Current IMRT QA: Pros and Cons

Stephen Kry

AAPM
July, 2020
Current IMRT QA

• Patient specific, pre-treatment measurements
• The current standard of care for QA of IMRT treatments
• Lots of devices, methods, analysis, interpretation, etc.
• Discuss pros and cons of this general philosophical approach
Pro - 1

• Completes the link between what is planned and what is delivered.
  – Verify that the intended dose is delivered
    • It is the delivered dose that will determine outcomes
    • There are many steps between the TPS image and the delivered dose
  – Verify the deliverability of the plan
    • Dry run of treatment streamlines patient treatment
There are errors to be caught! IMRT QA is detection opportunity

- IROC phantom data:
  - 10-17% of results fail loose tolerance. (Carson 2016; Edward 2020)
  - Beam modeling shortcomings have been identified in the majority of these cases. (Kerns 2017, Edward 2020)
- Detailed clinical series have also identified errors (Mans 2010)
  - Data transfer
  - Accidental plan modification
  - Suboptimal beam modeling
Pro-3

- Good measurement systems have the potential to detect a lot of failure modes
  - Not just calculation errors or delivery errors
  - Techniques like EPID transmission dosimetry can identify:
    - Anatomical changes
    - Patient setup errors
    - (Mans 2010, Olaciregui-Ruiz 2019)
Pro-4

• Well established
  – There is a long history of this approach to verifying complex treatments
  – Lots of available guidance in terms of literature and TG reports
  – Lots of available community experience
  – Don’t have to invent anything
• Clear cases of value
  – QA has caught errors, caused interventions
    • Dosimetric disagreement, deliverability, anatomical changes, data transfer errors.
      • (Mans 2010, Pulliam 2014)
  – Can highlight opportunity for improved planning techniques
    • (Letourneau 2013)
Cons-1

- Time consuming
  - Spend a lot of hours on this task
  - These are unpleasant hours as they are evenings/weekends
  - This task often falls on highly paid highly educated physicists

- Acceptable if it’s time well spent
Cons-2

- Incomplete evaluation
  - Rarely assessing dose in patient geometry
    - Geometrical array
  - Rarely assessing dose in heterogeneous environment
    - Just on array surface
    - Even calculations into patient anatomy often use simplistic dose calculation algorithms
  - Rarely capture dose in framework for clinical interpretation
    - No DVH info, hard to relate %pixels passing to clinical judgements
  - Anatomy changes during treatment
Cons-3

- We usually don’t act on failures
  - When IMRT QA fails, we usually repeat measurements and repeat until we get a passing result (Pulliam 2014)

MD Anderson experience of 301 failed ion chamber-based IMRT QA results
Cons-3 Continued

• We usually don’t act on failures
  – Survey of 1,455 institutions highlighted similar results (Mehrens 2020)
  – Main approach to failing IMRT QA is re-measure
    • Other strategies: use relative mode, change passing criteria, replan
    • Most approaches: make the current situation work

• It is understandable that we don’t do much with these issues
  – Not an easy solution (replanning is a lot of work)
  – IROC shows lots of errors originate with beam model. This isn’t the time to be fixing a beam model…..
  – Patient already on table

• If we don’t act on problems, why are we doing this??
  – Current approach is not working
Cons-4

- Devices don’t catch errors
  - Dosimetrically unacceptable plans are called fine
  - Low sensitivity, high specificity
    - Doesn’t fail bad plans, but doesn’t fail good plans. Doesn’t fail anything!!!
    - IROC phantom results not predicted by inst. IMRT QA (Kry 2019)

<table>
<thead>
<tr>
<th>Device</th>
<th># Tests</th>
<th># Poor</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>337</td>
<td>59</td>
<td>5 (3 identified)</td>
<td>99</td>
</tr>
<tr>
<td>MapCheck</td>
<td>121</td>
<td>20</td>
<td>5 (1 identified)</td>
<td>100</td>
</tr>
<tr>
<td>ArcCheck</td>
<td>93</td>
<td>16</td>
<td>0 (0 identified)</td>
<td>100</td>
</tr>
<tr>
<td>EPID</td>
<td>58</td>
<td>16</td>
<td>0 (0 identified)</td>
<td>100</td>
</tr>
<tr>
<td>Ion chamber</td>
<td>44</td>
<td>8</td>
<td>25 (2 identified)</td>
<td>94</td>
</tr>
</tbody>
</table>
Cons-4

• Devices don’t catch errors
  – This phantom result:
    • 8% systematic underdose
    • ArcCheck QA, 3%/3mm, absolute dose mode
    • 97% and 100% of pixels passed

• Also don’t necessarily catch patient issues
  – Per-Fraction: translations of 2 cm before detected (Hseih 2017)
  – EPID: Translations of 1 cm not always well detected (Olaciregui-Ruiz 2019)
  – Rotations less than 8 degrees not well detected (Olaciregui-Ruiz 2019)
Clinical QA thresholds are unrealistic

- Even TG-218 suggested criteria don’t appear to be adequate
- To detect 80% of poor or failing IROC phantom results: (Kry 2019)

<table>
<thead>
<tr>
<th>Criteria</th>
<th>IROC phantom result</th>
<th>Threshold (% pixels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3%/3mm</td>
<td>Fail (&gt;7% error)</td>
<td>99.7</td>
</tr>
<tr>
<td>2%/2mm</td>
<td>Fail (&gt;7% error)</td>
<td>100</td>
</tr>
<tr>
<td>3%/3mm</td>
<td>Poor (&gt;5% error)</td>
<td>99.8</td>
</tr>
<tr>
<td>2%/2mm</td>
<td>Poor (&gt;5% error)</td>
<td>99.2</td>
</tr>
</tbody>
</table>

- These criteria are not clinically implementable
Summary

• Conceptually current IMRT QA is very important, very robust technique to probe plan, delivery, even the patient.

But, overwhelmingly,

• Methods/devices don’t actually work well
  – Usually don’t catch errors
• Program of IMRT QA doesn’t work well
  – Even when IMRT QA indicates a problem, we don’t/can’t act on it.

• We need to improve on the current status of IMRT QA
END