IMRT vs. VMAT: Assessing the dosimetric equivalence of two treatment delivery techniques via independent DICOM RT data analysis

Poster No.: R-0076
Congress: 2014 CSM
Type: Scientific Exhibit
Authors: M. Young; HOBART/AU
Keywords: Pelvis, Computer applications, CT, Radiation therapy / Oncology, Intensity Modulated Radiotherapy (IMRT), VMAT and Tomotherapy, Radiotherapy techniques
DOI: 10.1594/ranzcr2014/R-0076

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 ECR's endorsement, sponsorship or recommendation of the third party, information, product or service. ECR 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 ECR 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 and any other multimedia files are not available in the pdf version of presentations.
www.myESR.org
Aim

This poster demonstrates how DICOM data from a treatment plan can be used to compare different techniques to inform the clinical implementation of a new treatment technique. The aim is to assess dosimetric equivalence of volumetric modulated arc therapy (VMAT) compared to intensity modulated radiation therapy (IMRT), using DICOM data analysis software, prior to clinical implementation of VMAT at the Royal Hobart Hospital.
Methods and materials

A planning study was undertaken to assess the dosimetric equivalence of a new treatment planning and delivery technique (VMAT) compared to current treatment planning and delivery technique (IMRT) used in the treatment of high risk prostate cancer patients with significant risk of pelvic lymph node involvement at the Royal Hobart Hospital. Assessment of the different planning techniques (IMRT vs VMAT) was undertaken via an independent assessment of DICOM RT data exported from the Pinnacle treatment planning system and analyzed using 20 different scoring metrics generated by QualityReports (QR) software (Version 1.2.1, Sun Nuclear Corp. Florida).

Background

Treatment of prostate gland and pelvic lymph nodes has been undertaken with IMRT since 2008 at the Royal Hobart Hospital [1, 2]. The standard technique has been a 9 field evenly spaced beams (40 degree spacing, gantry range 200 > 160 degrees) plan treating the primary prostate volume to 78Gy and pelvic lymph nodes to 56-58Gy. (Figure 1) [3]

Treatment delivery with 9 field IMRT is time consuming with treatment taking approximately 12-15 minutes each day, including pre-treatment imaging.

In 2013 a VMAT capable machine was installed in Hobart. VMAT has been shown to reduce treatment delivery times significantly whilst maintaining dosimetric requirements [4]

Prior to the clinical implementation of VMAT this planning study was undertaken to compare the techniques and assess whether VMAT was dosimetrically equivalent to our current practice.

The planning study is a form of comparative effectiveness analysis [5] with scoring providing an objective numerical assessment of the plans. Previous planning studies [4-10] have shown that treatment plans equivalent or better than 3D conformal or IMRT plans can be produced with VMAT. Scoring also allows for future quality improvement to be assessed with the aim of reducing variability (e.g. between planners) and increasing the overall mean score or plan quality (Figure 2) [11]

Methods & Materials

VMAT plans were generated using a commercial treatment planning system (Pinnacle Version 9.2) for 22 patients who had been planned and approved for treatment with the
standard technique between Jan 2013 & March 2014. All treatment plans were completed and approved prior to assessment with the plan quality metric algorithm to avoid planner bias (knowledge of being scored) and selection bias.

<table>
<thead>
<tr>
<th>Patient Details</th>
<th>Cohort</th>
<th>Mean</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>67.9</td>
<td>54</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>Stage</td>
<td>T1c - T3a</td>
<td>T1c</td>
<td>T3a</td>
<td></td>
</tr>
<tr>
<td>Gleason Score</td>
<td>GS7 n= 5</td>
<td>GS8 n = 6</td>
<td>GS9 n =11</td>
<td></td>
</tr>
<tr>
<td>PSA</td>
<td>16.5</td>
<td>3.4</td>
<td>50</td>
<td></td>
</tr>
</tbody>
</table>

Table 1 Patient Cohort Details

VMAT plans used 2 arcs completing the equivalent gantry range of the 9 field standard plans (200 > 160 degrees). 2 arcs were chosen after initial investigations by the author indicated 1 arc may not be able to produce clinically acceptable plans for these patients.

Planning objectives and prescription normalization from the IMRT plan were maintained for the VMAT plan.

After completion of 2 optimization passes (60 iterations each) plans were calculated using a 0.2cm dose grid and Collapsed Cone Convolution dose algorithm.

The following DICOM data items from the completed IMRT and VMAT plans were exported from the treatment planning system

RP - Plan Prescription
RS - Plan structure set
RD - Dose grid
DICOM Image - CT Images

DICOM data was imported into QR and compared using a scoring algorithm with 20 score functions or metrics based on the clinical protocol used at the RHH [11,12] & RTOG protocols (e.g. RTOG_0815) [13]. Score functions are built by entering data points based on dose, volume, region of interest and score into a spread sheet within the QR software. This is then graphically displayed as per figures 3-8.
Scoring structure was based on Radiation Oncology Resources Plan Challenge methodology with PTV coverage taking priority followed by organs at risk in order of clinical significance [11]

Figures 9 & 10 provide a summary of the 20 score functions / metrics.

The scoring algorithm was applied to each IMRT and VMAT plan.

Population Histograms were generated for each technique for comparison. Metric score data was also exported to MS Excel to identify statistically significant differences in overall plan scores and individual score functions.
Fig. 1: Typical Nodal (Left) and Primary (Right) Target Volumes and 9 field beam arrangement

© Royal Hobart Hospital - HOBART/AU
**Fig. 2:** Aim of Continual Improvement - Higher mean performance, lower variability. (Courtesy of Greg Robinson, Sun Nuclear Corp, FL, USA)

© Greg Robinson, Sun Nuclear Corporation. Florida, USA.
Fig. 3: PTV Coverage and Maximum Dose Score Functions
© Royal Hobart Hospital - HOBART/AU

Fig. 4: PTV78 Conformation Number, Mean Dose & Homogeneity Index Score Functions
© Royal Hobart Hospital - HOBART/AU
**Fig. 5:** PTV 56 Score Functions. Note both score functions are scalable for 58Gy & 55.1Gy dose levels.

© Royal Hobart Hospital - HOBART/AU

**Fig. 6:** Rectum Score Functions

© Royal Hobart Hospital - HOBART/AU
Fig. 7: Bladder & Small Bowel Score Functions
© Royal Hobart Hospital - HOBART/AU

Fig. 8: Femur & Peripheral Tissue Score Functions
© Royal Hobart Hospital - HOBART/AU
Fig. 9: Summary of Score Functions 1-10

© Royal Hobart Hospital - HOBART/AU
Fig. 10: Summary of Score Functions 11-20

© Royal Hobart Hospital - HOBART/AU
Results

Overall scores for IMRT and VMAT plans for the 22 patients were compared using a paired t-test.

The plans optimized with VMAT resulted in a higher mean score of 128.03 vs 121.24 (Figure 11) which was statistically significant (p-value = 0.027). A higher score indicates a superior plan according to the plan quality algorithm used.

The breakdown of individual metric scores (Figure 12) shows a difference in V78 for PTV78, V56 for PTV56, D0.03 for peripheral tissue & D0.03 for PTV78. There is a statistically significant difference between these scores for VMAT vs IMRT with p-values 0.039, 0.022, 0.029 and <<0.01. (Table 2)

<table>
<thead>
<tr>
<th>Metric Number</th>
<th>Metric Title</th>
<th>IMRT Mean Score</th>
<th>VMAT Mean Score</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>[PTV78] V[78.0Gy] (%)</td>
<td>19.09091</td>
<td>24.2265</td>
<td>0.039</td>
</tr>
<tr>
<td>3</td>
<td>[PTV78] D[0.03cc] (Gy)</td>
<td>0.58371</td>
<td>1.071863</td>
<td>&lt;&lt;0.01</td>
</tr>
<tr>
<td>8</td>
<td>[PTV56] V[56.0Gy] (%)</td>
<td>5.061886</td>
<td>6.458273</td>
<td>0.022</td>
</tr>
<tr>
<td>20</td>
<td>[PERIPHERAL TISSUE] D[0.03cc] (Gy)</td>
<td>0.476234</td>
<td>0.823454</td>
<td>0.029</td>
</tr>
</tbody>
</table>

Table 2 Score breakdown where statistically significant difference (p <0.05) exists

In addition to the score differences the absolute values for 9 metrics were also statistically significant (Table 3). All favoured VMAT except the Small Bowel V60 metric which resulted in an average 0.54cc increase which is likely to be of minimal clinical significance.

<table>
<thead>
<tr>
<th>Metric Number</th>
<th>Metric Title</th>
<th>Absolute difference (VMAT-IMRT)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>[PTV78] V[78.0Gy] (%)</td>
<td>2.247741</td>
<td>&lt;&lt;0.01</td>
</tr>
</tbody>
</table>
3  | [PTV78] D[0.03cc] (Gy)  | -0.82953 | 0.015
7  | [PTV78] Homogeneity Index [78.0Gy]  | -0.01507 | <<0.01
8  | [PTV56] V[56.0Gy] (%)  | 3.275914 | 0.012
9  | [PTV56] V[53.2Gy] (%)  | 0.204186 | <<0.01
10 | [RECTUM] V[65.0Gy] (%)  | -0.23033 | 0.048
11 | [RECTUM] V[40.0Gy] (%)  | -1.16282 | 0.012
17 | [SMALL BOWEL] V[60.0Gy] (cc)  | 0.540018 | 0.035
20 | [PERIPHERAL TISSUE] D[0.03cc] (Gy)  | -3.49417 | <<0.01

Table 3 Metrics where statistically significant difference (p<0.05) in raw data exists

Population distributions demonstrate the spread of scores. Figure 13* demonstrates the higher mean score for VMAT plans with the distribution of scores shifted slightly to the right. Minimum and maximum scores are similar for the two techniques.

Figure 14 shows the 5 highest scoring IMRT plans are all clustered together with a slight degradation of scores in the VMAT plans due to using the same planning objectives. This could be potentially solved by modifying the planning objectives and re-optimizing the VMAT plans. This is an area for potential improvement that would increase the mean score (i.e. improve overall plan quality)

Similarly the 5 lowest scoring IMRT plans are highlighted in Figure 15. The VMAT plans demonstrate a greater spread, with higher scores from improved coverage adding to overall higher mean score. Re-optimization could potentially improve plan scores and overall plan quality.

*Histogram's bin width = 0.2 x optimal (Scott's Rule [15] suggests 0.5 x optimal but 0.2 chosen for best visual display of the 22 data points)
Fig. 11: Comparison of IMRT & VMAT Mean Scores

© Royal Hobart Hospital - HOBART/AU
Fig. 12: Score comparison per metric

© Royal Hobart Hospital - HOBART/AU

Fig. 13: Population Histogram Distribution IMRT (Blue) vs VMAT (Red)

© Royal Hobart Hospital - HOBART/AU
Fig. 14: 5 Highest Scoring Plans for IMRT (Blue) & Corresponding VMAT (Red) Scores

© Royal Hobart Hospital - HOBART/AU
**Fig. 15:** 5 Lowest Scoring Plans for IMRT (Blue) & Corresponding VMAT (Red) Scores

© Royal Hobart Hospital - HOBART/AU
Conclusion

This study has demonstrated that VMAT plans for this patient cohort using 2 arcs with the same range of gantry motion & plan objectives are at least equivalent to our current IMRT plans, suggesting planners could use the same strategies to create a clinically acceptable VMAT plan.

The results indicate that the VMAT plans are statistically better with a higher overall mean score; are better with regard to prostate gland and nodal volume coverage and also resulted in lower peripheral and overall maximal dose.

The scores for the PTV78 Volume receiving 78Gy and PTV56 volume receiving 56Gy metrics indicate that the IMRT plans may have been under-normalized [16]. The improvement in these metrics (p-values 0.039 & 0.022) indicates that VMAT is able to assist with this component of the plans. This is further illustrated by the average cumulative DVH (Figure 16) which shows minimal differences between the two delivery techniques. Differences are evident where the metrics showed statistical significance; at the shoulder of the PTV78 & PTV56 DVH lines, indicating improved coverage with VMAT. This equates to approximately 2-3% difference in coverage based on the raw data (see Results Table 2). Discussion with planning RTs indicates this may be a trade-off between coverage and organ-at-risk dose for the clinical plans. In addition, the difference in PTV78 dose to 0.03cc (p-value <<0.01) indicates the VMAT plans were able to achieve a lower overall maximum dose (approximately 0.8Gy lower) whilst improving coverage, as shown by the tail of the PTV78 line. This is supported by the improved homogeneity index (p <<0.01).

Rectal doses at both V65 and V40 were also statistically significant (p-values 0.048 & 0.012) which is most evident in Figure 16 at the V40 level. This equates to a reduction in percentage volume of approximately 1% with VMAT. Similarly VMAT was able to reduce the peripheral tissue dose (p-value 0.029) (not shown). Small bowel V60 however was slightly increased with VMAT (approximately 0.5cc, p-value = 0.035)

This project was designed to assess the equivalence of the two techniques whilst maintaining as much similarity as possible by using the same gantry range and plan objectives. It is likely the IMRT objectives used would not produce an optimal VMAT plan. This is further assessed in a second poster entitled Using population histogram distributions to guide training for IMRT treatment planning. Other advantages of VMAT are reduced treatment delivery time [6-10].
This study also demonstrates the advantages of an independent treatment plan assessment tool and its utility in performing comprehensive analysis of treatment plan objectives when comparing different planning techniques. Population histograms generated in this fashion could be used for guidance of future practice, building a library of cases that can be used to develop protocols or class solutions and inform knowledge based planning.
Fig. 16: Average Cumulative DVH. NB PTV56 line includes 18 patients treated to this dose level. PTV58 (4 patients) not shown but values for these patients included in PTV78 and organs at risk

© Royal Hobart Hospital - HOBART/AU
15. Scott, David (1992), Multivariate Density Estimation: Theory, Practice, and Visualization, John Wiley and Sons