Jennifer O'Daniel, Ph.D.
Duke University Medical Center
Spring Clinical AAPM Meeting, 2016

Digging into our Toolbox: Non-Gamma Analysis Tools for IMRT

DukeMedicine

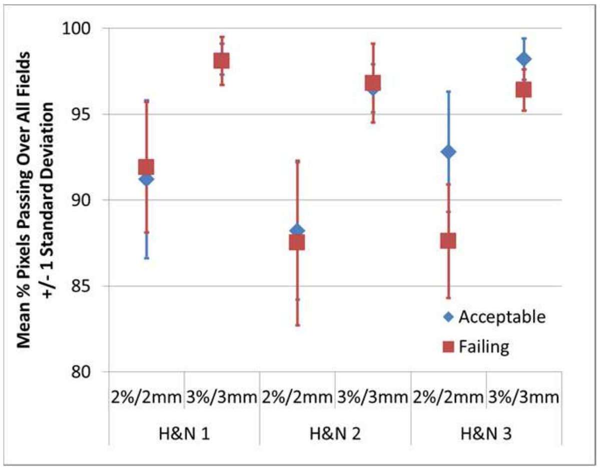
Acknowledgements/Disclaimer

- Appreciation for discussions with Daniel Low, Kyle Antes, Stephen Kry, and the faculty of Duke University.
- The opinions discussed are my own, and do not necessarily represent the opinions of Duke University.
- The discussion of commercial products is not intended as an endorsement/opposition to any specific product.

Questions to address

- Why look for alternatives to gamma analysis?
- What are the available alternatives to gamma analysis?
- How accurate are the alternatives?
- How do the results compare to gamma analysis?

Gamma analysis cannot separate acceptable vs. failing plans



High modulation

↓

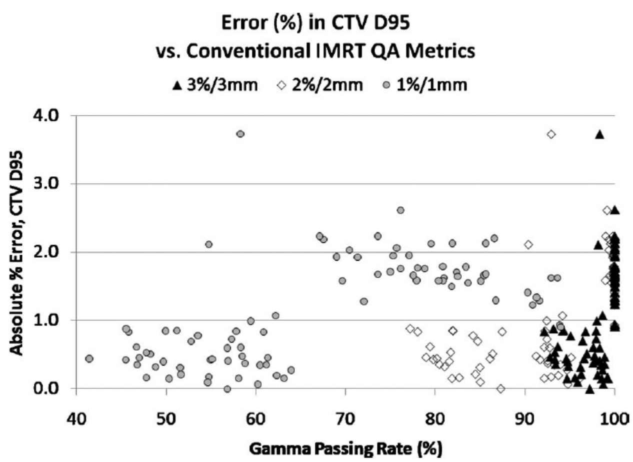
High dose gradients

↓

Errors obscured by DTA criteria

Kruse Med Phys 2010

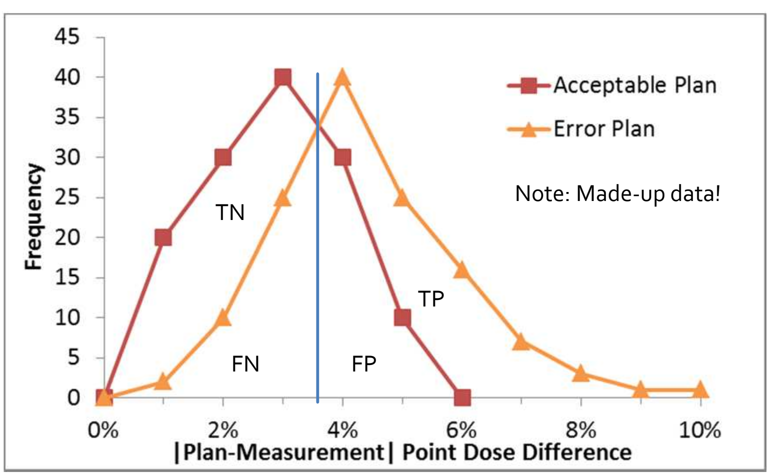
Gamma analysis cannot separate acceptable vs. failing plans



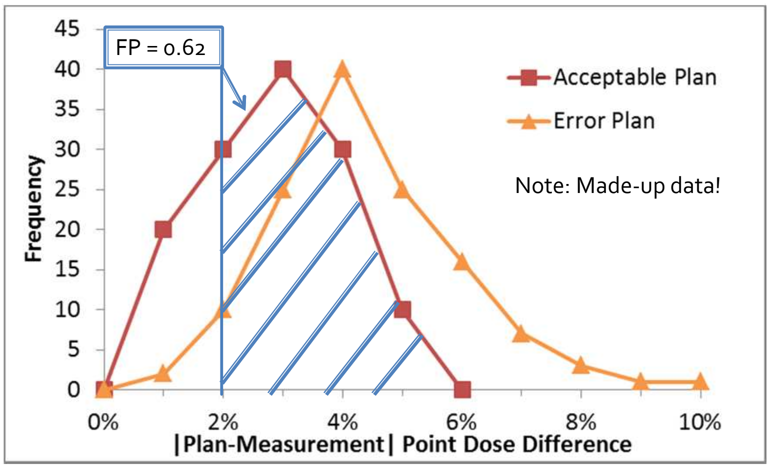
- 24 H&N plans x 4 beam model errors
- No correlation with gamma: cord, C parotid, larynx
- Greater errors correlate to greater gamma passing rates: I parotid, CTV

Nelms Med Phys 2011

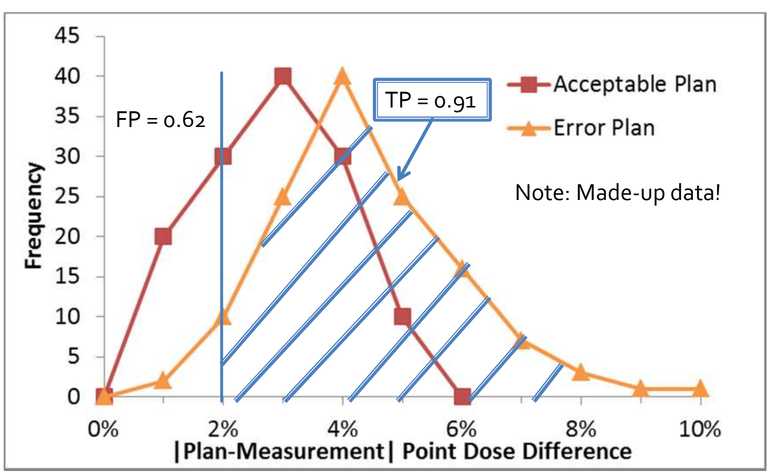
How to Evaluate a QA Test



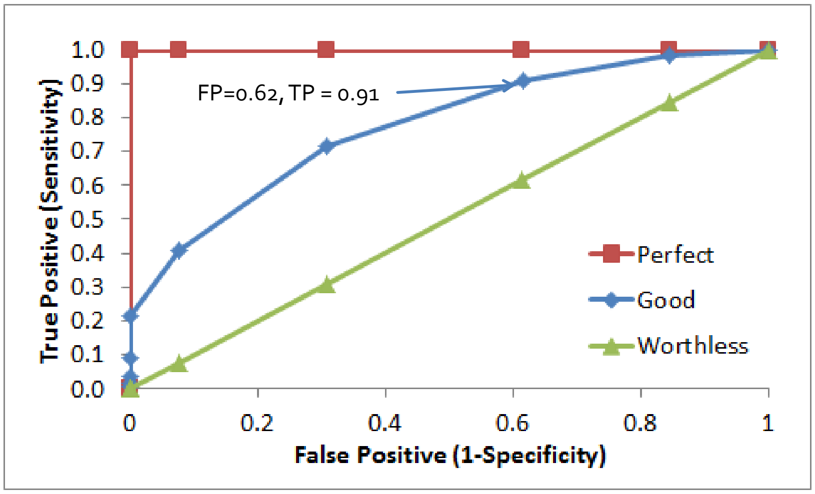
How to Evaluate a QA Test



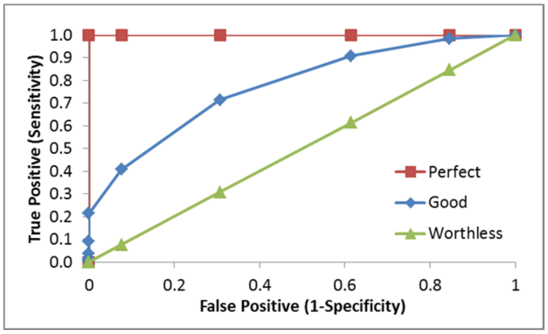
How to Evaluate a QA Test



Receiver Operating Characteristic Curve (ROC)



Area Under the Curve (AUC)



- Maximum = 1
- Minimum = 0.5

- Excellent ≥ 0.9
- Good ≥ 0.8
- Fair ≥ 0.7

SAMS 1: Which one of the following statements is false?

20%	1.	ROC plots the true positive rate vs. the false positive rate.
20%	2.	A useful test is one that has a high true positive rate and a low false positive rate.
20%	3.	The cutoff threshold selected determines the sensitivity and specificity of the test.
20%	4.	AUC > 0.8 indicates good test accuracy.
20%	5.	AUC < 0.80 indicates good test accuracy.

10

SAMS 1: Which one of the following statements is false?

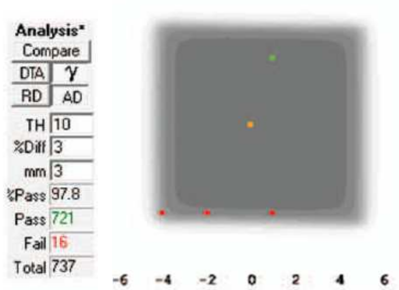
1. ROC plots the true positive rate vs. the false positive rate.
2. A useful test is one that has a high true positive rate and a low false positive rate.
3. The cutoff threshold selected determines the sensitivity and specificity of the test.
4. AUC > 0.8 indicates good test accuracy.
5. **AUC < 0.80 indicates good test accuracy.**

T Fawcett, "An introduction to ROC analysis," Pattern Recogn. **27**, 861-874 (2006).

What are the available alternatives to 2D gamma analysis?

- Single point measurement
- 1D line profile
- 2D dose distribution with non-gamma analysis
- 3D dose distribution

Single Point Measurement



- 10x10 field
- ArcCheck 3% G/3mm
  - 97.8% passing
- Ion chamber @ iso
  - -3.8%
- Solution: correct phantom density in TPS

Nelms, Med Phys **40**, 111722 (2013)

Ion Chamber – The Best?

- 26 IMRT plans, 15 acceptable & 9 unacceptable per multi-ion chamber phantom

Equipment/Technique	AUC
Ion chamber	0.94
Helical diode array	0.81
2D diode array (AP composite)	0.80
2D diode array & Film (Planned angle composite)	0.65 & 0.76
2D diode array (AP field-by-field)	0.61

McKenzie 2014 Med Phys **41**

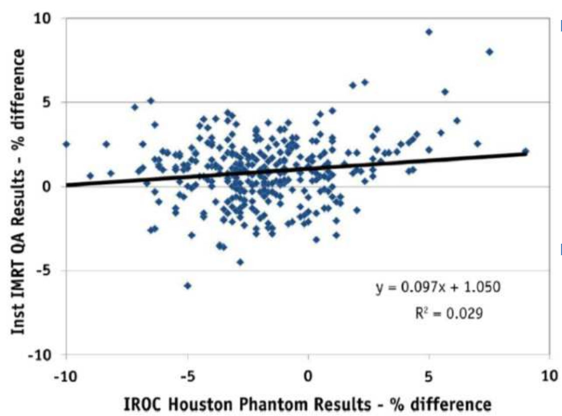
Ion Chamber – Worthless?

- 855 irradiations of IROC head-and-neck phantom
- Examined sensitivity and specificity of institution's IMRT QA to predict phantom measurement passing/failing

		IROC Houston	
		Fail	Pass
Institution QA	Fail	2	3
	Pass	120	730

Kry 2014 IJROBP

Ion Chamber – Worthless?



- Nearly no relationship between IROC TLD and institution ion chamber
- AUC values:
  - Film = 0.70
  - Ion chamber = 0.66
  - 2D diode array = 0.61

Kry 2014 IJROBP

What changed?

- Standardization of personnel & equipment
- Method of determining plan acceptability
  - Multiple IC measurement (10-15 x 0.057cc), deviation > 3% local
  - vs. IROC phantom
    - 70% failures ← TLD deviation > 7%
    - 10% failures ← Film < 85% passing 7%/4mm gamma
    - 20% failures ← Both TLD & film fail
- Failure is acceptable vs. failure is unacceptable

What do we do with a failing ion chamber measurement?

301 of 13,008 IMRT QAs failed the first ion chamber measurement

Remeasure 1	188 passed	66 failed	**47 not remeasured**
Remeasure 2	30 passed	11 failed	**25 not remeasured**
Remeasure 3	5 passed	2 failed	**4 not remeasured**
Remeasure 4			**2 not remeasured**

Pulliam 2014 JACMP

What do we do with a failing ion chamber measurement?

- 74% eventually passed re-measurement
- 26% failed consistently
  - 3 cases (1% of failures) replanned
  - "Small proportion" scaled MU
  - Remainder treated "as-is" after consultation with MD
- Are we missing QA failures due to clinical time pressure?

Pulliam 2014 JACMP

SAMS 2: Fill in the blanks: _________ measurements had the greatest accuracy at sorting acceptable from failing plans, while __________ measurements had the lowest accuracy.

20%	1.	Ion chamber, 2D AP field-by-field
20%	2.	Ion chamber, helical diode array
20%	3.	Helical diode array, 2D AP field-by-field
20%	4.	Helical diode array, ion chamber
20%	5.	2D AP field-by-field, ion chamber

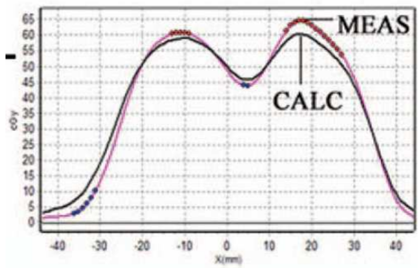
10

SAMS 2: Fill in the blanks: _________ measurements had the greatest accuracy at sorting acceptable from failing plans, while __________ measurements had the lowest accuracy.

1. **Ion chamber, 2D AP field-by-field**
2. Ion chamber, helical diode array
3. Helical diode array, 2D AP field-by-field
4. Helical diode array, ion chamber
5. 2D AP field-by-field, ion chamber

E. McKenzie *et al.*, "Toward Optimizing Patient Specific IMRT QA techniques in the accurate detection of dosimetrically acceptable and unacceptable patient plans," Med Phys **41**, 121702 (15pp.) (2014).

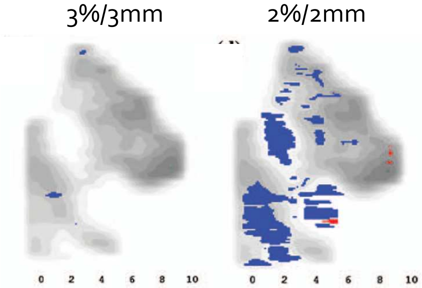
1D Line Profile



Nelms, Med Phys **40**, 111722 (2013)

- MapCHECK 3%G/3mm
  - 99.3% passing
- Line profile
  - Calculation underestimates peaks and overestimates valleys
- Solution: re-measure profiles with diode instead of Farmer chamber

2D Dose Distribution



Nelms, Med Phys **40**, 111722 (2013)
Chan Technol Cancer Res Treat **12** (2013)

- MapCHECK 3%G/3mm
  - 99.4% passing
- 2%L/2mm
  - cold stripes in field
- Solution: correct TPS tongue-and-groove setting

Miminum gamma parameters to detect MLC errors in VMAT

Per Leaf Shift	Open/Closed MLC Banks	Shifted MLC Banks
0.5 mm	2% / 1 mm	Not Evaluated
1.0 mm	2% / 2 mm	2% / 1 mm
2.0 mm	2% / 2 mm	2% / 1 mm
3.0mm	Not Evaluated	2% / 2 mm

Heilemann Med Phys (2013)
Kim Rad Oncol (2014)

SAMS 3: With regards to Nelms 2013, which of the following QA techniques overlooked errors during IMRT commissioning?

1. Measurement-guided dose reconstruction (MGDR)
2. 1D dose distribution profile
3. Ion chamber
4. 2D gamma analysis using 3% global/3mm criteria
5. 2D gamma analysis using 2% local/2mm criteria

20%	
20%	
20%	
20%	
20%	

10

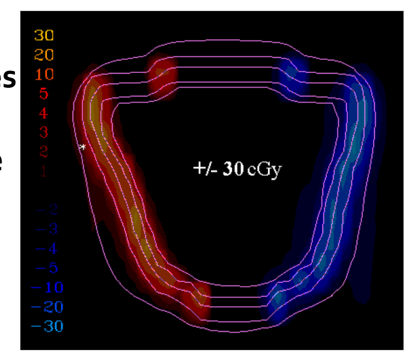
SAMS 3: With regards to Nelms 2013, which of the following QA techniques overlooked errors during IMRT commissioning?

1. Measurement-guided dose reconstruction (MGDR)
2. 1D dose distribution profile
3. Ion chamber
4. **2D gamma analysis using 3% global/3mm criteria**
5. 2D gamma analysis using 2% local/2mm criteria

B Nelms *et al.*, "Evaluating IMRT and VMAT dose accuracy: Practical examples of failures to detect systematic errors when applying a commonly used metric and action levels," Med Phys **40**, 111722 (15 pp.) (2013).

Other 2D Analysis Methods: Dose Gradient Compensation

Why?
Explicitly removes geometric uncertainties due to the QA measurement process.



Simulated 2mm offset

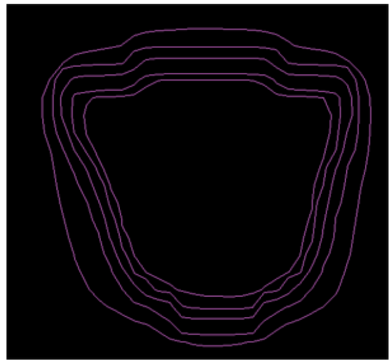
↓

Dose difference map without gradient compensation

- Geometric tolerance (GT) specified by user
- Δdose corrected$_{ij}$ = Δdose$_{ij}$ – (gradient$_{ij}$ x GT)

Moran, JACMP **6**, 62-73 (2005)

Other 2D Analysis Methods: Dose Gradient Compensation



GT = 2mm

No difference between calculation & simulated 2mm error

- Geometric tolerance (GT) specified by user
- Δdose corrected$_{ij}$ = Δdose$_{ij}$ – (gradient$_{ij}$ x GT)

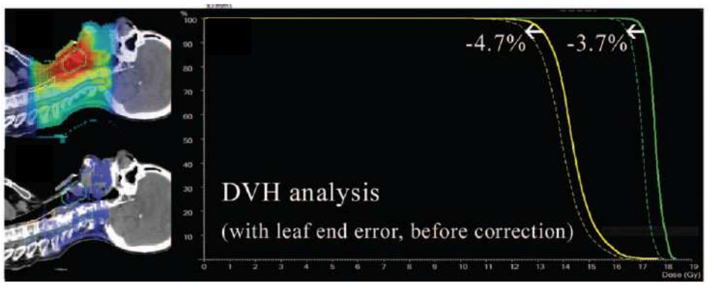
Moran, JACMP **6**, 62-73 (2005)

Other 2D Analysis Methods: Normalized Agreement Test (NAT)

- Based on dose difference and DTA
- Differences from gamma analysis
  - NAT = 0 when either dose difference or DTA criteria is met
  - NAT = 0 for areas of underdoseage outside the PTV
    - Calculated dose < 75% maximum dose
    - Measured dose < calculated dose
- Why? Focus on areas of biological significance

Childress & Rosen, IJROBP **56**, 1464-79 (2003)

3D Dose Distribution



- MapCHECK 3%G/3mm: 99.2% passing
- 3D Dose Reconstruction: 4% dose reduction in targets
- Solution: adjust leaf end offset table

Nelms, Med Phys **40**, 111722 (2013)
Jarry, Med Phys **38**, 3581 (2011)

3D Dose Distribution

- Direct 3D measurement
  - Polymer gels, radiochromic polyurethane
  - Review article: Baldock Phys Med Biol **55** (2010)
  - IC3DDose annual conference (www.ic3ddose.org)
- Pseudo-3D measurement
  - Multiple 2D planes
  - Ion chamber or diode detectors
  - Interpolation/calculation to 3D dose distributions
  - In phantom or in patient

Commercially Available Pseudo-3D Systems

- Sun Nuclear: 3DVH with MapCheck and ArcCheck
- Scandidos: Delta4 DVH with Delta4 pre-treatment (PT) phantom
- PTW: VeriSoft with Octavius4D
- IBA: COMPASS with MatriXX

TPS Dose Perturbation

Beam & MLC Model
CT Dataset
Treatment Plan
Calculated Patient 3D Dose
Calculated QA Dose
Measured QA Dose

Estimated 3D Dose

TPS Dose Perturbation 3DVH (IMRT)

Calculated Patient 3D Dose

↓

( )

↓

Ray-trace each beam through the patient dose grid

↓

Calculate per beam error maps (meas – calc 2D dose difference)

↓

Perturb each patient dose voxel (function of depth/energy)

↓

Estimated Patient 3D Dose

SunNuclear, 3DVH white paper

TPS Dose Perturbation 3DVH (VMAT)

Measured cylindrical phantom dose

↓

Determine sub-arcs based on time/gantry angle

↓

Sum measured dose grid on phantom from all sub-arcs*

↓

Calculate 3D dose correction grid on phantom (meas vs calc)

↓

Apply 3D dose correction grid to calculated patient 3D dose

↓

Estimated Patient 3D Dose

*Hidden complexity
Nelms Med Phys 2012

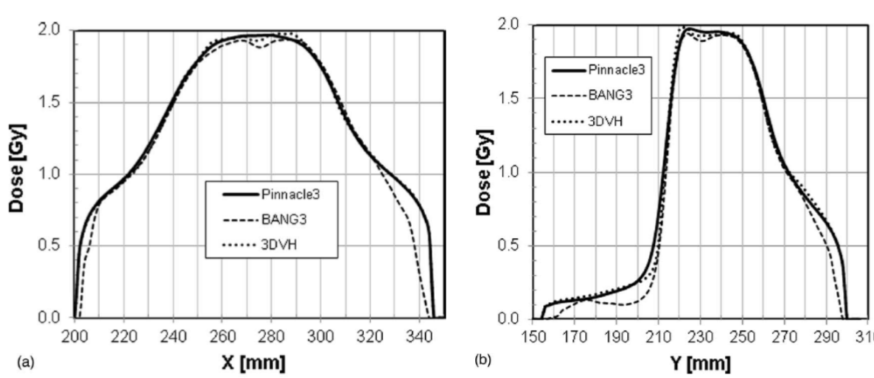
TPS Dose Perturbation Validation 3DVH

Equipment	Reference	Measured Difference
Ion chamber, water equivalent	Olch 2012	1.3% +/- 1.5%
	Nelms 2012*	0.1% +/- 1.0%
	Opp 2013*	<2%
OSLD, water equivalent	Opp 2013*	<2%
OSLD, lung equivalent	Opp 2013*	<4% low modulation <14% high modulation
Film (γ 1%G/2mm)	Nelms 2012*	88.6%
Film (γ 2%G/2mm)	Nelms 2012*	96.1%
	Olch 2012	97.7%
Film (γ 3%G/3mm)	Nelms 2012*	99.5%

IMRT: Olch Med Phys 2012

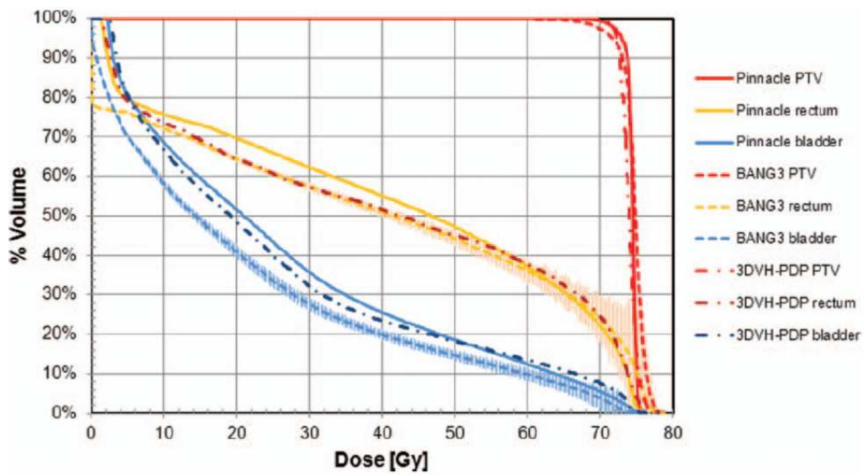
*VMAT: Nelms Med Phys 2012
Opp JACMP 2013

TPS Dose Perturbation Validation 3DVH



VMAT: Watanabe Med Phys 2013

TPS Dose Perturbation Validation 3DVH



VMAT: Watanabe Med Phys 2013

TPS Dose Perturbation Delta4 PT

Measurement in phantom 2 x 2D Diode Detectors

↓

Scale TPS dose in phantom based on measurement

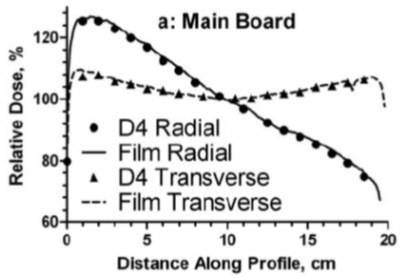
↓

Phantom 3D Dose

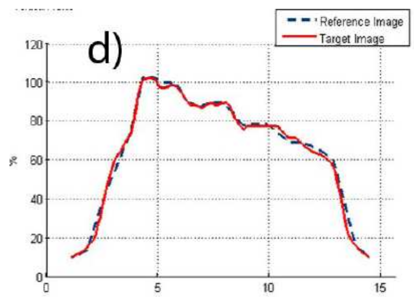
Feygelman JACMP **10** (2009)

TPS Dose Perturbation Validation Delta4 PT

- Good agreement between film and Delta4 in phantom



Static Field



IMRT Plan

Feygelman JACMP **10** (2009)

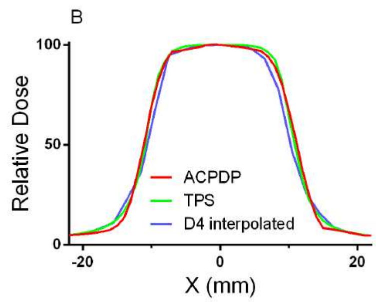
TPS Dose Perturbation Validation Delta4 PT

- Good agreement between film and Delta4 in phantom
- Good agreement with point doses
  - Independent diode: 0.7%
  - TPS point doses: 0.4% ± 0.5%

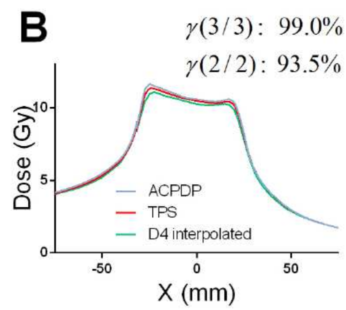
Feygelman JACMP **10** (2009)

TPS Dose Perturbation Validation Delta4 PT

- Good agreement between 3DVH (ACPDP) and Delta4 PT profiles



Static Field

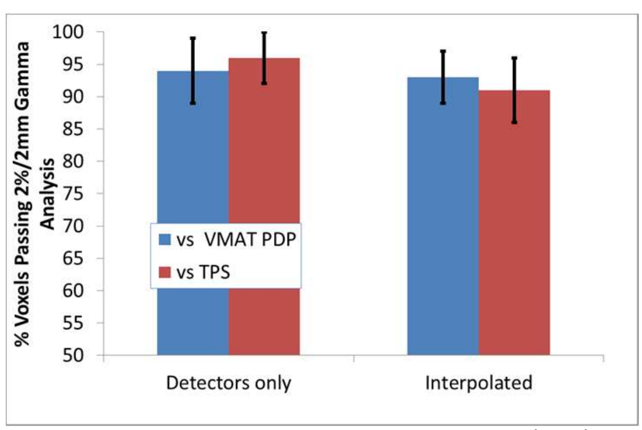


SBRT Lung VMAT

Stambaugh JACMP **15** (2014)

TPS Dose Perturbation Validation Delta4 PT

- 2% / 2mm gamma passing rates > 90%



Stambaugh JACMP **15** (2014)

Forward Calculation Algorithms

Beam & MLC Model

Patient CT Dataset

Treatment Plan

Measured Phantom Dose

Estimated 3D Dose in Phantom or Patient

Forward Calculation Algorithms Compass

Measurement in phantom (2D detector attached to gantry)

Calculate dose in **patient**
(collapsed cone convolution/superposition algorithm)

Boggula Phys Med Biol 55 (2010)

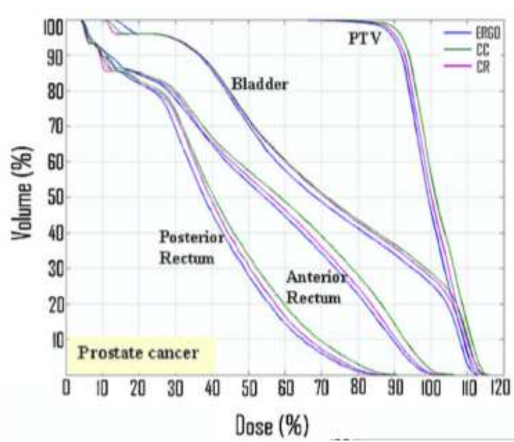
Compass, Collapsed Cone Validation

- 12 VMAT plans in anthropomorphic phantoms
  - Ion chamber vs. PBC TPS: 0.1 ± 1.3%
  - Ion chamber vs. Compass Meas: 1.2 ± 1.1%
  - Film vs. PBC TPS (3%/3mm): 90.9 ± 12.5%
  - Film vs. Compass (3%/3mm): 96.1 ± 5.9%

Boggula Phys Med Biol 55 (2010)

Compass, Collapsed Cone Validation

- Good agreement with DVH in heterogeneous pelvic phantom
- ERGO = PBC TPS
- CC = Compass Calculation
- CR = Compass Dose Reconstruction



Boggula Phys Med Biol 55 (2010)

Forward Calculation Algorithms Delta4 PT + DVH

Measurement in phantom 2 x 2D Diode Detectors

↓

Scale TPS dose in phantom based on measurement

↓

Phantom 3D Dose

↓

Estimate 2D energy fluence per control point

↓

Calculate 3D dose on **patient** CT dataset (pencil beam dose algorithm)

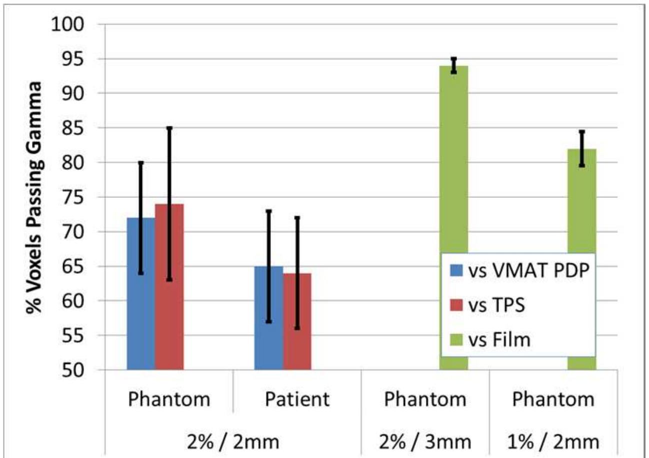
Scandidos, Delta4DVH Anatomy
Gustafsson 2013 white paper, http://scandidos.com/home/documents/references

Delta4 Pencil Beam Validation

- Good agreement with pinpoint ion chamber in static fields
  - 2x2cm2: 1.5 ± 0.5%
  - 10x10cm2: 0.9 ± 0.9%

Hauri JACMP 2014

Delta4 Pencil Beam Validation

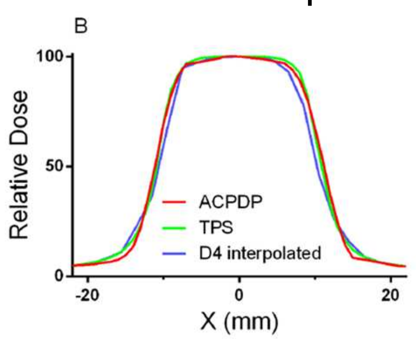


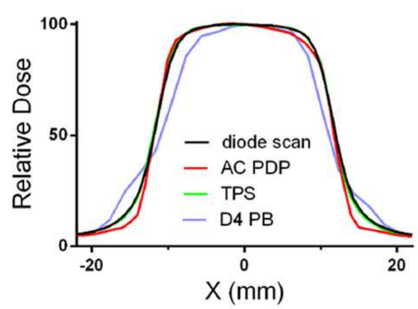
Stambaugh JACMP 2014

Hauri JACMP 2014

Delta4 Pencil Beam Validation

Lateral profiles of 2cm x 2cm field





Interpolated Measurement:
Good agreement

Pencil Beam Calculation:
Clear Penumbra Difference

Stambaugh JACMP 2014

Forward Calculation Algorithms Verisoft/Octavius 4D

Measurement in phantom (rotating 2D detector on table)

↓

Calculate **phantom 3D dose**
(based on PDDs entered during commissioning)

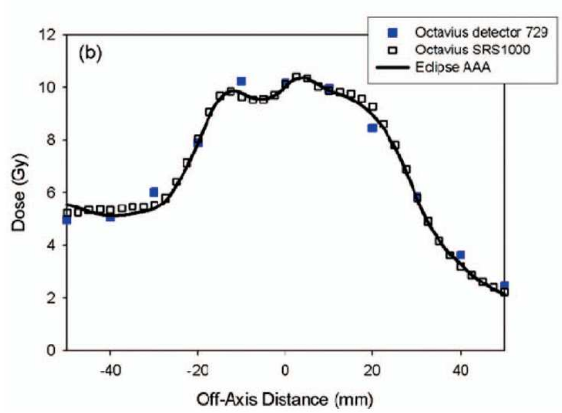
↓

Calculate dose on **patient** CT dataset **within area of measurement**
(TPR, determination of water-equivalent depth)

PTW, Octavius 4D
Allagaier 2013 white paper, http://www.ptw.de/2403.html

Verisoft/Octavius 4D Validation

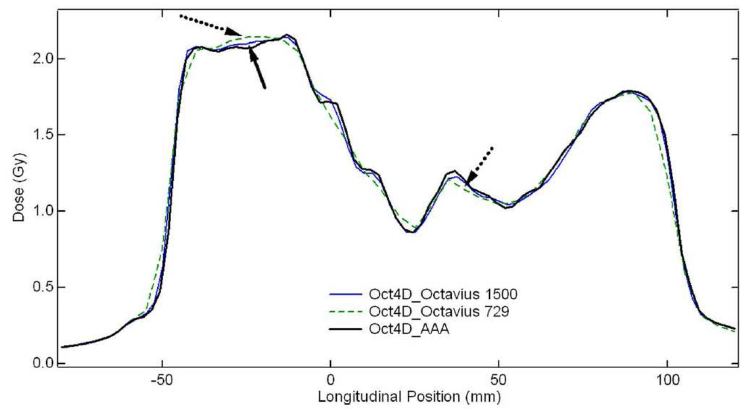
- The Octavius 729 ion chamber spacing of 1cm causes failures in steep dose gradient regions. The Octavius SRS1000 spacing of 2.5mm solves those issues.
- Good agreement with ion chamber at isocenter (0.8-1.3%)



McGarry Med Phys **40** (2013)

Verisoft/Octavius 4D Validation

- The Octavius 1500 spacing of 0.7cm also improves results



Van Esch Med Phys **41** 2014

Verisoft/Octavius 4D Validation

- DVHs only available for structures that are completely inside the measurement volume
- Maximum dose TPS vs. Octavius 729
  - Homogeneous plans ~ ±2%
  - Lung plans ~ ±6%

Allgaier white paper 2013
http://www.ptw.de/2403.html?&cId=390

3D Dose Measurements

- Several options for robust 3D dose distribution measurements in phantom
  - 3D gamma analysis
- Ability to convert to <u>patient</u> 3D dose distributions is improving
  - Required for true DVH-based QA

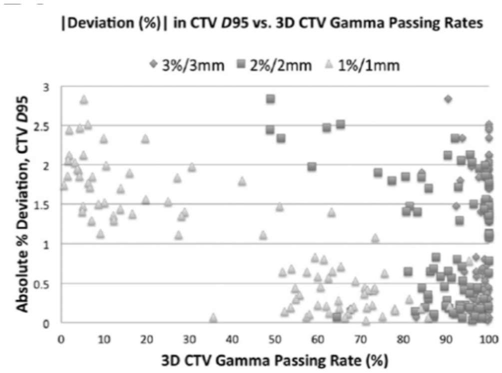
3D Gamma vs DVH

- 96 head-and-neck IMRT plans with introduced errors
- 3D patient CT scans
  - Planned error-free vs. simulated error
  - Evaluate both full 3D-γ & ROI-γ
- Pearson product-moment correlation coefficient
  - Perfect negative correlation r = -1

Zhen Med Phys 2011

3D Gamma vs DVH

- ROI-specific gamma better correlated to DVH than whole patient gamma
- Tighter gamma metrics increase correlations



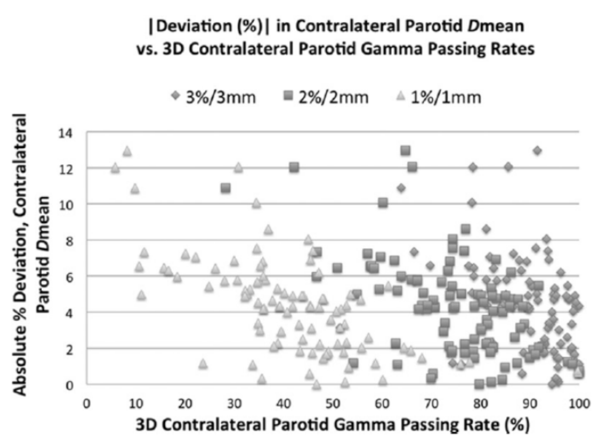
ROI- γ: 1% Local/1mm

CTV D95: r = -0.86

Zhen Med Phys 2011

3D Gamma vs DVH

For non-target structures, ROI-γ has little correlation to DVH.



ROI- γ: 1% Local/1mm

Spinal cord D1cc: r = -0.5

Spinal cord Dmax: r = -0.4

Parotid mean dose: r = -0.6

Larynx mean dose: no significant correlation

Zhen Med Phys 2011

Conclusions

- Alternatives to gamma analysis
  - Simple = ion chamber
  - Complex = 3D DVH measurement/calculation
- Commissioning
  - Take advantage of all possible QA routes
- Routine IMRT QA
  - Moving towards 3D DVH measurements

Thanks for your attention!