TG218- Measurement Methods and Tolerance Levels for Patient-Specific IMRT QA

Moyed Miften, Ph.D.
Professor and Chief Physicist

Department of Radiation Oncology
University of Colorado School of Medicine

TG218 Members

Patient-Specific IMRT Verification QA Measurement

- Designed to identify discrepancies between planned and delivered doses
- Detect gross errors in the radiation delivery
- Minimizes reliance on the concept that all potential sources of error in the IMRT process are known, characterized, and contained
- Ensuring the safety of patient, fidelity of treatment, and that the patient receives the desired treatment plan

Patient Specific IMRT QA Guidance Documents

Patient specific pre-treatment quality assurance (QA)	Because of the complexity of IMRT planning and delivery, pre-treatment patient-specific quality assurance has been recommended in guidance documents from ASTRO, ACR, and AAPM.[15,19,26,35]

Perform or oversee the pre-treatment quality assurance checks including:

a. Verify integrity of the information transferred to the treatment management system for the patient plan and the QA plan, including correct transfer of gantry, collimator, table, and jaw positions, and calculated monitor units etc.
b. Verify correctness of MLC leaf positions, sequences, and fractional monitor units
c. Verify the accuracy of monitor units used for the patient dose calculation

ASTRO's safety white paper on IMRT

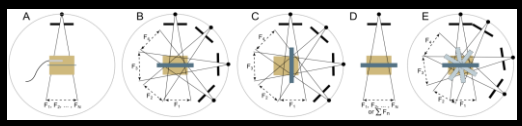
Why TG218

- There is little systematic guidance on patient-specific IMRT verification QA
- There are no discussion on the pros and cons of the different delivery methods for QA measurements
- How to assess the clinical relevance of failed IMRT plans
- What are the course of actions a clinical physicist can undertake to deal with failed patient-specific IMRT QA plans
- Radiation oncology clinics have developed their own patient-specific IMRT QA procedures
- QA procedures differ in scope and depth, acceptable tolerance levels, delivery methods, verification tools, analysis methodologies, and the type of verified calculation vs. measured data

TG218 Charge

- To review literature and reports containing data on the achieved agreement between measurements and calculations for IMRT, VMAT, and tomotherapy techniques.
- To review commonly used measurement methods: composite of all beams using the actual treatment parameters, perpendicular composite, and perpendicular field-by-field. Discuss pros and cons of each method.
- To review methodologies for absolute dose verification with ion-chamber and 2D detector arrays
- To investigate the dose-difference/DTA and γ verification metrics, their use and vendor-implementation variability, including the choice of various parameters used to perform the IMRT QA analysis.

Delivery Methods



True Composite (film & chamber)	True Composite (Device in coronal direction)	True Composite (Device in sagittal direction)	Field-by-Field OR Composite ALL Fields Summed (gantry @ 0º)	Composite ALL Fields Summed (device perpendicular to gantry)

Perpendicular Field-by-Field (PFF)

- The radiation beam is perpendicular to the plane of the measurement device
- The device can be placed on the couch or attached to the gantry head.
- The dose from each of the IMRT beams is delivered and analyzed.

Perpendicular Composite (PC)

- The radiation beam is always perpendicular to the measurement device detector plane.
- The device can be placed on the couch or attached to the gantry head.
- The doses from all IMRT radiation beams are delivered and subsequently summed.

True Composite (TC)

- All of the radiation beams are delivered to a stationary measurement device in a phantom placed on the couch using the actual treatment beam geometry for the patient.
- This method most closely simulates the treatment delivery to the patient.

Delivery Methods: Pros

- PFF and PC: Every part of every field is sampled, fast acquisition.
- PC: only one dose image to analyze. More uniform dose for analysis than PFF.
- TC: provide an actual dose summation in a 2D slice of the 3D dose. Only one dose image to analyze.

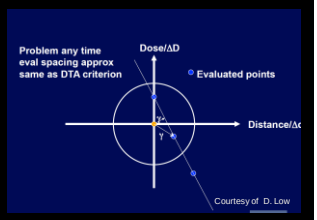
Adapted from A. Olch

Delivery Methods: Cons

- PFF, PC: no 3D summation. Can't know significance of regional errors in each beam.
- PFF, PC: can get any γ result you want for relative dose mode by normalizing to a different place.
- PC: errors from each field may cancel on summation.
- TC: Does not sample every part of each beam.

Dose difference, DTA, and γ analysis

- Practical considerations
  - Normalization
  - Spatial resolution
  - Interpretation



γ IMRT QA Evaluation

- 100% passing is ideal but not practical
- γ statistics should be checked in a structure by structure basis.
- γ tool should be used as an indicator of problems, not as a single indicator of plan quality
- Clinical interpretation of failure results is a challenging QA process
- Quality measures are intended to set a requirement for the performance of a system

Adapted from D. Low

Vendor Algorithm Testing: γ Passing Rates

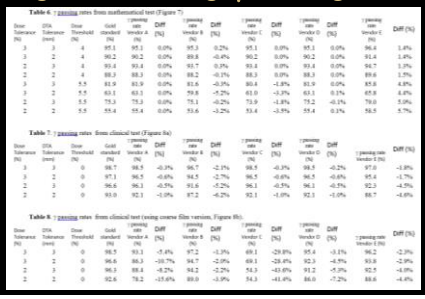


Table 6: γ passing rates from mathematical test (Figure 7)

Dose Tolerance (%)	DTA Tolerance (mm)	Dose Threshold (%)	Gold standard (%)	γ passing rate Vendor A (%)	Diff (%)	γ passing rate Vendor B (%)	Diff (%)	γ passing rate Vendor C (%)	Diff (%)	γ passing rate Vendor D (%)	Diff (%)	γ passing rate Vendor E (%)	Diff (%)
3	3	4	95.1	95.1	0.0%	95.3	0.2%	95.1	0.0%	95.1	0.0%	96.4	1.4%
3	2	4	90.2	90.2	0.0%	89.8	-0.4%	90.2	0.0%	90.2	0.0%	91.4	1.4%
2	3	4	93.4	93.4	0.0%	93.7	0.3%	93.4	0.0%	93.4	0.0%	94.7	1.3%
2	2	4	88.3	88.3	0.0%	88.2	-0.1%	88.3	0.0%	88.3	0.0%	89.6	1.3%
3	3	5.5	81.9	81.9	0.0%	81.6	-0.3%	80.4	-1.5%	81.9	0.0%	85.8	4.8%
3	2	5.5	63.1	63.1	0.0%	59.8	-5.2%	61.0	-3.3%	63.1	0.1%	65.8	4.4%
2	3	5.5	75.3	75.3	0.0%	75.1	-0.2%	73.9	-1.8%	75.2	-0.1%	79.0	5.0%
2	2	5.5	55.4	55.4	0.0%	53.6	-3.2%	53.4	-3.5%	55.4	0.1%	58.5	5.7%

Table 7: γ passing rates from clinical test (Figure 8a)

Dose Tolerance (%)	DTA Tolerance (mm)	Dose Threshold (%)	Gold standard (%)	γ passing rate Vendor A (%)	Diff (%)	γ passing rate Vendor B (%)	Diff (%)	γ passing rate Vendor C (%)	Diff (%)	γ passing rate Vendor D (%)	Diff (%)	γ passing rate Vendor E (%)	Diff (%)
3	3	0	98.7	98.5	-0.3%	96.7	-2.1%	98.5	-0.3%	98.5	-0.2%	97.0	-1.8%
3	2	0	97.1	96.5	-0.6%	94.5	-2.7%	96.5	-0.6%	96.5	-0.6%	95.4	-1.7%
2	3	0	96.6	96.1	-0.5%	91.6	-5.2%	96.1	-0.5%	96.1	-0.5%	92.3	-4.5%
2	2	0	93.0	92.1	-1.0%	87.2	-6.2%	92.1	-1.0%	92.1	-1.0%	88.7	-4.6%

Table 8: γ passing rates from clinical test (using coarse film version, Figure 8b).

Dose Tolerance (%)	DTA Tolerance (mm)	Dose Threshold (%)	Gold standard (%)	γ passing rate Vendor A (%)	Diff (%)	γ passing rate Vendor B (%)	Diff (%)	γ passing rate Vendor C (%)	Diff (%)	γ passing rate Vendor D (%)	Diff (%)	γ passing rate Vendor E (%)	Diff (%)
3	3	0	98.5	93.1	-5.4%	97.2	-1.3%	69.1	-29.8%	95.4	-3.1%	96.2	-2.3%
3	2	0	96.6	86.3	-10.7%	94.7	-2.0%	69.1	-28.4%	92.3	-4.5%	93.8	-2.9%
2	3	0	96.3	88.4	-8.2%	94.2	-2.2%	54.3	-43.6%	91.2	-5.3%	92.5	-4.0%
2	2	0	92.6	78.2	-15.6%	89.0	-3.9%	54.3	-41.4%	86.0	-7.2%	88.6	-4.4%

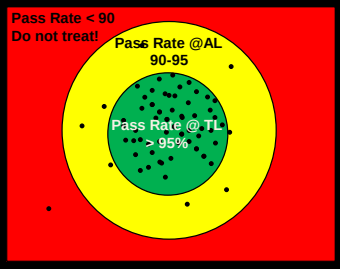
Action Limits (ALs)

- Quality measures (QMs) → set a requirement for the performance of IMRT QA
- Action Limits
  - → degree to which the QMs are allowed to vary
  - → thresholds for when an action is required
  - → based on clinical judgment
    - acceptability of a certain level of deviation from a QM

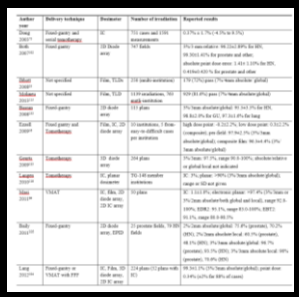
Tolerance Limits (TLs)

- TLs → boundary within which a process is considered to be operating normally
- Measurements outside of a TL provide a warning that a system is deviating
  - – investigate to see if an issue can be identified and fixed
- Intent → fix issues before they become a clinical problem (i.e. data outside of ALs)

What Should We Expect?



Literature Review



TG218 Recommendations

- IMRT QA measurements should be performed using TC
  - QA device has negligible angular dependence or the angular dependence is accurately accounted for in the vendor software.
- IMRT QA measurements should be performed using PFF if the QA device is not suitable for TC measurements, or for TC verification error analysis.
- IMRT QA measurements should not be performed using PC which is prone to masking delivery errors.

TG218 Recommendations

- Analysis of IMRT QA measurement and plan should be performed in absolute dose mode, not relative dose.
- A dose calibration measurement compared against a standard dose should be performed before each measurement session
  - factor the variation of the detector response and accelerator output into the IMRT QA measurement.
- Global normalization should be used. Global normalization is deemed more clinically relevant than local normalization.
  - global normalization point should be selected whenever possible in a low gradient region with a value that is ≥ 90% of the maximum dose in the plane of measurement.

Recommendations

- Local normalization is more stringent than global normalization for routine IMRT QA.
  - It can be used during the IMRT commissioning process and for troubleshooting IMRT QA.
- Dose threshold should be set to exclude low dose areas that have no or little clinical relevance but can bias the analysis.
  - setting the threshold to 10% in a case where the OAR dose tolerance exceeds 10% of the prescription dose.
  - allows the γ passing rate analysis to ignore the large area of dose points that lie in very low dose regions which, if included, would increase the passing rate

Recommendations

- Tolerance limits: the γ passing rate should be ≥ 95%, with 3%/2mm and a 10% dose threshold.
- Action limits: the γ passing rate should be ≥ 90%, with 3%/2mm and a 10% dose threshold.
  - If the plan fails this AL, evaluate the γ failure distribution and determine if the failed points lie in regions where the dose differences are clinically irrelevant
  - If the γ failure points are distributed throughout the target or critical structures and are at dose levels that are clinically relevant, the plan should not be used
  - It may be necessary to review results with a different detector or different measurement geometry

Recommendations

- For any case with γ passing rate < 100%,
  - the γ distribution should be carefully reviewed rather than relying *only* on distilled statistical evaluations
  - review of γ results should not be limited to only the %points that fail, but should include other relevant γ values
  - an analysis of the maximum γ value and the %points that exceed a γ value of 1.5 should be performed.
  - For a 3%/2 mm, a γ value of 1.5 could indicate a dose diff of 4.5% in a shallow dose gradient region or a DTA of ~3.0 mm in a steep dose gradient region.

Recommendations

- Reviewing dose differences directly without γ or using local dose normalization and tighter dose difference/DTA criteria.
- γ should be reviewed on a structure by structure basis
- Track γ passing rates across patients and for the same tumor sites to look for systematic errors in the system.
- Software tools that compare measured and calculated DVHs of structures are preferred over analysis in phantoms.
- DVH analysis can be used to evaluate the clinical relevance of QA results, especially when the γ passing rate fails the tolerance limits or is inconsistent.

Steps to Check Marginal/Failed IMRT QA

- Phantom/device setup
- Beam characteristics
- MLC
- TPS

Summary

- Each IMRT QA method has advantages and disadvantages with variable ability to identify delivery-to-plan differences.
- True composite provides at least a 2D plane out of a 3D dose distribution
- None of the methods provide us the error in delivery of the 3D dose to the patient's PTV or critical organs.
- Deriving clinical indications from failing points is challenging
- Defining IMRT tolerance and action levels improve the IMRT QA process

Thank You

CU Anschutz Medical Campus