A comparison of evaluation methods for comparing contours of swallowing organs in the head and neck

Poster No.: R-0148
Congress: 2014 CSM
Type: Scientific Exhibit
Authors: S. Smith, L. Gholamrezaei, J. Benson, L. McLean, G. Whittington, K. Foo; SYDNEY/AU
Keywords: Anatomy, Head and neck, CT, Comparative studies, Radiation therapy / Oncology, Quality assurance
DOI: 10.1594/ranzcr2014/R-0148

Any information contained in this pdf file is automatically generated from digital material submitted to EPOS by third parties in the form of scientific presentations. References to any names, marks, products, or services of third parties or hypertext links to third-party sites or information are provided solely as a convenience to you and do not in any way constitute or imply RANZCR/AIR/ACPSEM's endorsement, sponsorship or recommendation of the third party, information, product or service. RANZCR/AIR/ACPSEM is not responsible for the content of these pages and does not make any representations regarding the content or accuracy of material in this file.

As per copyright regulations, any unauthorised use of the material or parts thereof as well as commercial reproduction or multiple distribution by any traditional or electronically based reproduction/publication method ist strictly prohibited.

You agree to defend, indemnify, and hold RANZCR/AIR/ACPSEM harmless from and against any and all claims, damages, costs, and expenses, including attorneys' fees, arising from or related to your use of these pages.

Please note: Links to movies, .ppt slideshows, .doc documents and any other multimedia files are not available in the pdf version of presentations.
The aim of this study is to evaluate clinical methods of comparison of organs used for swallowing in the head and neck region. There have been many studies comparing PTVs and other larger organs ie bladder, though few who compare small irregular shapes such as those organs involved in swallowing and therefore there is no consensus as to the optimal method. Our aim especially seeks to determine the difference between comparing contours of swallowing organs at risk (SWOARs) to clinically significant PRVs and more conventional methods of contour comparison.

Methods and materials

A retrospective study comprising five patients who received radical radiation treatment to the head and neck region were investigated by six study participants; two Radiation Oncologists (RO) and four Radiation Therapists. Each of the six participants used an anatomical atlas detailing the contouring parameters of nine swallowing organs in the head and neck as outlined in Christianen et al's 2011[1] paper. The superior, middle and inferior pharyngeal constrictor muscles (SPCM, MPCM, IPCM), cricopharyngeal muscle (CM), oesophageal inlet muscle (OIM), cervical oesophagus (CE), Base of Tongue (BOT), glottic larynx (GL) and supraglottic larynx (SL) were contoured on every patient by each observer, resulting in 270 contours.

The gold standard of each of the nine structures was considered to be that of the most experienced RO, after having been qualitatively deemed acceptable by the second RO involved. The nine structures created by the other five observers were compared using absolute volume, conformity index and percentage of volume encompassed by a planning organ at risk volume (PRV) ie a 3mm expansion on the gold standard (Figure 1).

Absolute volume was observed by using the data supplied by the treatment planning system (Eclipse V10.4, Varian Medical Systems Palo Alto, CA) and represented as a percentage volume of the gold standard. The Conformity index (CI) in this study is defined as the ratio of the intersection and the union of two delineated volumes, where an outcome of 1 represents identical structures and 0 represents no overlap at all (Figure 2). The CI was calculated as a mean value per observer for each individual SWOAR. Each contour was also qualitatively assessed by two Radiation Oncologists.
Fig. 1: PRV for Middle Pharyngeal Constrictor Muscle Yellow = Gold standard Blue = PRV ie 3mm expansion on gold standard
**Fig. 2:** Conformity Index

The Conformity Index (CI) is calculated as follows:

\[
CI = \frac{\text{Intersection}}{\text{Union}}
\]
Results

A qualitative assessment of all structures by the two ROs suggested all contours were clinically acceptable.

The conformity index for each patient was measured as an average for each contour per observer and is demonstrated in figure 3. The contour with the highest conformity observed using this comparison method was the Cervical Oesophagus with average CIs for each observer #0.9. Using the CI method the Cricopharyngeal Muscle & Supraglottic larynx least conform to the gold standard.

Each observer's contours were analysed as to the amount of volume, as a percentage, that was encompassed by the PRV. The PRV comparative analysis showed that all observers achieved a minimum of 95% of the average of their structure encompassed by the PRV across all patients.

Analysis of the absolute volume, conformity index and PRV technique showed discrepancies between the methods. The three largest discordances can be seen in Table 1.

Measures of absolute volume and conformity index are commonly used to compare contouring of PTVs and organs at risk [2-3] both being easy to measure and assess. Despite being the most reported measure in the literature [4] unfortunately absolute volume does not account for variation in position or shape. Our results show that even though the CE has the largest deviation to the gold standard in absolute volume, the CI and PRV method contraindicate the deviation with high measures of concordance.

The Supraglottic larynx received relatively low CIs across observers and a large variation in absolute volume the SL was the only structure to achieve 100% of volume within the PRV in all instances and was deemed acceptable by clinical observation.

Conformity Index is well represented as a measure of contour comparison in the literature and has produced useful results in reporting on regular and relatively larger structures including contouring of the prostate [5-6] and breast radiotherapy structures [7-9]. These results are valid and useful in these situations where the centre of mass is known to be similar. Our results present a high CI for the CE which can be attributed to the larger nature of the CE contour in comparison to the other SWOARs. Unfortunately CI used alone or for small irregular structures is not as valuable as the measure does not give information regarding the size and location of the variation and is too sensitive to minor variations which are occur in the small structures of the SWOARs as demonstrated in Figure 4.
It is the opinion of the investigators that the PRV method is the most clinically significant, allowing for some small variation of the contours. The PRV method is not however sensitive to contours that are considerably small though remain within the PRV (Figure 5a&b). Therefore this method could not be used in isolation and so requires collaboration with other comparative techniques. One suggestion for future studies is to measure the lower limit of the contour with a reduction volume created from the gold standard.

Images for this section:

**Fig. 3:** Conformity Index for each patient as an average for each contour per observer (n=225)

<table>
<thead>
<tr>
<th>Organ</th>
<th>Absolute Volume (% of gold standard)</th>
<th>Absolute Volume Range (% of gold standard)</th>
<th>CI</th>
<th>CI Range</th>
<th>PRV (% structure within PRV)</th>
<th>PRV Range (% structure within PRV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supraglottic</td>
<td>49.3</td>
<td>49.3-83.8</td>
<td>0.66</td>
<td>0.47-0.78</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Larynx</td>
<td>115.5</td>
<td>102.2-134.9</td>
<td>0.92</td>
<td>0.9-0.95</td>
<td>97</td>
<td>97-100</td>
</tr>
<tr>
<td>CE</td>
<td>102.9</td>
<td>88.2-115.4</td>
<td>0.74</td>
<td>0.64-0.87</td>
<td>98</td>
<td>96-100</td>
</tr>
</tbody>
</table>

*Table 1:* Comparison of discordance between measures of comparison
Low conformity index
Contours not clinically different
(small impact on plan optimisation)

Create Planning Risk Volume (PRV) around gold standard contour using expansion margin

Assess different operator contour against PRV

Fig. 4: Comparison of CI and PRV methods
**Fig. 5:** PRV Results. Both images depict a far smaller observer contour than the gold standard yet would receive a PRV score of 100% Yellow: observer contour Green: Gold standard Red: PRV
Conclusion

The nature of the organs in the head and neck outlined in this study highlights the restrictive nature of conventional methods of contour comparison. The PRV method of comparison did not find the same discordance, as was evident when using the conventional methods, though still poses its disadvantages. The re-entrant shape, size and location of some of these structures necessitates the development of more clinically viable and novel measures.

Personal information

References


