An Overview of Patient-Specific Quality Assurance

Vimal Desai, Ph.D.

Patient-Specific QA Devices/Software

American Association of Physicists in Medicine Virtual Meeting

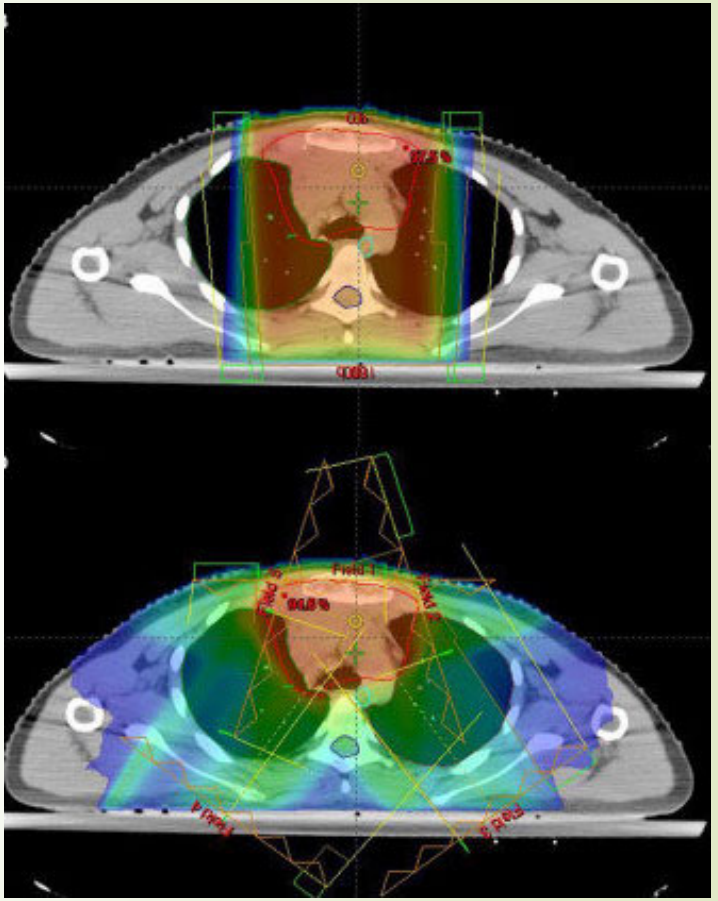
07/25/2021

Disclosures

- In this talk, several manufacturer-specific solutions are presented for educational purposes but are NOT explicitly endorsed by myself or the AAPM
- I previously delivered a webinar for ScandiDos but am not affiliated with ScandiDos

Motivation

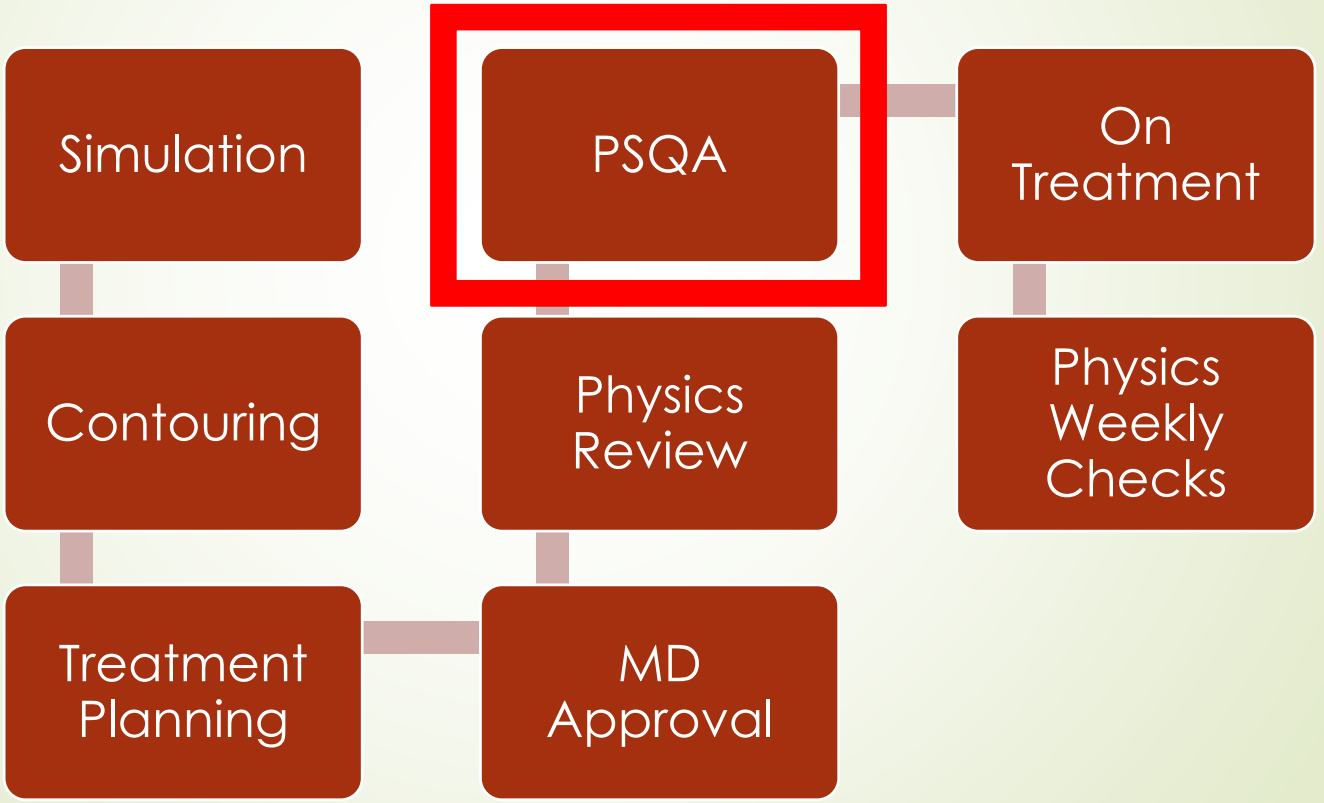
- Modern external beam radiotherapy is complex
  - Varying degrees of modulation to achieve more conformal dose distributions
  - Modulation usually achieved with multi-leaf collimator (MLC)
  - Assumptions and simplifications of beam modeling have larger impact on dose calculation accuracy in IMRT
  - Independent MU calculations and dose verification more important for fields heavily modulated using MLCs



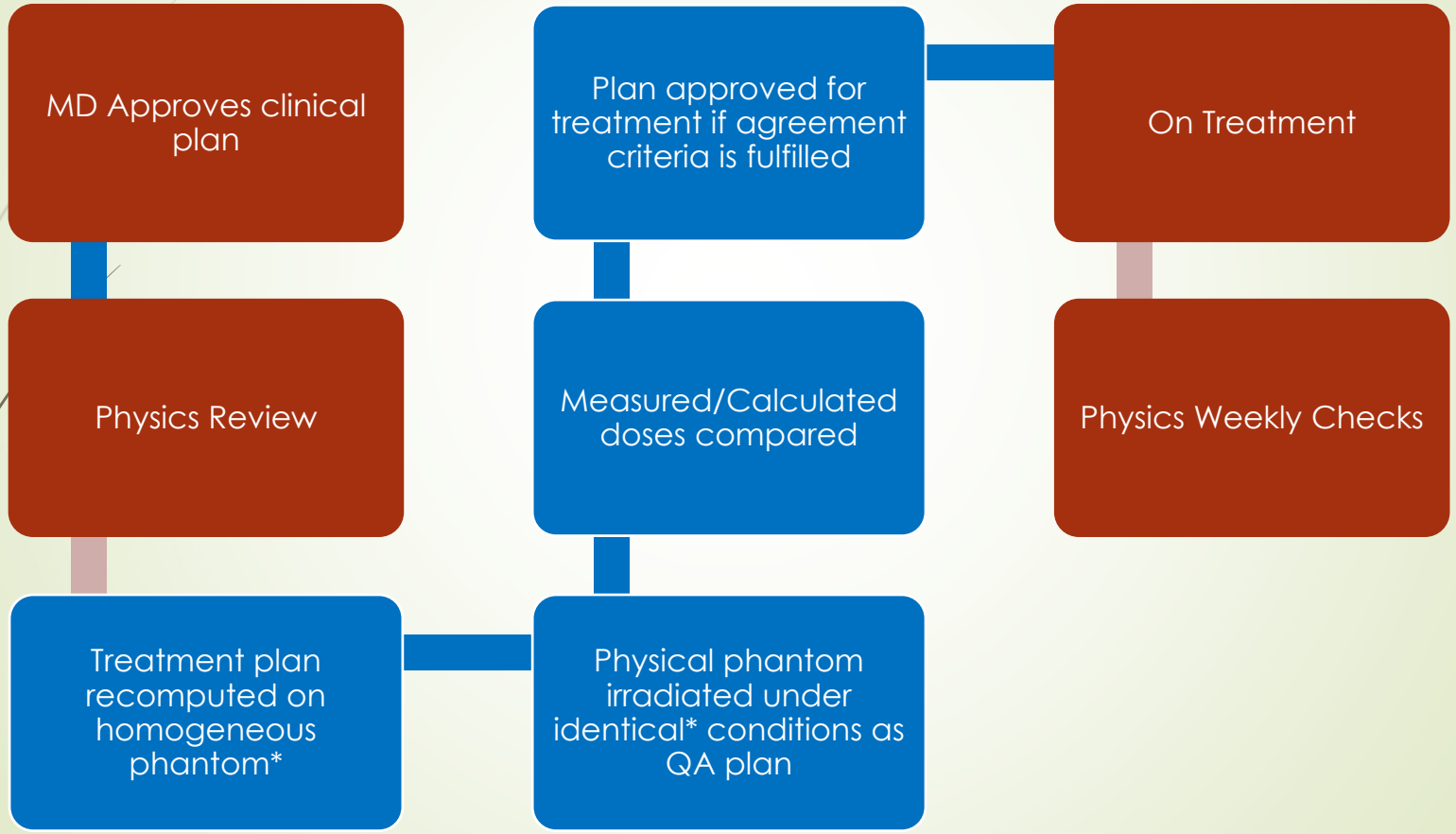
Comparison of 3DCRT and IMRT plan treating supradiaphragmatic Hodgkin Lymphoma

1. De Sanctis et al., Radiation Oncology 7(1):129

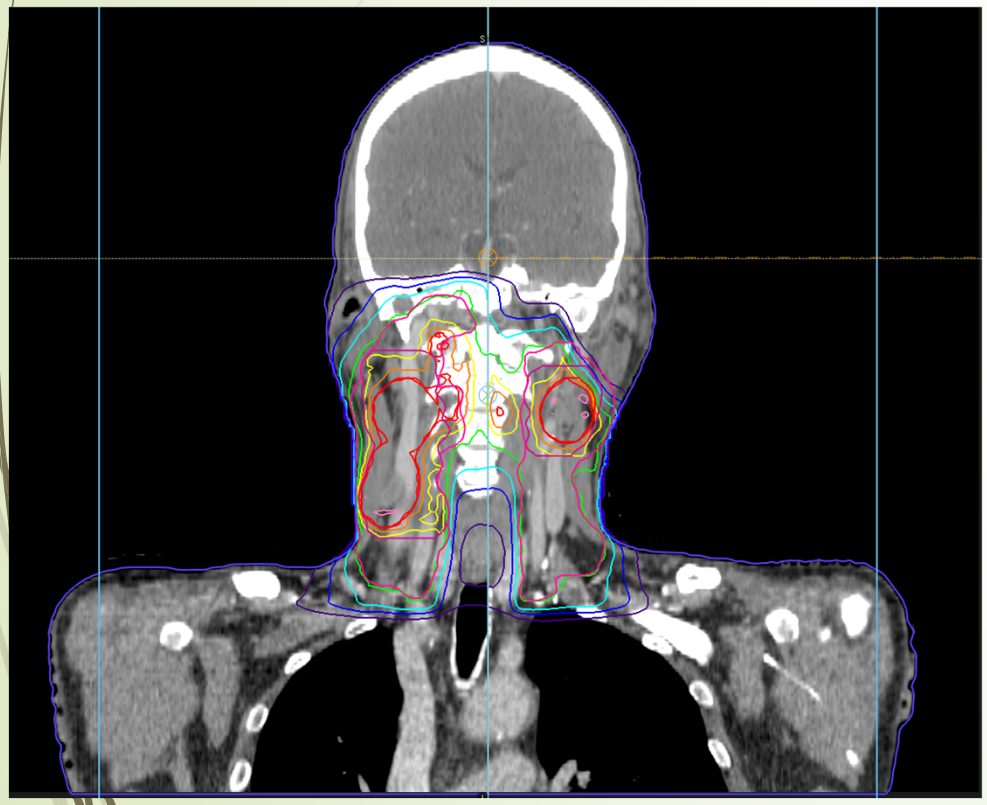
Generalized clinical workflow



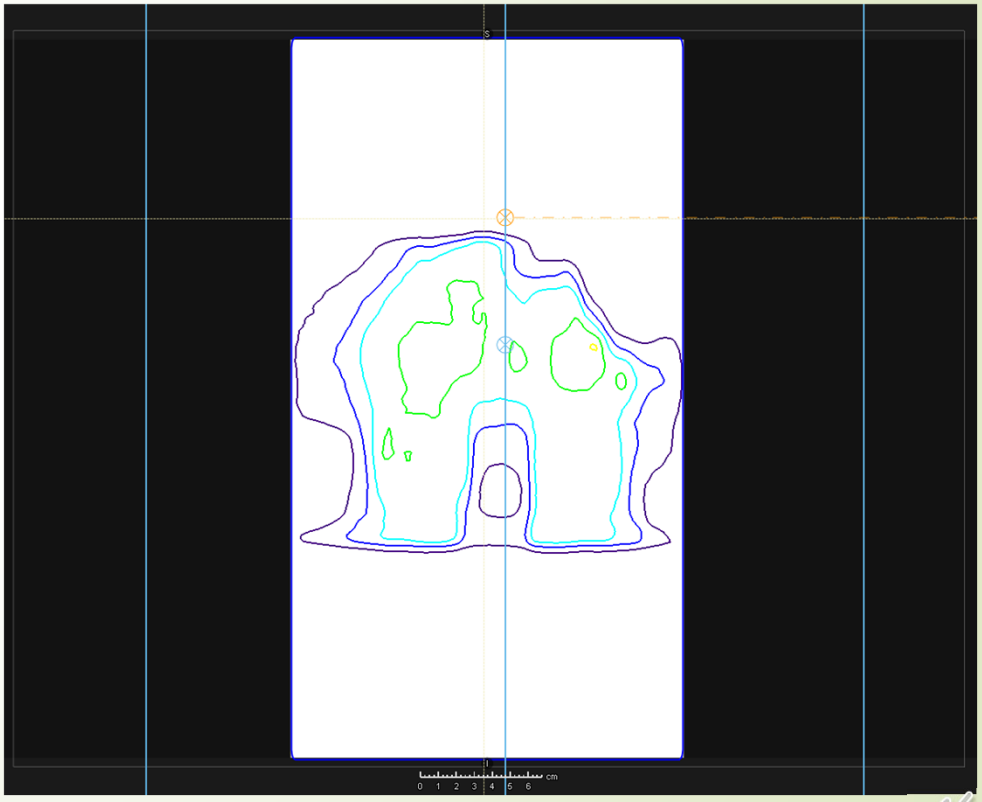
PSQA expanded workflow



Transfer of dose to homogeneous phantom



Bilateral neck clinical plan in coronal view



Bilateral neck QA plan calculated on homogeneous cylindrical phantom in coronal view

General interpretation of "good" PSQA results

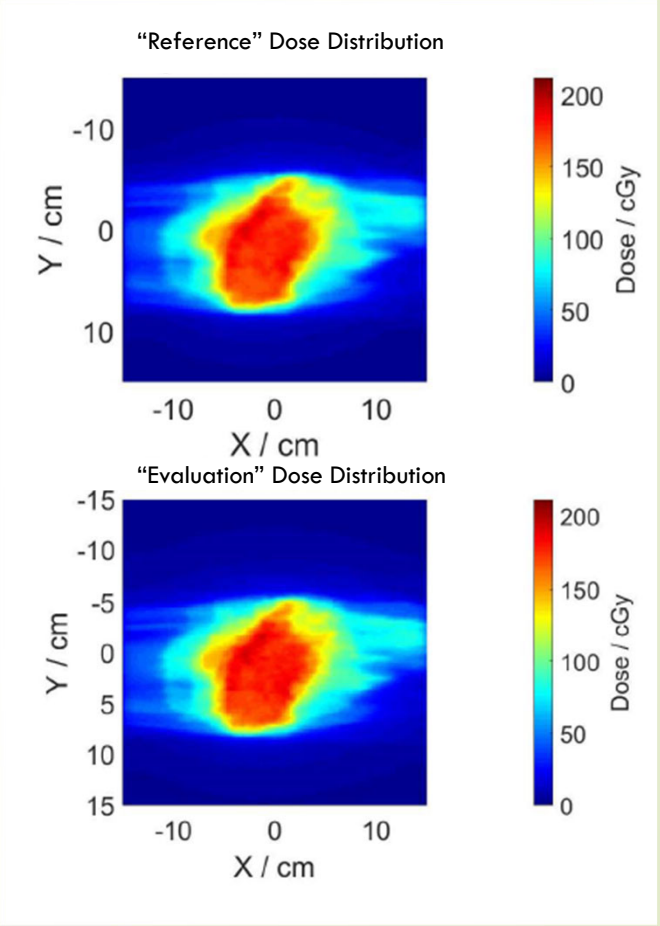
- If a PSQA measurement/calculation agree within the specified criteria…

- We CAN infer that:
  - The patient plan was successfully transferred to the machine (assuming clinical plan was accessed for QA delivery)
  - The patient plan does not push the machine beyond its mechanical limits
  - The delivered patient plan was calculated accurately on the phantom geometry by the TPS

- We CANNOT infer that:
  - The clinical plan accurately accounts for patient heterogeneities
  - The clinical plan is robust against patient positioning errors

Dosimetric comparisons – general terms

- "Reference" dose – physical dose calculated by the TPS on the homogeneous phantom geometry
- "Evaluated" dose – physical dose measured by a radiation detector in the homogeneous phantom geometry
- Some comparison techniques are sensitive to which dose distribution is the "reference" dose and others are not

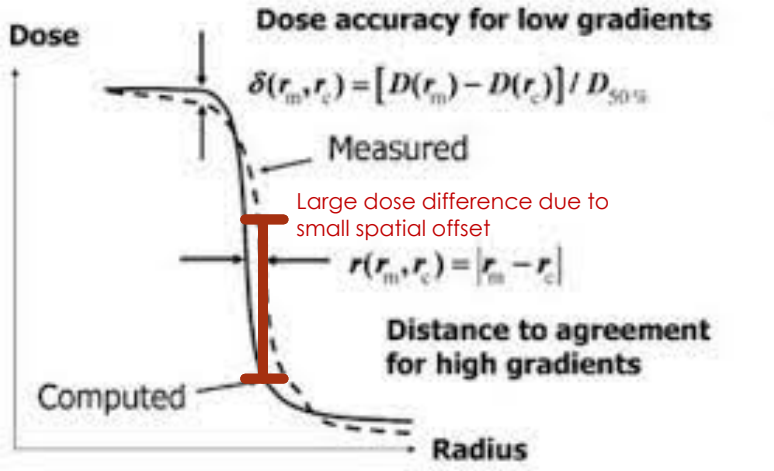


1

1. Desai V.K., Thesis, UW-Madison

Clinically relevant dose comparison considerations

- Dose errors – is something causing the evaluated dose to deviate significantly from the reference dose?
  - Simplest comparison – take a difference between two distributions
  - High gradient regions can have huge dose differences even if only slightly misaligned
- Spatial errors – is the dose being compared sensitive to positional uncertainties?
  - Realistic goal for spatial uncertainty should be based on achievable patient setup uncertainty



Comparison of dose in low and high gradient region

1. Das, I.J. et al. Practical Radiation Oncology 7(2):E145-E155

Dose comparison methods

- **Dose difference test**
  - $\delta(\vec{r}) = D_e(\vec{r}) - D_r(\vec{r})$
  - Dose difference, $\delta(\vec{r})$, at location $(\vec{r})$ is difference between evaluated and reference dose
  - Ideal test when two dose distributions are perfectly spatially superimposed

- **Appropriate** to use independently in low-dose gradient regions
  - In these regions, dose changes slowly with location so spatial uncertainties have minimal impact on dosimetric disagreement

- **Inappropriate** to use independently in high-dose gradient regions
  - In these regions, dose changes rapidly with location so dose differences can often be attributed to small spatial misalignments which are otherwise clinically acceptable

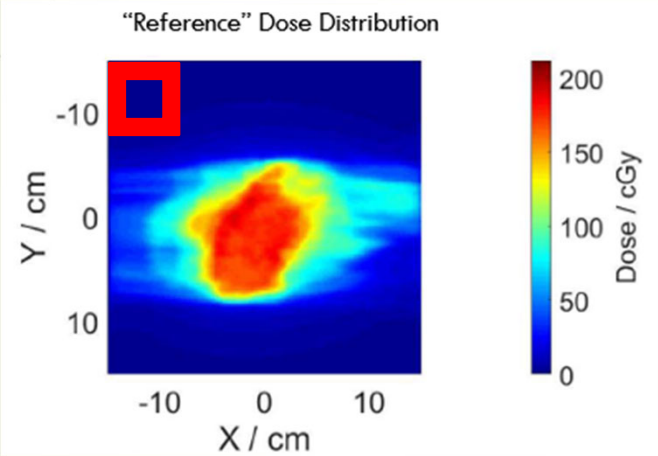
Dose comparison methods

- **Distance to agreement (DTA)**[1]
  - In steep dose gradients, distance between two dose distributions should be considered as part of acceptance criterion
  - DTA for a point in reference dose distribution is the closest location in the evaluated dose distribution with the same dose as the point in the reference distribution
    - Evaluated dose distribution is being searched – therefore DTA is NOT invariant to which distribution is "reference" and which one is "evaluated"
- **Appropriate** for use in high-dose gradient regions in conjunction with a dose difference test
- **Inappropriate** to use in low-dose gradient regions
  - Oversensitive to small differences in dose
- Unlike dose difference, DTA is NOT clinically relevant when evaluated alone in a low gradient region

1. Van Dyk, J., et al. International Journal of Radiation Oncology Biology and Physics 26:261-273

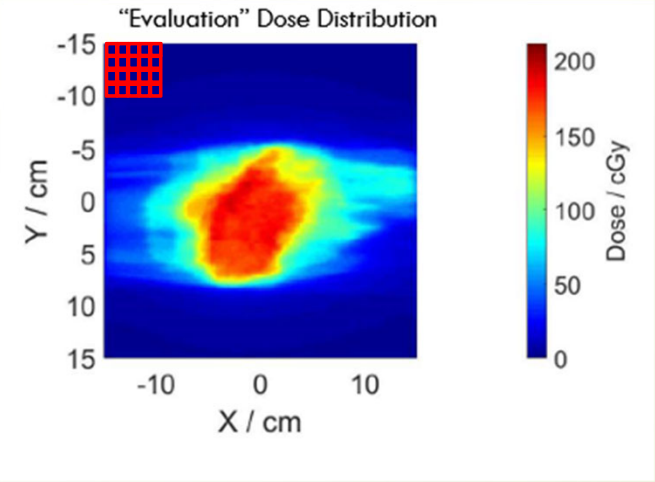
DTA not invariant to "reference dose"

Reference Voxel



Assume 3 mm TPS dose voxels

Evaluation Voxels



[1]

Assume 72 DPI film scan – resolution is ~0.35 mm

- Reference distribution can have ANY resolution (DTA calculates point by point in reference distribution) .
- Evaluated dose distribution should have at least the same or greater resolution and dimensionality than the reference distribution

1. Desai V.K., Thesis, UW-Madison

Dose comparison methods

- **Composite test[1]**
  - Reference point passes if either/both DTA and dose difference test pass
  - Failure only if point fails BOTH individual tests
  - Binary pass/fail, no indication of magnitude of failure

- **Gamma (γ) test[2]**
  - Calculates displacement in multidimensional space considering both dose difference and DTA

$$\Gamma(\vec{r}_e, \bar{r}_r) = \sqrt{\frac{r^2(\vec{r}_e, \bar{r}_r)}{\Delta d^2} + \frac{\delta^2(\vec{r}_e, \bar{r}_r)}{\Delta D^2}}$$

1. Harms, WB Sr et al., Medical Physics 25:1830-1836
2. Low, D. et al., Medical Physics 25:656-661

Dose comparison methods

- **Gamma (γ) test**
  - Calculates displacement in multidimensional space considering both dose difference and DTA

$$\Gamma(\vec{r}_e, \bar{r}_r) = \sqrt{\frac{r^2(\vec{r}_e, \bar{r}_r)}{\Delta d^2} + \frac{\delta^2(\vec{r}_e, \bar{r}_r)}{\Delta D^2}}$$

DTA Criteria Dose difference Criteria

$$\gamma(\vec{r}_e, \bar{r}_r) = \min\{\Gamma(\vec{r}_e, \bar{r}_r)\} \forall \{\vec{r}_e\}$$

Gamma test

- DTA criteria represented in two spatial dimensions in this example (x and y)
- Dose difference is vertical axis
- Combination of dimensions subtends an ellipsoid
- Magnitude of vector must be < 1 to be within ellipsoid surface and pass
- For a single **reference** point, we are considering multiple **evaluation** points

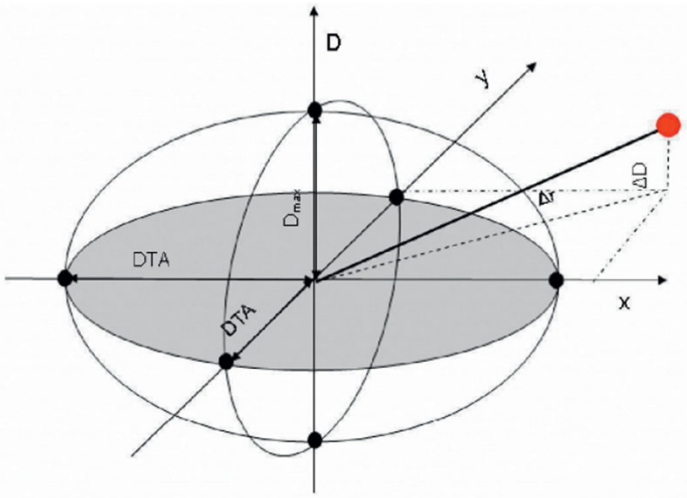


Fig. 3. The concept of gamma verification [5]: x, y, D – spatial and dose dimensions; DTA – distance-to-agreement; D_{max} – max dose deviation; Δr, ΔD – local spatial and dose divergence of the analyzed point

1

Graphical representation of combined dose-difference and DTA tests with two spatial dimensions and one dose dimension

1. Winiecki et al., Rep Pract Oncol Radiother 14(5):162-168

Normalization

- Two main types when considering dose differences: global and local
- **Global normalization**:
  - $\delta_{\text{Global}} = \frac{D_{\text{eval}} - D_{\text{ref}}}{D_{\text{norm}}} \times 100$
- Global normalization implies a D_{norm} value applied to <u>all</u> point pairs
  - Maximum dose in reference data
  - Prescription dose
- Example: Assume 6% dose difference criteria
  - Measured (evaluation) dose at a location: 1.6 Gy
  - Calculated (reference) dose at same location: 1.5 Gy
  - Normalized to 2.0 Gy prescription so D_{norm} = 2 Gy
  - Comparison <u>passes</u> criteria

$$\delta_{\text{Global}} = \frac{1.5\ \text{Gy} - 1.6\ \text{Gy}}{2.0\ \text{Gy}} \times 100 = -5.00\%$$

Normalization

- Two main types when considering dose differences: global and local
- **Local normalization**:
  - $\delta_{\text{Local}} = \frac{D_{\text{eval}} - D_{\text{ref}}}{D_{\text{ref}}} \times 100$
- Local normalization does not make use of any D_{norm} value
  - Direct comparison between two values
  - Extremely unforgiving in low dose regions
- Example: Assume 6% dose difference criteria
  - Measured (evaluation) dose at a location: 1.6 Gy
  - Calculated (reference) dose at same location: 1.5 Gy
  - 2.0 Gy prescription ignored in comparison
  - Comparison fails criteria

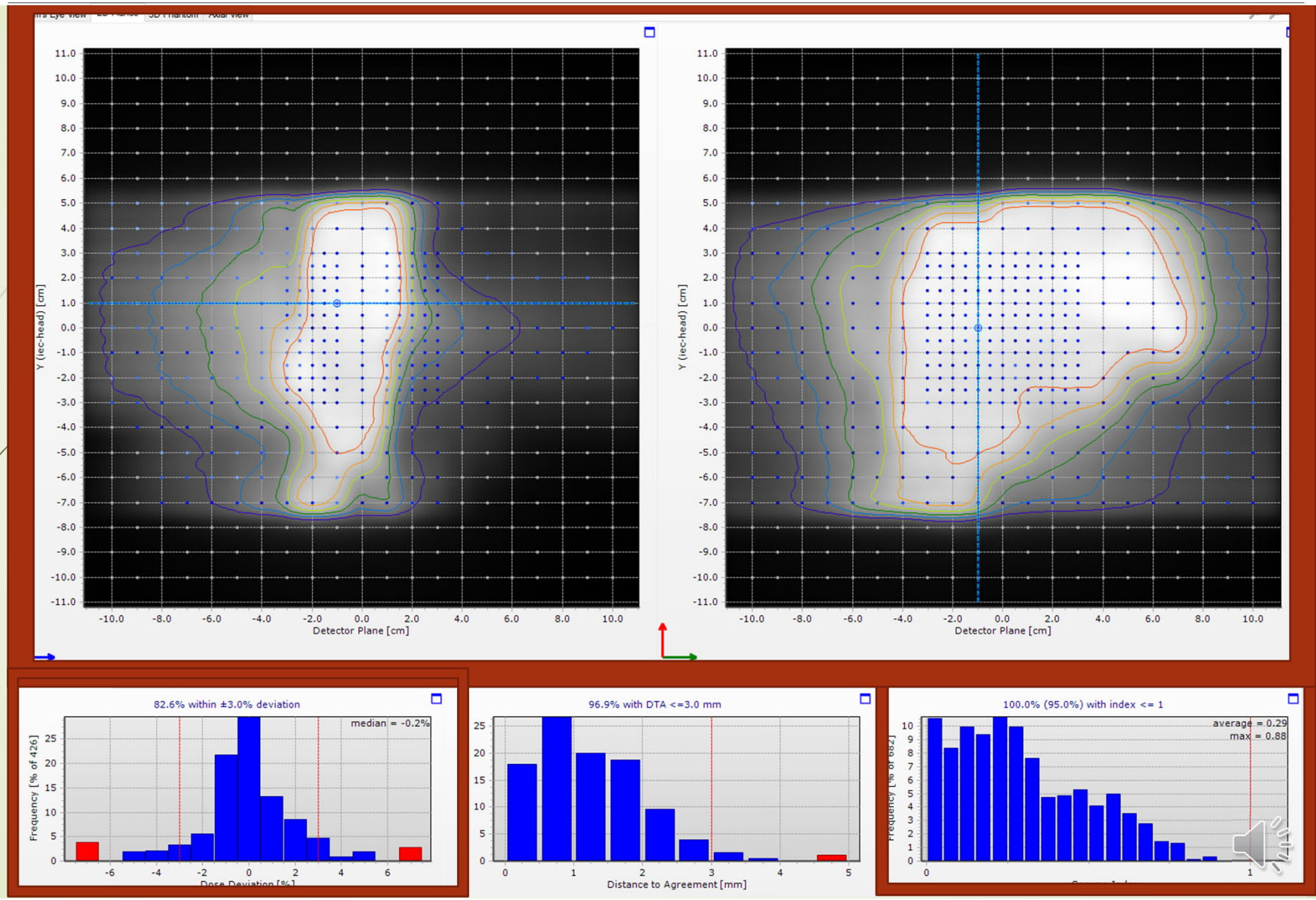
$$\delta_{\text{Local}} = \frac{1.5\ \text{Gy} - 1.6\ \text{Gy}}{1.5\ \text{Gy}} \times 100 = -6.67\%$$

Distilling information from gamma tests

- Ideal situation: Dose difference and DTA criteria specific to various organs at risk and measurements made on a representative heterogeneous phantom
  - Stricter values for spinal cord
  - Looser values for less critical organs
- Reality: Need to more crudely separate out the "clinically meaningful" dose
  - Achieve with dose thresholding
  - If a reference dose value is less than user-specified threshold, exclude from gamma analysis
  - Common dose threshold values are 10% or 20% of D_{max} or D_{RX}

Interpretation of gamma analysis

- Advantage of gamma anlaysis is that it factors in magnitude of failure
  - A failing point with a gamma value of 1.01 is much less concerning than one failing with a gamma value of 10.1
- Total gamma pass rate [%] is $\left(\frac{\text{Number of Points in Plan with Gamma} \leq 1}{\text{Total Number of Points Evaluated in Plan}}\right) * 100$
- **A gamma pass rate alone should not be interpreted as a clinically acceptable plan** (independent of pass/fail criteria)
  - Visual assessment of gamma "heat maps" should be performed
  - Distribution of gamma values should be examined
  - Median of gamma distribution should be assessed


Y (iec-head) [cm]
Detector Plane [cm]
82.6% within ±3.0% deviation
median = -0.2%
Frequency [% of 426]
96.9% with DTA <=3.0 mm
Distance to Agreement [mm]
100.0% (95.0%) with index <= 1
average = 0.29
max = 0.88

Select recommendations of TG-218

- Select takeaways in context of topics discussed in this talk – see TG-218 for full list
  - Global normalization should be used as it's more clinically relevant
  - A dose threshold should be used to exclude low-dose areas with little or no clinical relevance since they can bias the gamma analysis
  - When using gamma analysis and global normalization:
    - **Universal tolerance: Gamma pass rate should be ≥95% with a 3%/2mm criterion and a 10% (of Rx) dose threshold**
    - Universal action limit: Gamma pass rate ≥90% with a 3%/2mm criterion and a 10% dose threshold

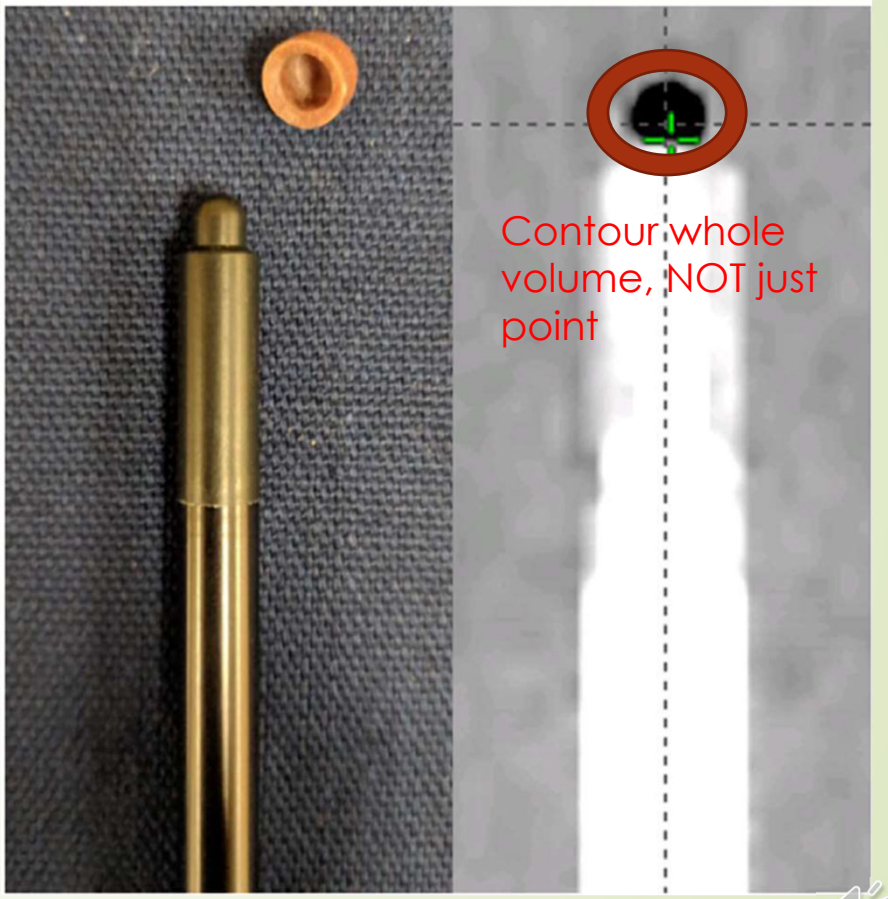
Radiation measurement devices for PSQA

- Ideal Detector Characteristics
  - Measure absolute dose
  - Integration with existing clinical systems and/or clinical experience
  - Easy to set up
  - Reusable
  - High spatial resolution
  - 3D dose measurement
  - Tissue equivalent
  - Stability and total lifetime
  - Energy, dose rate, and directionally independent
  - Affordability

- Reference for detectors being discussed: AAPM TG-120 - Dosimetry tools and techniques for IMRT

Major types of radiation measurement devices for PSQA

- **Ionization chambers**
  - Pros
    - Absolute dosimeters
    - Stable
    - Directionally invariant (cylindrical)
    - Various sizes available
    - Reusable and instantaneous readout
    - Largely energy-independent
  - Cons
    - Volume averaging
    - Polarity/Leakage issues
    - Only capable of "point" measurement without a scanning arm



Exradin A26 micro chamber (left) and CT scan of chamber in phantom (right)

1. Desai V.K., Thesis, UW-Madison

Major types of radiation measurement devices for PSQA

- **Diodes**
  - Pros
    - Negligible volume averaging
    - High sensitivity
    - Reusable
    - In-vivo measurements (no bias)
  - Cons
    - Require calibration for absolute dosimetry
    - Directional dependence
    - Only capable of "point" measurement without a scanning arm
    - Radiation damage occurs over time
    - Noticeable energy dependence
    - High Z materials present (not tissue-equivalent)

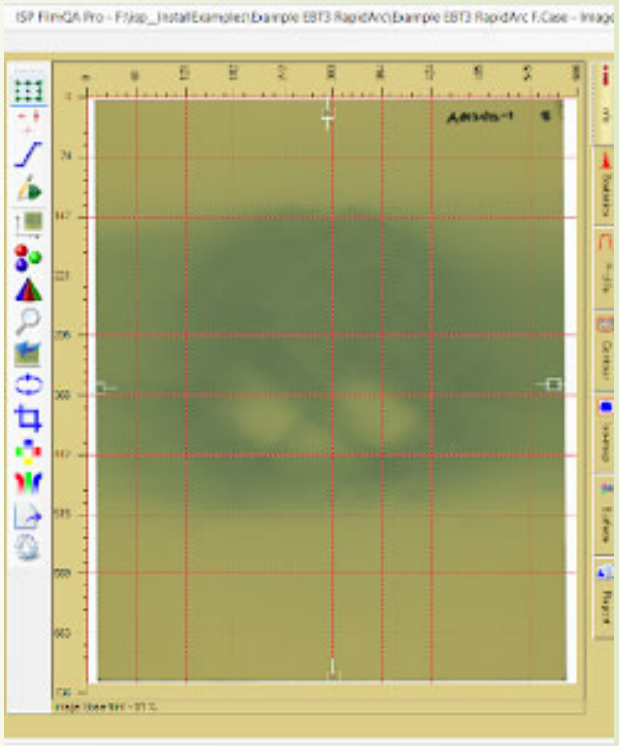
1

EDGE Detector from Sun Nuclear

1. (Image source) Sun Nuclear Corporation, Melbourne, FL, USA, www.sunnuclear.com

Major types of radiation measurement devices for PSQA

- **Radiochromic film**
  - Pros
    - Atomic composition and density close to water
    - 2D detector
    - High spatial resolution
    - No chemical processing required, simple flatbed scanner works fine
  - Cons
    - Each lot/box requires calibration for absolute dosimetry
    - Must be handled carefully, scanned at specific time intervals, and follow a suitably established scan protocol
    - Non-reusable and non-permanent
    - Larger dose uncertainty (~3-5%)
    - Suitable max dose is ~10 Gy



Example clinical dose distribution delivered to GAFChromic EBT3 film analyzed in FilmQA Pro software

1. Source: Ashland, Bridgewater, NJ, USA, www.ashland.com

Major types of radiation measurement devices for PSQA

- **Array Detectors**
  - Pros
    - 2D or 3D dose measurement
    - Reusable
    - Efficient to use
    - Instantaneous readout
  - Cons
    - Require calibrations for absolute dose
    - Low spatial resolution

MapCHECK® 3

1

ArcCHECK® MR

2

Delta^4 Phantom+ MR

3

1. (Image source) Sun Nuclear Corporation, Melbourne, FL, USA, www.sunnuclear.com
2. (Image source) Sun Nuclear Corporation, Melbourne, FL, USA, www.sunnuclear.com
3. (Image source) Delta^4 by ScandiDos, Uppsala, Sweden, www.delta4family.com

Major types of radiation measurement devices for PSQA

- **EPID**
  - Pros
    - Built into TDS (no additional equipment required)
    - Efficient to use once commissioned for pretreatment QA
    - Can be used to detect pretreatment or on-treatment errors
  - Cons
    - Requires very careful calibration (not meant to be used as an absolute dosimeter)
    - Oversensitive to low energies (scatter, off-axis, field size, and phantom attenuation implications)

Varian TrueBeam with EPID extended [1]

1. (Image source) Radiation Oncology Systems, San Diego, CA, USA, www.oncologysystems.com

Non-measurement-based approaches

- **Independent secondary monitor unit calculation**
- Purpose: Redundant (but independent) verification that the total MU or time calculated by the primary TPS is reasonable
- Secondary calculation does NOT need to exactly reproduce the dose distribution calculated by the TPS
  - It just needs to indicate whether the values are generally reasonable or not
  - Compromises in the secondary check calculation algorithm could be made to increase efficiency (but the independent check can be just as complex as or even more complex than the commissioned TPS)
  - Should be totally independent of delivery system
- See AAPM TG-114 for more information about monitor unit verification

Non-measurement-based approaches

- **Log file analysis**
  - Purpose: Verification that treatment delivery was executed as intended
  - Efficient means of setup and delivery verification on a per-fraction basis
  - Trusting machine's ability to self-report (not independent)
  - Can be difficult to translate mechanical discrepancies into dosimetric ones

Conclusions and Future Direction of PSQA

- PSQA helps to ensure appropriate implementation of IMRT in clinical setting
  - QA tools and methodologies must be understood carefully in order to verify that PSQA is being properly verified
- PSQA explicitly tests for:
  - Integrity of data transfer from TPS to TDS
  - Physical deliverability of plan by the treatment machine
  - Agreement of measured dose and dose calculated by TPS using a specific phantom geometry
- PSQA is NOT explicitly testing for:
  - Beam model accuracy including ability of TPS to calculate dose in presence of heterogeneities
  - Robustness of plan against patient motion or positioning error

Conclusions and Future Direction of PSQA

- Ongoing debate: Do we need measurements at all?
  - Deliverability issues attempting to be predicted using complexity metrics and machine learning
- Idealized future solution: independent verification of patient dose on a per fraction basis
- Vendor partnerships important for improving PSQA workflow