circular MLC aperture. The dosimetric tracking accuracy was quantified by measuring the dose of low and high modulation VMAT plans to a moving Delta4PT phantom that reproduced the same eight trajectories and comparing this dose with a static reference dose (2%/2mm gamma evaluation). A linear Kalman filter prediction was used to account for the tracking system latency in the geometric and dosimetric accuracy experiments.

Results: The mean tracking system latency was 145ms (SD=7.4ms) for MLC tracking and 143ms (SD=17ms) for the tracking backup jaws. Compared to non-tracking, MLC tracking reduced the mean root-mean-square geometric targeting error from 2.7mm to 0.5mm (parallel to MLC leaves) and from 2.4mm to 1.1mm (perpendicular). Dosimetrically, MLC tracking reduced the motion induced gamma failure rate from 30.0% to 9.4% (prostate) and from 41.2% to 3.4% (lung) on average (see Table).

Conclusions: TrueBeam MLC tracking performance was thoroughly investigated for the first time. The system has similar geometric and dosimetric performance as previously reported for prototype MLC tracking systems. The inclusion of backup jaw tracking effectively reduces radiation outside the field by shielding for inter-leaf leakage.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Uniform</th>
<th>Individualised</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTV1 mean dose</td>
<td>0.51</td>
<td>0.47</td>
</tr>
<tr>
<td>PTV2 mean dose</td>
<td>0.83</td>
<td>0.81</td>
</tr>
<tr>
<td>PTV3 mean dose</td>
<td>0.82</td>
<td>0.91</td>
</tr>
<tr>
<td>SC max dose</td>
<td>0.73</td>
<td>0.67</td>
</tr>
<tr>
<td>BS max dose</td>
<td>0.26</td>
<td>0.73</td>
</tr>
</tbody>
</table>

Table 1: Correlation coefficients for the uniform and individualised contours

Conclusions: These preliminary results suggest that a simple tool, based on the individualised contour and area of overlap, can help predict dosimetric changes resulting from weight loss in HN VMAT patients. This could provide an objective assessment of weight loss during treatment, improving the workflow for HN ART, potentially acting as a trigger for plan adaptation.

Poster: Physics track: Management of interfraction changes/deformation and adaptive radiotherapy

PO-0934
Predicting the dosimetric consequences of weight loss from on-treatment CBCT images for head and neck VMAT patients
W.J. Beasley1, A. McWilliam1, C.G. Rowbottom2
1The Christie NHS Foundation Trust, Christie Medical Physics and Engineering, Manchester, United Kingdom
2The Clatterbridge Cancer Centre NHS Foundation Trust, Physics Department, Liverpool, United Kingdom

Purpose/Objective: Weight loss is common during head and neck (HN) radiotherapy, often resulting in reduced plan quality relative to that intended. At this institution ART is performed only when necessary, and patients are monitored for anatomic changes through weekly CBCT imaging. If radiographers suspect that observed weight loss might be significant, the patient images are referred to physics for dosimetric review. This process is time consuming, and only a small number of reviews reveal clinically significant dosimetric changes. The decision to refer a patient for dosimetric review is highly subjective, and with the implementation of ART likely to increase in the future, there is a need for tools that introduce objectivity into this process. A simple contour-based tool, which creates a patient-specific contour that can be used to objectively assess weight loss, is presented.

Materials and Methods: An individualised patient contour, created by contracting the external patient contour by distances depending on the control point weighting in the VMAT plan, was created using an in-house Python script. An additional uniform contraction of the external contour was also created to serve as a baseline for comparison. These ART contours were created for ten patients and applied to on-treatment CBCT images that had been previously analysed due to observed weight loss during treatment. The patient contour was outlined on the CBCT images and the area of overlap with the different ART contours was recorded. This area of overlap was then compared with changes in mean dose to PTVs, and changes in maximum doses to 1 cm³ (1 cc max) for the spinal cord (SC) and brainstem (BS). The correlation between the area of overlap and change in dosimetric parameters was then measured to determine whether the individualised contour could predict the dosimetric effects of weight loss.

Results: The area of overlap for both contour types was correlated with the change in mean dose to the targets and the change in 1 cc max for the SC and BS, as shown in the table. With the exception of PTV1 mean dose, which showed moderate correlation for both contour types, all dosimetric parameters showed a strong correlation with the individualised area of overlap. Although the correlation for the SC max dose was slightly stronger for the uniform contraction, all other parameters showed an equivalent or stronger correlation with the area of overlap for the individualised contour than for the uniform contour. Furthermore, the uniform contour failed to show any correlation with the max dose to the BS.

Evaluation of deformable image registration methods for dose monitoring in head and neck adaptive radiotherapy
B. Rigaudo1, A. Simon1, J. Castelli1, M. Gobeli1, J.D. Ospina Arango1, G. Cazoulat1, O. Henry1, P. Haigron1, R. De Crevoisier2
1INSERM U1099, Laboratoire Traitement du Signal et de l’Image, Rennes, France
2INSERM U1099, Centre Eugene Marquis, Rennes, France
Purpose/Objective: In head and neck cancer (HNC) adaptive radiation therapy (ART), the two purposes were to compare the accuracy of different deformable images registration (DIR) methods and to quantify their impact for dose accumulation, in healthy structures.

Materials and Methods: Fifteen HNC patients had a planning CT (CT0) and weekly CTs during the 7 weeks of intensity-modulated radiation therapy (IMRT), with or without contrast agent injection. Ten DIR combinations were tested, combining two registration methods (demons or B-spline Free-Form Deformation (FFD)), with or without image preprocessing (sigmoid filtering to enhance soft tissues and delineation mapping) and with sum of squared differences (SSD) or mutual information (MI) metric. Two observers identified 14 landmarks (LM) on each CT-scan to compute the LM registration error (mm) and the LM dose accumulation error (Gy). The cumulated doses in the parotid glands (PG) estimated by each method were compared.

Results: The two most effective DIR methods were the FFD and the demons, both with the MI metric and the filtered CTs. The corresponding LM registration accuracy (precision) were 2.44 mm (1.30 mm) and 2.54 mm (1.33 mm), respectively. The corresponding LM estimated cumulated dose accuracy (dose precision) were 0.85 Gy (0.93 Gy) and 0.88 Gy (0.95 Gy), respectively.

The inter-observer distance accuracy (precision) were 2.01 mm (1.29 mm), The inter-observer cumulated dose accuracy (dose precision) were 0.68 Gy (0.75 Gy), respectively. The median (SD) mean planned dose for the PGs was 30.22 Gy (7.76 Gy). Using the ‘FFD with MI on filtered CTs’ method to calculate the cumulated mean PG dose, 66% of the PGs presented an increase of the mean dose of 3.38 Gy (SD= 2.82 Gy, range: 0.38-11.69 Gy), and 33% of the PGs presented a decrease of the mean dose of 1.52 Gy (SD= 1.08 Gy, range: 0.06-3.22 Gy), compared to the mean planning dose. The mean uncertainty (difference between maximal and minimal estimated cumulated doses considering all the 10 methods) to estimate the cumulated mean PG dose was 4.03 Gy (SD= 2.27 Gy, range: 1.06-8.91 Gy).

Conclusions: The choice of the metric and/or of the image preprocessing is at least as important as the registration method. If the estimated local accumulated dose has to be considered carefully, the most accurate method provides the means to detect over- or under-irradiation for healthy tissues.

PO-0936

Relationship between geometric and dosimetric accuracy of auto-contouring in head and neck VMAT treatment planning

A. McWilliam¹, W. Beasley¹, C.G. Rowbottom²

¹The Christie NHS Foundation Trust, Christie Medical Physics and Engineering, Manchester, United Kingdom
²The Clatterbridge Cancer Centre, Physics Department, Liverpool, United Kingdom

Purpose/Objective: Automatic contouring (AC) will be an essential component of an adaptive radiotherapy (ART) strategy, enabling efficient treatment planning and plan adaptation. Although there has been much interest in the geometric accuracy of AC algorithms, there has been little consideration of the dosimetric accuracy of automatic contours used for treatment planning. Furthermore, it is not clear how the metrics used to assess geometric accuracy relate to dosimetric accuracy when these automatic contours are used for treatment planning. In this study the relationship between geometric and dosimetric accuracy is assessed for commonly-used metrics.

Materials and Methods: A commercial AC algorithm was used to retrospectively segment ten head and neck patients. Ground truth contours were created for each patient by combining contours from five clinicians by means of the simultaneous truth and performance level estimation (STAPLE) algorithm. The geometric accuracy of the auto-contours relative to the STAPLE-contours was assessed by a number of commonly-used metrics, as well as some novel metrics. VMAT plans were created for each patient, according to standard departmental protocols, using the auto-contours for optimisation. The difference in dose (ΔDmean or ΔDmax) to the auto-contours relative to the dose to the STAPLE-contours provided a measure of the dosimetric accuracy of the auto-contours. This dosimetric accuracy was then compared with the geometric accuracy indicated by the different metrics, and the correlation coefficient calculated for each structure.

Results: The mean ΔDmean for the parotids, submandibulars (SMG) and larynx, and mean ΔDmax for the spinal cord (SC) and brainstem (BS) is shown in the table. For the parotids, DSC showed poor correlation with ΔDmean (R = 0.39), with the strongest correlate being centroid separation (R = 0.82). For the larynx, DSC showed the strongest correlation, although the correlation was moderate (R = 0.59). The large amount of overlap of the SMG with the target volume resulted in no correlation with ΔDmean for any metric. The RMS-difference, a new metric based on the distance-to-agreement histogram, was found to provide a strong correlation with ΔDmax for the SC and BS (R = 0.61, R = 0.82).

<table>
<thead>
<tr>
<th>Structure</th>
<th>Mean dosimetric difference</th>
<th>Inter-observer range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parotids</td>
<td>-8.4%*</td>
<td>-4.7% - 3.8%</td>
</tr>
<tr>
<td>Larynx</td>
<td>-4.1%*</td>
<td>-2.1% - 4.0%</td>
</tr>
<tr>
<td>SMG</td>
<td>0.7%*</td>
<td>-1.0% - 1.1%</td>
</tr>
<tr>
<td>SC</td>
<td>1.7%*</td>
<td>-</td>
</tr>
<tr>
<td>BS</td>
<td>-1.4%*</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 1: Mean dosimetric difference between auto and STAPLE contours. *mean dose/* max dose

Conclusions: Several metrics commonly used to assess the accuracy of automatically-generated contours did not show strong correlation with dosimetric agreement between auto- and STAPLE-contours. In particular, DSC, one of the most commonly-used metrics, showed poor correlation with ΔD for most structures. The RMS-difference provided good correlation for structures for which the maximum dose is most important (SC and BS). It is important to consider a variety of metrics when assessing the acceptability of AC algorithms, depending on the specific structure and dosimetric parameter of clinical interest.