A comparison study of three planar IMRT QA methods

Introduction: Since the introduction of intensity modulated radiation therapy (IMRT), a large number of quality assurance (QA) methods have been developed to test the accuracy of IMRT treatments. One of the most widely used patient-specific IMRT QA techniques is 2D dosimetric comparison between planning and measurement. The 2D comparison is accomplished by ‘composite’ delivery of an entire treatment plan and field-by-field (FBF; per-beam) exposure using film and 2D array detectors. The composite delivery is further categorized as patient-gantry-angle composite (PGAC) and single-gantry-angle composite (SGAC). In the SGAC technique, all fields in a plan are delivered with a single gantry angle and the sum of all the fields are compared with an analogous calculated treatment plan. Nelms and Simon\(^1\) showed that 32.8 % of institutions are still using the SGAC method for 75-100% of their patients and 64.1% are using field by field (FBF; per-beam) analysis for all of their patients. Even if these QA methods are widespread, there are few extensive comparison studies on the relevance of the different techniques. The goal of this research is to investigate comparability of the three different planar IMRT QA methods using a common conventional IMRT QA performance metric, the gamma passing rate\(^2\) and improve the understanding of the current QA methods.

Methods and Materials: IMRT verification plans were generated by the Varian Eclipse treatment planning system (TPS) for twelve heavily modulated head and neck (H&N) and eleven moderately modulated prostate treatment plans retrospectively selected for this study as summarized in Table 1. The Varian TrueBEAM STX\(^\text{TM}\) with high-definition MLC (HD120; 2.5 mm and 5 mm leaf width) was used to deliver the IMRT QA plans. Verification of IMRT delivery by Radiological Physics Center (RPC) showed that the ratio of RPC value/our value of IMRT delivery was found to be less than 2% for PTV and for OAR 0 mm (the criteria for acceptance set by RPC: less than 7% for PTV and 4 mm for OAR).

For each of these patients, three different IMRT QA methods were performed using Sun Nuclear Mapcheck2 (a diode array containing 1527 diode detectors) sandwiched with water equivalent MapPhan-MC2 of a 5 cm buildup to the detectors as shown in Figure 1. The gamma criteria of 1%-1mm (dose difference and distance-to-agreement (DTA)\(^3\)), 2%-2mm, and 3%-3mm were used to determine the percentage of points passed for each verification plan. Absolute dose comparison (using absolute dose data; not normalized to a selected dose point such as a point of the maximum dose) with 10% threshold (a threshold used to exclude low dose points from appearing to fail) was adopted for the QA analysis as shown in Figure 2. For the FBF analysis, the overall passing rate for each plan was the weighted average (the weighting based on the number of MUs for the each field) of the passing rates for the individual fields. Statistical analysis (Student’s t-test) was employed to examine the agreement between two different methods: SGAC vs. PGAC (C1), FBF vs. SGAC (C2), and FBF vs. PGAC (C3).

Results: The average passing rates of the gamma test were shown in Table 2. For the 12 H&N cases, they were 99.0 (1.7)% (PGAC), 99.5 (0.6)% (SGAC), and 99.6 (0.3)% (FBF) with 3%-3mm criteria. The p-values of the t-test with 1%-1mm/2%-2mm/3%-3mm were 0.43/0.46/0.35 (C1), 0.10/0.06/0.83 (C2), and 0.92/0.90/0.29 (C3) as shown in Table 3. For the 11 prostate cases, the average passing rates were 99.7 (0.5)% (PGAC), 100.0 (0.0)% (SGAC), and 100.0 (0.1)% (FBF) with 3%-3mm criteria. The p-values of 1%-1mm/2%-2mm/3%-3mm were 0.89/0.17/0.10 (C1), 0.54/0.60/0.19 (C2), and 0.85/0.16/0.14 (C3). Based on the statistical analysis and the average pass rates, it was determined that all three IMRT QA methods had comparable results with the various criteria of the QA performance in 95% confidence level for both H&N and prostate cases.

Conclusion: All three IMRT QA methods were found to have comparable passing rates with different gamma criteria. This indicates that any of these three methods can be used for routine IMRT QA in a clinic. However, using the tighter gamma criteria such as 1%-1mm, the passing rate for each patient was significantly different among the three QA techniques for both H&N and prostate cases (for instance, H&N patient #8 had the maximum difference of 32.1%). Therefore, considering uncertainty of QA measurements, gamma criteria of 3%-3mm or higher are recommended for the QA analysis to produce meaningful and consistent results as suggested by other groups.\(^4,5\)

References:
\(^4\)M. Stock, B. Kroupa, and D. Georg, “Interpretation and evaluation of the \(\gamma\) index and the \(\gamma\) index angle for the verification of IMRT hybrid plans,” Phys Med Biol 50, 399-411 (2009)
Figure 1. IMRT QA measurement setup using Mapcheck® and MapPHAN®.

Table 1. Summary of IMRT patients

<table>
<thead>
<tr>
<th>Patient #</th>
<th># of beams</th>
<th>Total MUs</th>
<th>Treatment site</th>
<th># of beams</th>
<th>Total MUs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12</td>
<td>1087</td>
<td>Nasopharynx</td>
<td>7</td>
<td>494</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>879</td>
<td>Pharynx</td>
<td>7</td>
<td>388</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>583</td>
<td>Esophagus</td>
<td>9</td>
<td>472</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>500</td>
<td>Rt Parotid</td>
<td>9</td>
<td>317</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>420</td>
<td>Esophagus</td>
<td>9</td>
<td>550</td>
</tr>
<tr>
<td>6</td>
<td>18</td>
<td>1077</td>
<td>Tongue</td>
<td>9</td>
<td>366</td>
</tr>
<tr>
<td>7</td>
<td>12</td>
<td>808</td>
<td>Larynx</td>
<td>9</td>
<td>438</td>
</tr>
<tr>
<td>8</td>
<td>12</td>
<td>877</td>
<td>Larynx</td>
<td>7</td>
<td>572</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>507</td>
<td>Esophagus</td>
<td>7</td>
<td>387</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>337</td>
<td>Rt Parotid</td>
<td>8</td>
<td>433</td>
</tr>
<tr>
<td>11</td>
<td>14</td>
<td>810</td>
<td>Tonsil</td>
<td>7</td>
<td>363</td>
</tr>
<tr>
<td>12</td>
<td>10</td>
<td>775</td>
<td>Tonsil</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Note: split beams are separately included in the number of beams.

Figure 2. 2D dose comparison of H&N patient #1 in Table 1 using the Sun Nuclear Mapcheck diode array detector: (a) PGAC, (b) SGAC, and (c) FBF (one out of 12 fields shown). The gamma test with ‘AD’ (absolute dose comparison) and 10% threshold was used for the analysis. Blue dots show measurement is lower than calculation and red means measurement is higher.

Table 2. Passing rates of the gamma test for the three IMRT QA techniques with the different criteria (dose difference/DTA).

<table>
<thead>
<tr>
<th>Patient #</th>
<th>1%/1 mm</th>
<th>2%/2 mm</th>
<th>3%/3 mm</th>
<th>1%/1 mm</th>
<th>2%/2 mm</th>
<th>3%/3 mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>H&amp;N</td>
<td>PGAC</td>
<td>SGAC</td>
<td>FBF</td>
<td>PGAC</td>
<td>SGAC</td>
<td>FBF</td>
</tr>
<tr>
<td>1</td>
<td>70.8</td>
<td>45.0</td>
<td>62.8</td>
<td>94.9</td>
<td>82.3</td>
<td>91.7</td>
</tr>
<tr>
<td>2</td>
<td>76.4</td>
<td>56.4</td>
<td>67.5</td>
<td>95.6</td>
<td>89.5</td>
<td>94.0</td>
</tr>
<tr>
<td>3</td>
<td>87.1</td>
<td>89.3</td>
<td>80.5</td>
<td>99.3</td>
<td>98.2</td>
<td>96.9</td>
</tr>
<tr>
<td>4</td>
<td>42.7</td>
<td>63.4</td>
<td>64.3</td>
<td>77.8</td>
<td>92.7</td>
<td>91.4</td>
</tr>
<tr>
<td>5</td>
<td>79.7</td>
<td>65.4</td>
<td>60.8</td>
<td>98.2</td>
<td>91.2</td>
<td>92.2</td>
</tr>
<tr>
<td>6</td>
<td>76.3</td>
<td>58.7</td>
<td>74.9</td>
<td>98.9</td>
<td>92.4</td>
<td>94.6</td>
</tr>
<tr>
<td>7</td>
<td>64.8</td>
<td>78.6</td>
<td>80.0</td>
<td>96.1</td>
<td>96.9</td>
<td>96.9</td>
</tr>
<tr>
<td>8</td>
<td>82.0</td>
<td>49.9</td>
<td>67.7</td>
<td>97.5</td>
<td>92.8</td>
<td>96.9</td>
</tr>
<tr>
<td>9</td>
<td>56.1</td>
<td>77.2</td>
<td>73.9</td>
<td>97.6</td>
<td>94.7</td>
<td>95.1</td>
</tr>
<tr>
<td>10</td>
<td>50.6</td>
<td>66.9</td>
<td>66.2</td>
<td>83.4</td>
<td>92.6</td>
<td>93.5</td>
</tr>
<tr>
<td>11</td>
<td>77.7</td>
<td>72.8</td>
<td>73.4</td>
<td>98.4</td>
<td>95.6</td>
<td>96.0</td>
</tr>
<tr>
<td>12</td>
<td>75.3</td>
<td>62.8</td>
<td>72.4</td>
<td>97.6</td>
<td>89.3</td>
<td>95.2</td>
</tr>
</tbody>
</table>

Prostate   | PGAC    | SGAC    | FBF     | PGAC    | SGAC    | FBF     |
| 1         | 63.9    | 60.3    | 80.7    | 97.2    | 88.4    | 99.4    |
| 2         | 62.8    | 81.9    | 82.9    | 97.9    | 100.0   | 99.5    |
| 3         | 72.9    | 78.7    | 68.0    | 96.6    | 100.0   | 98.8    |
| 4         | 64.2    | 83.0    | 72.2    | 93.5    | 100.0   | 97.8    |
| 5         | 61.9    | 74.4    | 71.0    | 89.8    | 96.3    | 93.8    |
| 6         | 57.3    | 68.8    | 70.4    | 93.3    | 97.5    | 96.1    |
| 7         | 59.1    | 66.7    | 60.3    | 96.2    | 98.6    | 97.1    |
| 8         | 82.5    | 68.8    | 64.1    | 99.7    | 100.0   | 99.2    |
| 9         | 74.2    | 60.6    | 59.5    | 96.1    | 100.0   | 98.8    |
| 10        | 84.0    | 59.8    | 60.2    | 100.0   | 99.2    | 95.9    |
| 11        | 75.2    | 62.0    | 58.4    | 97.8    | 98.7    | 95.2    |
| 12        | 75.3    | 62.8    | 72.4    | 97.6    | 89.3    | 95.2    |

Mean       | 70.0    | 65.5    | 70.4    | 94.3    | 92.6    | 94.3    |
| Std       | 13.6    | 12.6    | 6.5     | 6.8     | 4.3     | 2.0     |

Table 3. The p-values based on the Student’s t-test. If the p-value is greater than 0.05, it is concluded that there is no significant difference between two methods in 95% confidence level.

<table>
<thead>
<tr>
<th>H&amp;N</th>
<th>1%/1 mm</th>
<th>2%/2 mm</th>
<th>3%/3 mm</th>
<th>Prostate</th>
<th>1%/1 mm</th>
<th>2%/2 mm</th>
<th>3%/3 mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>PGAC-SGAC</td>
<td>0.432</td>
<td>0.098</td>
<td>0.919</td>
<td>PGAC-SGAC</td>
<td>0.460</td>
<td>0.060</td>
<td>0.901</td>
</tr>
<tr>
<td>SGAC-FBF</td>
<td>0.173</td>
<td>0.603</td>
<td>0.155</td>
<td>PGAC-FBF</td>
<td>0.347</td>
<td>0.826</td>
<td>0.294</td>
</tr>
<tr>
<td>PGAC-FBF</td>
<td>0.887</td>
<td>0.543</td>
<td>0.846</td>
<td>PGAC-SGAC</td>
<td>0.097</td>
<td>0.192</td>
<td>0.137</td>
</tr>
</tbody>
</table>