0.77</td>
</tr>
</tbody>
</table>

Differences in DSC scores for each contourer varied with familiarity of the contouring convention. For Contourers 1 and 2, the differences between edited auto-contours versus manual contours were minor with only small differences in DSC observed. Typically the edited auto-contours had greater agreement with the atlas.

Contourer 3, who was not familiar with the contouring style of the other two observers, showed larger differences in DSC. The majority of these differences were improvements in DSC when the edited auto-contours were compared to the manual results.

Figures 1 and 2 show manual and edited auto-contours drawn on 2 different patients by all 3 contourers compared to the atlas based auto segmentation. They illustrate the potential change in clinical opinion caused by editing an atlas-based auto segmentation compared to drawing a contour from scratch.

Conclusions
Atlas-based auto-contouring using deformable image registration was found to improve the conformity to the atlas contours of Contourer 3 for the parotid and submandibular glands.

DSC for Contourers 1 and 2 were largely unaffected by auto-contours. This suggests that familiarity with contouring convention, or lack thereof, is an important factor in assessing auto-contouring efficacy and that auto-contouring may improve conformity to a standard.

Within a clinical setting, or for clinical trials, atlas-based auto contouring may have benefit in training and quality assurance in developing consensus.