UNIVERSITY of MARYLAND
SCHOOL OF MEDICINE
DEPARTMENT OF RADIATION ONCOLOGY

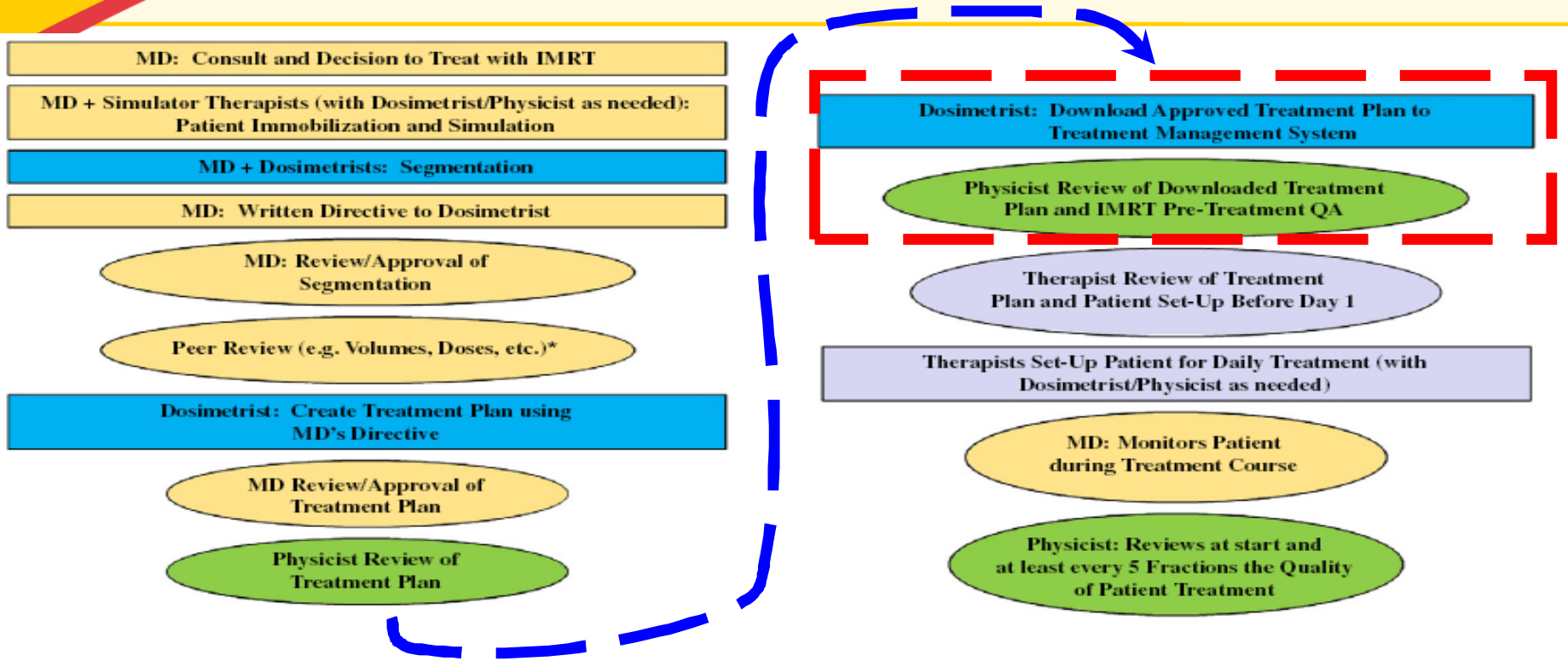
Is it safe, non-measurement based Patient Specific IMRT QA?

Byong Yong Yi, PhD

2018 Spring AAPM

Disclosure: None

Steps of IMRT and Scope of Talk



ASTRO White Paper, Safety considerations for IMRT, PRO 1: 190-195, 2011

Why Patient Specific QA (PSQA) for IMRT?

A series of New York Times articles:

Hazard to patients when patient-specific IMRT QA was not performed after a change to a patient's treatment plan was made.

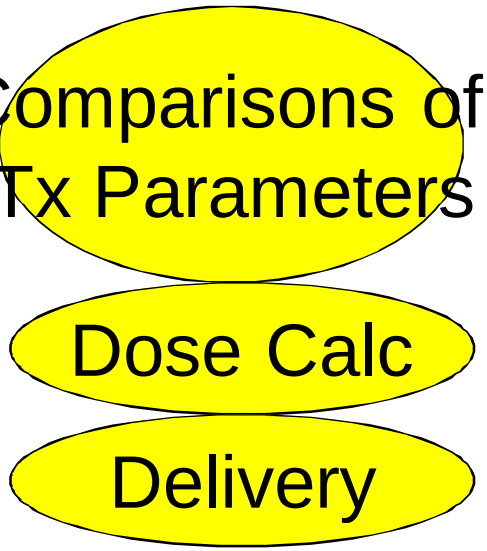
AAPM TG 218 Tolerance for measurement-based IMRT PSQA , 2018

Essential Components of IMRT PSQA

Comparison of Planned and Transferred including,
Gantry, Coll, Table, FS, MU, MLC, fractional MU...

Dose Calculation and/or MU calc Verification

Delivery Verification



*Dose comparison is one of the PSQA components.
Comparison of plan parameter is often less emphasized.

AAPM Report 82, IMRT Guidance document 2003

Are these valid statements for PSQA?

Measurement is ALWAYS the ground truth in physics world;

However, Is measurement the ONLY method?.

1. *Only measurements can check defects of the commissioned planning system: TPS dose model may not have been commissioned as desired ex) small size segments,*

2. *Measurement is the only way to confirm that the movement is fast and accurate enough to produce the planned delivery,*

3. *Current measurement methods of PSQA measure actual delivered dose and confirm the congruence of beam parameters between planned and delivered.*

PSQA Methods

- Measurement based
  - Phantom: Slab Phantom, Cubic, Cylindrical
  - Detector: Chamber and Film
    - 2D Detector Array
- Calculation based
  - In house Monte Carlo
  - Independent commercial 3D Software's
- Hybrid or reconstruction
  - Calculation using delivered information such as;
    - EPID images, Dynalog files etc

AAPM TG 119, 218 and
ASTRO White Paper

'How to' is still on going discussion issue

AAPM Report 82 (2003): Measurement and Calculation with frequent QA

ACR/ASTRO Guideline (2009) : Measurement is only mentioned

ASTRO White Paper (2011): Measurement and Calculation

TG 218 (2018): Tolerance of Measurement and Hybrid

AAPM Point/Counterpoint;

It is STILL necessary … measurements before delivery	2011
PSQA should be performed using software	2013

2018 AAPM Spring Meeting Best Poster Competition

PO-BPC-Fyer-31	Hybrid QA
PO-BPC-Fyer-14	Comparison of Software, Measurements and Hybrid
PO-BPC-Fyer-11	Quantitative Evaluation not using gamma

Limitations of Suggested Measurement Methods

*-Only 2D; Not whole Volume (*Hybrid methods may.),*

-Setup Uncertainty,

- Measurement dose not necessarily compare ALL of the parameters, especially MLC positions and/or Gantry positions and their MU partial weights,

-One 'fixed' phantom condition: Not considers various clinical situations.

Lack of transfer of the MLC files is a known cause of a catastrophic failure. (ASTRO White Paper)
Then, what about the lack of confirmation of all of these parameters? → *Assumption that the parameters have been confirmed since the measurement PSQA passed may not be correct.*

QA Program will be weakened if;

Measurement based	Non-Measurement based
Wrong detector: poor resolution or inadequate spacing for the gradients in the intensity maps	*Poor algorithms which make them inadequate for dosimetric verification of complex geometries*
QA failures are approached solely by repeating measurements at multiple different positions in the dose distribution until a point passes rather than identifying the root cause	
or QA failures are approached by the application of too generous dose/distance criteria for agreement	
Not checking the accuracy of the data transfer to the treatment management system.	

ASTRO White Paper, 2011

Potential Myths on Delivery Test by Measurements - I

1. Only measurements can check defects of the planning system: TPS dose model may not be commissioned as desired ex) small size segments

→ Calculation-based method also can pickup, if it is independent and well tested.

(One of the prerequisites for Calculation-Based PSQA)

2. Delivery test is essential: Measurement is only way to confirm that the movement is fast and accurate enough to produce the planned delivery

→ The TPS and/or the checking software should be able to pick up the machine limitations;

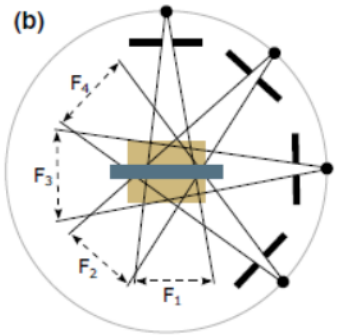
More frequent and comprehensive MLC QA should pick up failures prior to treatments.

Siochi and Molineu, PSQA using Software, Med Phys.2013

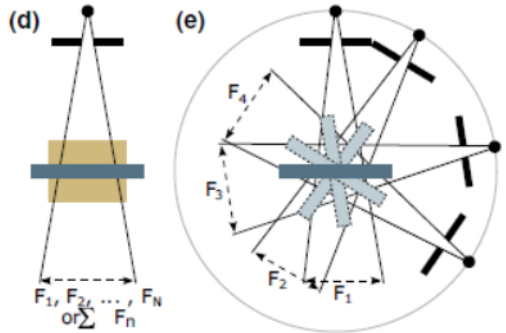
Potential Myths on Delivery Test by Measurements - II

3. Current measurement methods of PSQA measure actual delivered dose and confirms the congruence of beam parameters between planned and delivered;

→ *Only 2D measurements*



Total Composite

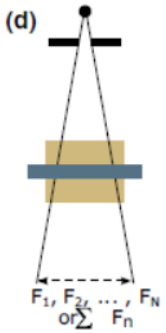


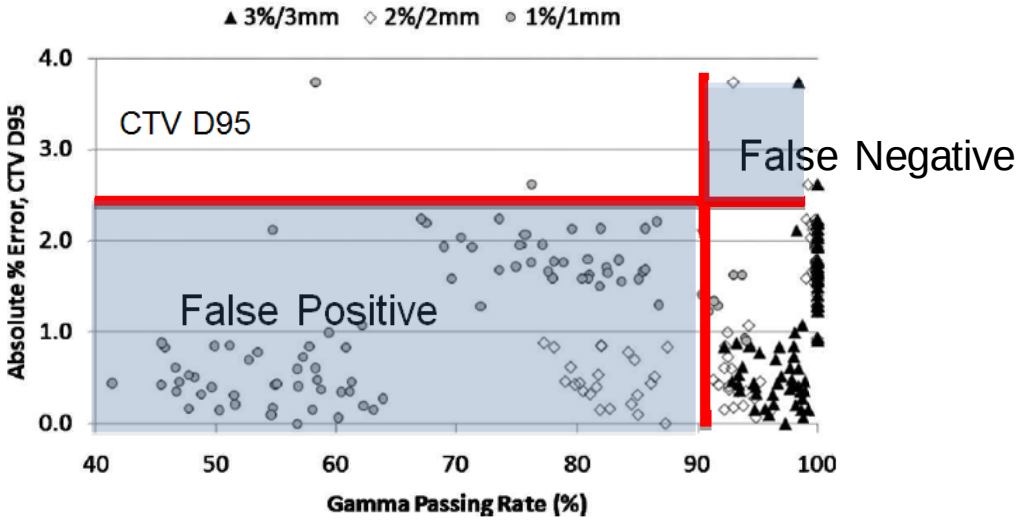
Perpendicular Field-by-by Field and Perpendicular Composite

From TG 218, 2018

→ *Limited spatial resolutions: May not sensitive enough to detect small variations*

continued





Gamma index per beam QA often shows less sensitivity.

Nelms, Per Beam, Planar PSQA MedPhys 2011

Practical questions on Measurement-based PSQA - I

We need measurements because this measurement will confirm if the plan is deliverable, the delivered dose will be as planned and the Tx machine functions as intended.

1. More than one linear accelerators with same beam quality:

**Patient may be treated at a different treatment machine than it is QA'ed.*

Do we need to QA again if we want to switch the machine?

2. Fractionated Treatments: More than one fractions

Is one time QA sufficient since we do not know when the machine malfunctions?

Practical questions on Measurement-based PSQA - 2

3. What should we do when fails: Re-measurements?

-Pulliam, 2012

- *13,002 PSQA and 302 (2.3%) failure*
- *222 cases passed after repeated measurements (1.7%)*
- *Final failure: 0.6%*

-McKenzie, 2014

- *Average of reproducibility of the PSQA is less than 2%*

→ *This explains that more than 2/3 of Pulliam's failed cases passed (above) after repeated measurements*

Calculation Based may be a solution; Advantages

Dose comparison with the patient CT geometry, not phantom,

Heterogeneity Correction Considered,

It may pick up various clinical situations, which phantom based measurements may not be possible (ex: bolus)

DVH comparisons,

Dose comparison of each anatomy,

Done at office hour,

No setup error,

…

Prerequisites for Calculation-based PSQA

Commissioning of an independent dose calculation engine;

→ Same efforts as RTP commissioning

→ Comparisons of enough cases of Measurements

Comprehensive Machine QA, especially MLC QA

→ ex) Weekly MLC QA

A software tool which confirms the beam deliverability and compares the treatment parameters

University of Maryland Solution from 2009

-DICOM RT comparison: RTP vs Aria

Comparisons of Tx Parameters

-In-house Monte Carlo Engine (Naqvi 2003 Phys Med Biol)

Commercial Dose calculation SW from 2015

3D Dose Calc and 3D Comparison

Comparison of Dose Distribution

3D Gamma Calculation

Dose Calc

Robust and Comprehensive Machine QA

If fails or questions then to measure as one of the steps to find the causes

Delivery

If verification fails;

To find root causes considering clinical scenarios;

Same set of CT, planning etc,

Bolus,

Patient support devices.

Considering limitations of the both of the calculation engines;

CT Number to density table,

Leaf Transmission, etc

Never to consider MU scaling unless it is confirmed to be necessary.

Reasons identified and these are verified from measurements.

Measurement;

Same as the measurement-based method.

In-house Software for Plan Parameter comparison

Plan: Plan Name, Number of Fractions, Beam Names

Radiation: MU, Radiation Type, Energy,

Segment: Numbers, Weights, Segment shapes (MLC)

Patient Orientation, Bolus

Machine Geometry: Coll, Gant, Couch, SSD

Field Size: X and Y Jaw

Isocenter Coordinates including Tx Fields and Images

…

Commissioning of Dose Calc Engine

Monte Carlo Engine

- *MC Engine Development: Naqvi, Phys Med Biol 2003*
- *Dosimetry test and MLC model confirmation*
- *PSQA comparisons using a same phantom: ~ 200 cases*
  - Film & Chamber: 150 cases
  - 2D array detector: 50 cases

Commercial Dose Calculation System

- *Beam and MLC model from the company.*
- *Same procedure as a RTP commissioning,*
- *Comparison to Planning system commissioning data,*
- *Measurements vs Calculations using a same phantom: ~40 cases*

U Maryland Experience

Commissioning of a Commercial system

-Commissioning tests were done to validate the model.

- *Configuration (PDDs, OF and OAR)*
- *Quick calcs test for TMRs and OFs.*
- *Phantom Plans (TPS vs Mobius vs Measurement)*

  Homogeneous (11 plans) and Heterogeneous (1 Plan)

  → *all energies and for 2 RTP systems*

-*Patient Plans (36 cases)*

-Also to understand the limitation of the calculation engine.

ex) Lower dose in air and in bone

U Maryland Experience

Development of Weekly MLC QA

Test Items…

- *Motion range of MLCs and Leaf bank,*
- *Extreme situations like gravity and interleaf interaction,*
- *Picket Fence,*
- *VMAT functionality,*

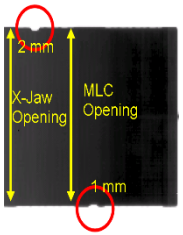
Practicalities…

- *Weekly using EPID: Monthly is too long and daily is not practical,*
- *Reasonable time for delivery and analysis: Therapist delivers Physicists checks,*
- *Not be too difficult to implement,*
- *Qualitative with the level of quantitative verification.*

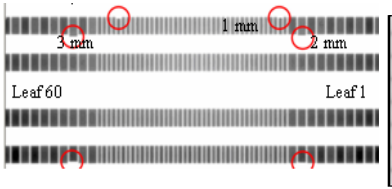
U Maryland Experience

University of Maryland - Weekly MLC Test Design

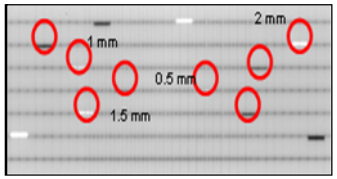
- MLC position
- MLC motion interdigitation
- Picket Fence
- Leaf-End leakage
- VMAT QA



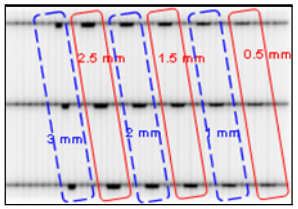
I. MLC position



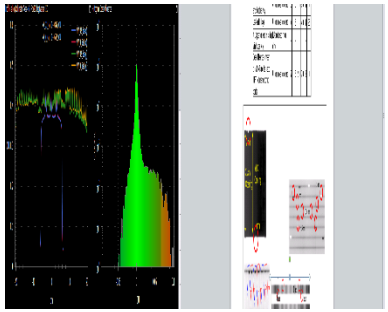
II. MLC motion interdigitation



III. Picket Fence



IV. Leaf End Leakage



V. VMAT: Speed of Gantry and Dose Rate

U Maryland Experience

Analysis of Weekly MLC QA

24 (3%) cases failures for 794 weekly QA;

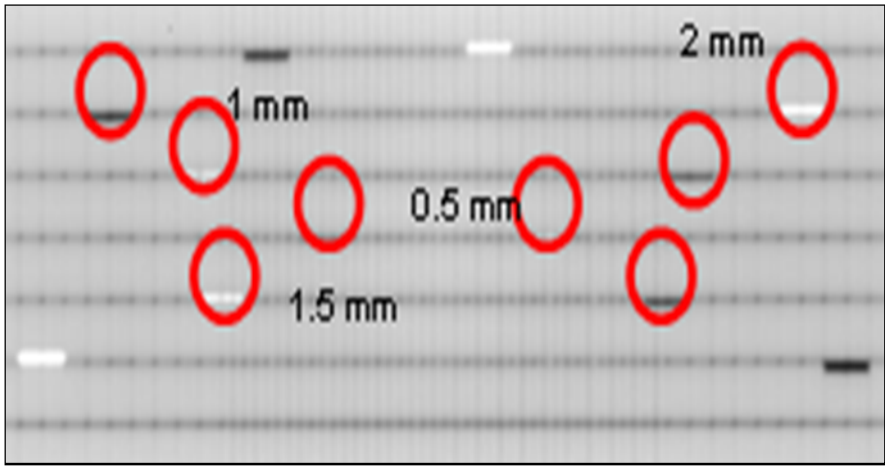
- *MLC Position 0.2%*
- *Interdigitation 0.1%*
- *Picket Fence 0.1%*
- *Leaf End Leakage 0%*
- *VMAT QA 2.5%*

Kalavagunta, Is Weekly MLC QA Necessary? AAPM 2016

U Maryland Experience

One example which the Weekly QA picked up

- *MLC system was upgraded from v6 to v7.6.*
- *MLC leaf gap was incorrectly setup during upgrade by the engineer.*
- *Weekly MLC QA after upgrade picked up a larger leaf gap than expected using the Picket Fence Test.*

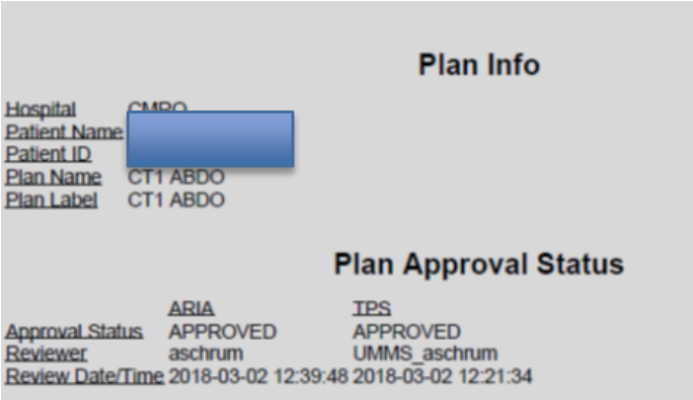


Kalavagunta, Is Weekly MLC QA Necessary? AAPM 2016

U Maryland Experience

Comparison of Plan Parameters RTP vs R&V -I

An in-house program to compare DICOM RT files (Aria vs TPS)



Plan Info

Hospital	CMRO
Patient Name	[redacted]
Patient ID	
Plan Name	CT1 ABDO
Plan Label	CT1 ABDO

Plan Approval Status

	ARIA	TPS
Approval Status	APPROVED	APPROVED
Reviewer	aschrum	UMMS_aschrum
Review Date/Time	2018-03-02 12:39:48	2018-03-02 12:21:34

Patient Name
Plan Name,
Approval Status
Creation Time

Comparison of Plan Parameters RTP vs R&V -II

An in-house program to compare DICOM RT files (Aria vs TPS)

# of Fractions, Beam Names
MU, Radiation Type, Energy,
Segment: Numbers, Weights, MLC
Patient Orientation, Bolus
Geometry: Coll, Gant, Couch, SSD
Field Size: X and Y Jaw

U Maryland Experience

Summary

These two plans are *IDENTICAL* in terms of MU, MLC shape, energy, collimator angle, gantry angle, gant rotation, couch angle, SSD, jaw positions, iso center, segment weights, wedge, bolus, patient position, applicator.

Beam List

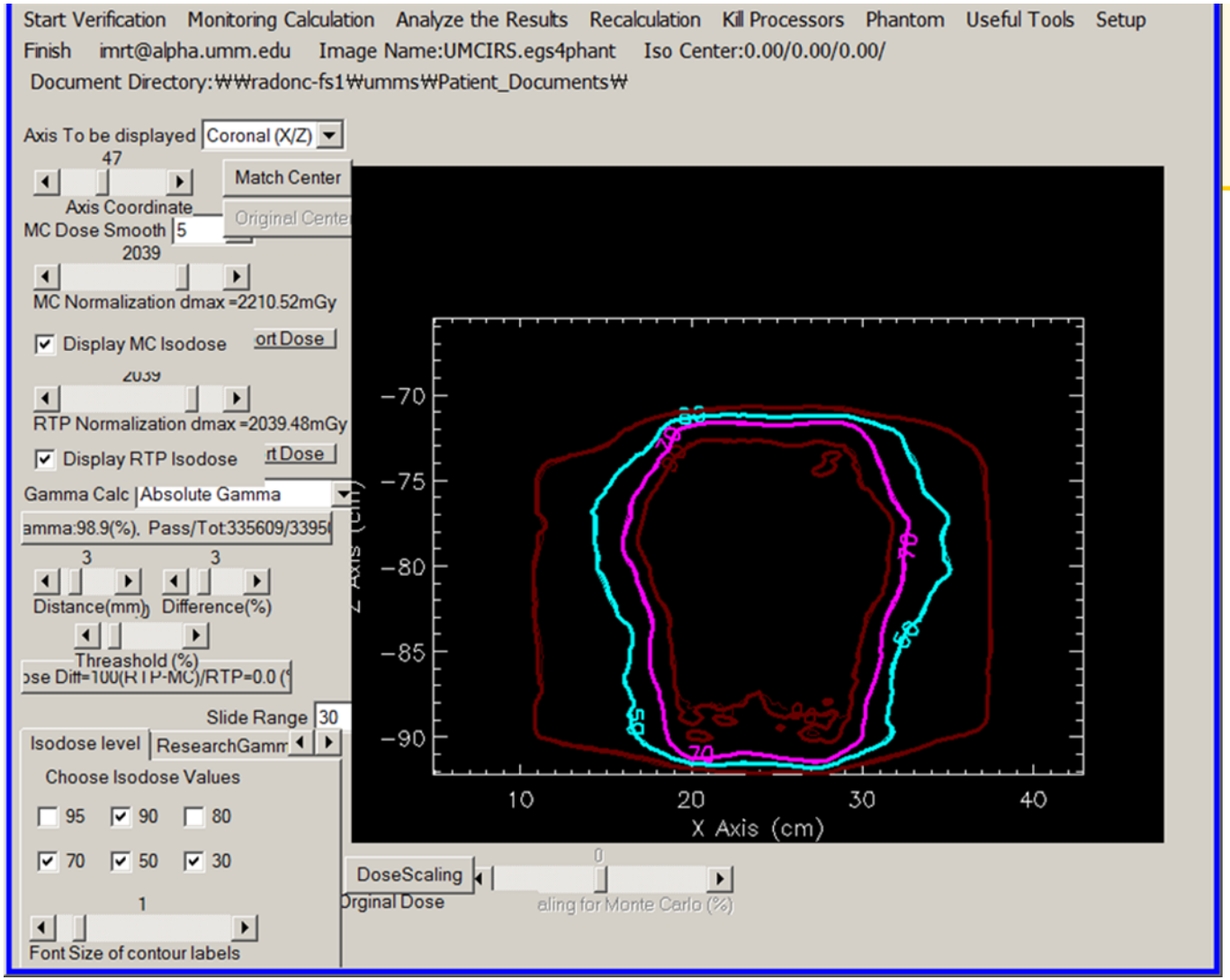
Beam	Parameter	Aria	TPS	Status	Tolerance
181-179 A					
	Fx #	23	23	✓	
	MU	175.67	175.67	✓	0.5
	Radation Type	PHOTON	PHOTON	✓	
	Segment#	180	180	✓	
	Energy	6	6	✓	0
	Collimator Angle(°)	350	350	✓	0.1
	Gantry Angle(°)	181/179	181/179	✓	0.1
	Gantry Rotation Direction	CW	CW	✓	
	Couch Angle(°)	0	0	✓	0.1
	SSD(cm)	78.60	78.60	✓	0.1
	Jaw(X)(cm)	(-8.37, 8.59)	(-8.37, 8.59)	✓	0.01
	Jaw(Y)(cm)	(-7.5, 11)	(-7.5, 11)	✓	0.01
	Iso Center(cm)	(3.88, 8.54, -128.04)	(3.88, 8.54, -128.04)	✓	0.01
	Max MLC Difference(cm)		0.0005	✓	0.01
	Max Segment Weight Difference(%)		0	✓	0.001
	Patient Setup Position	HFS	HFS	✓	
179-181 A					
	Fx #	23	23	✓	

Comparison of Plan Parameters RTP vs R&V -III

Isocenter Coordinates: Important for CBCT and kV

Beam isocenter coordinates comparison

Plan(Aria)	Beam	Type	Tol.Table	Aria	TPS	Machine(Aria/TPS)
CT1_ABDO	181-179 A	TREATMENT	Photons	(3.88, 8.54, -128.04)	(3.88, 8.54, -128.04)	VGS_CLINAC:UMMS_Clinac
CT1_ABDO	179-181 A	TREATMENT	Photons	(3.88, 8.54, -128.04)	(3.88, 8.54, -128.04)	VGS_CLINAC:UMMS_Clinac
CT1_ABDO	0 SU	SETUP	OBI	(3.88, 8.54, -128.04)	(3.88, 8.54, -128.04)	VGS_CLINAC:UMMS_Clinac
CT1_ABDO	CBCT	SETUP	OBI	(3.88, 8.54, -128.04)	(3.88, 8.54, -128.04)	VGS_CLINAC:UMMS_Clinac
CT1_ABDO	270 SU(87.6)	SETUP	OBI	(3.88, 8.54, -128.04)	(3.88, 8.54, -128.04)	VGS_CLINAC:UMMS_Clinac

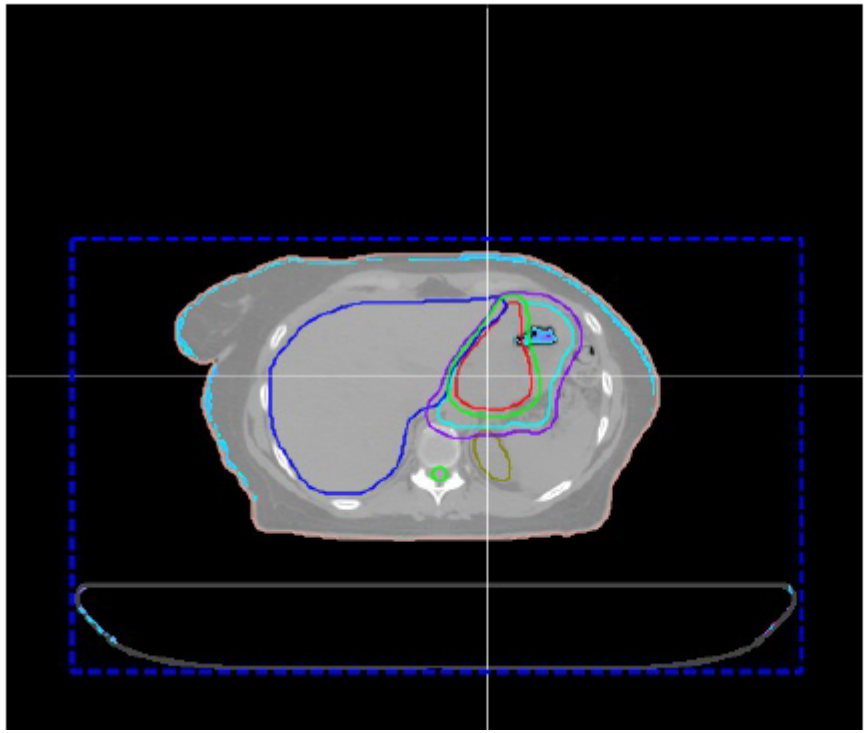


UMIV with MC

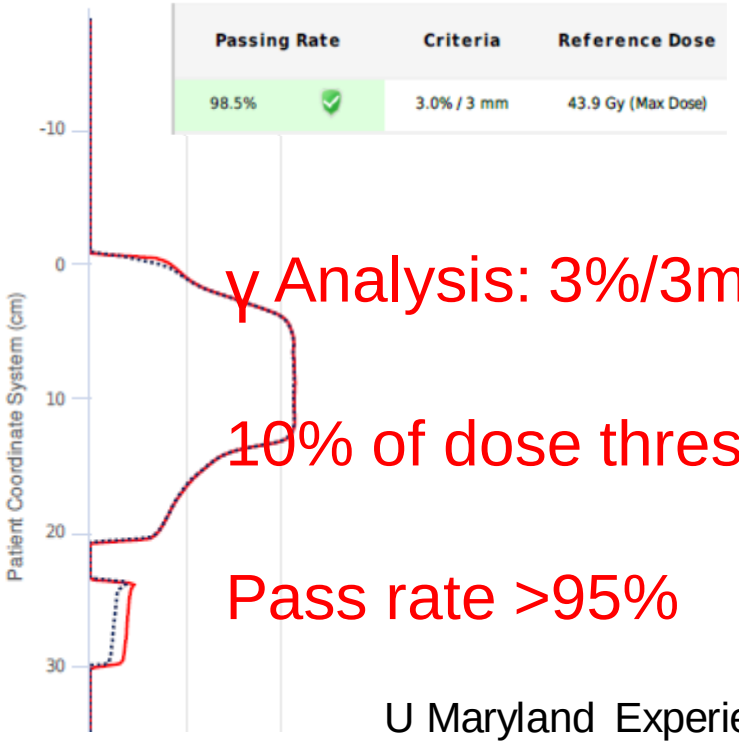
U Maryland Experience

Dose Calc: 3D Gamma Analysis

Transverse Plane at 0 cm from Isocenter



Vertical Dose Profile



γ Analysis: 3%/3mm

10% of dose threshold

Pass rate >95%

U Maryland Experience

Target Doses and Point Dose

TPS Name	Mean Dose			95% Coverage			Stray Voxel
	TPS	M3D	% Diff	TPS	M3D	% Diff	
CTV1	42.5 Gy	42.4 Gy	-0.09%	42.1 Gy	41.9 Gy	-0.4%	None
GTV	42.5 Gy	42.5 Gy	-0.08%	42.2 Gy	42.1 Gy	-0.22%	None
PTV1	42.2 Gy	42.1 Gy	-0.22%	41 Gy	40.6 Gy	-1.08%	None
PTV2	42.4 Gy	42.3 Gy	-0.16%	41.9 Gy	41.5 Gy	-0.73%	None

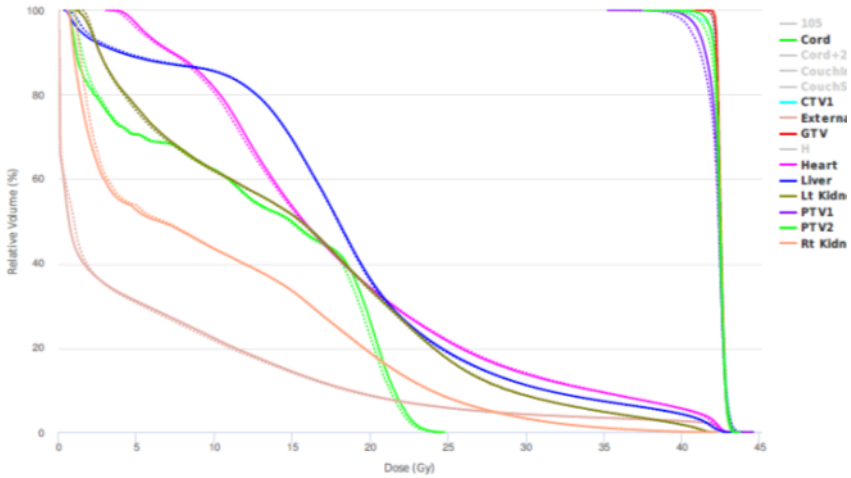
<3%

Beam	181-179 A	179-181 A
Energy (MV)	6	6
TPS MU	176	151
M3D MU	176	152
TPS Beam Dose (cGy)	105.28	79.25
M3D Beam Dose (cGy)	104.94	78.99
Dose Difference	-0.32%	-0.32%
Segments	179	179
X1 / X2 Jaws (cm)	8.4 8.6	8.7 8.3
Y1 / Y2 Jaws (cm)	7.5 11	7.5 11
Wedge	None	None
MLC	VMAT	VMAT
Rotation	VMAT	VMAT
Gantry	181° to 179°	179° to 181°
Collimator	350°	10°
Couch	0°	0°
Gantry Clearance (cm)	9.51	9.51

U Maryland Experience

Organ-at-Risk Dose <5%

TPS Name	Volume	3D Gamma (3.0% / 3 mm)	Mean Dose		
			TPS	M3D	% Diff
105	1.49 cc	100%	42.145 Gy	42.354 Gy	0.48%
Cord	22.1 cc	100%	12.654 Gy	12.593 Gy	-0.14%
Cord+2MM	40.5 cc	100%	12.64 Gy	12.577 Gy	-0.14%
CouchInterior	13497 cc	100%	1.554 Gy	0.504 Gy	-2.39%
CouchSurface	2150 cc	99.7%	1.393 Gy	0.386 Gy	-2.3%
CTV1	478 cc	99.3%	42.479 Gy	42.441 Gy	-0.09%
External	30130 cc	99.5%	5.814 Gy	5.783 Gy	-0.07%
GTV	216 cc	99.9%	42.501 Gy	42.465 Gy	-0.08%
H	0.23 cc	100%	37.384 Gy	37.43 Gy	0.1%
Heart	419 cc	100%	18.268 Gy	18.184 Gy	-0.19%
Liver	1475 cc	100%	18.502 Gy	18.499 Gy	-0.01%
Lt Kidney	131 cc	100%	15.359 Gy	15.292 Gy	-0.15%



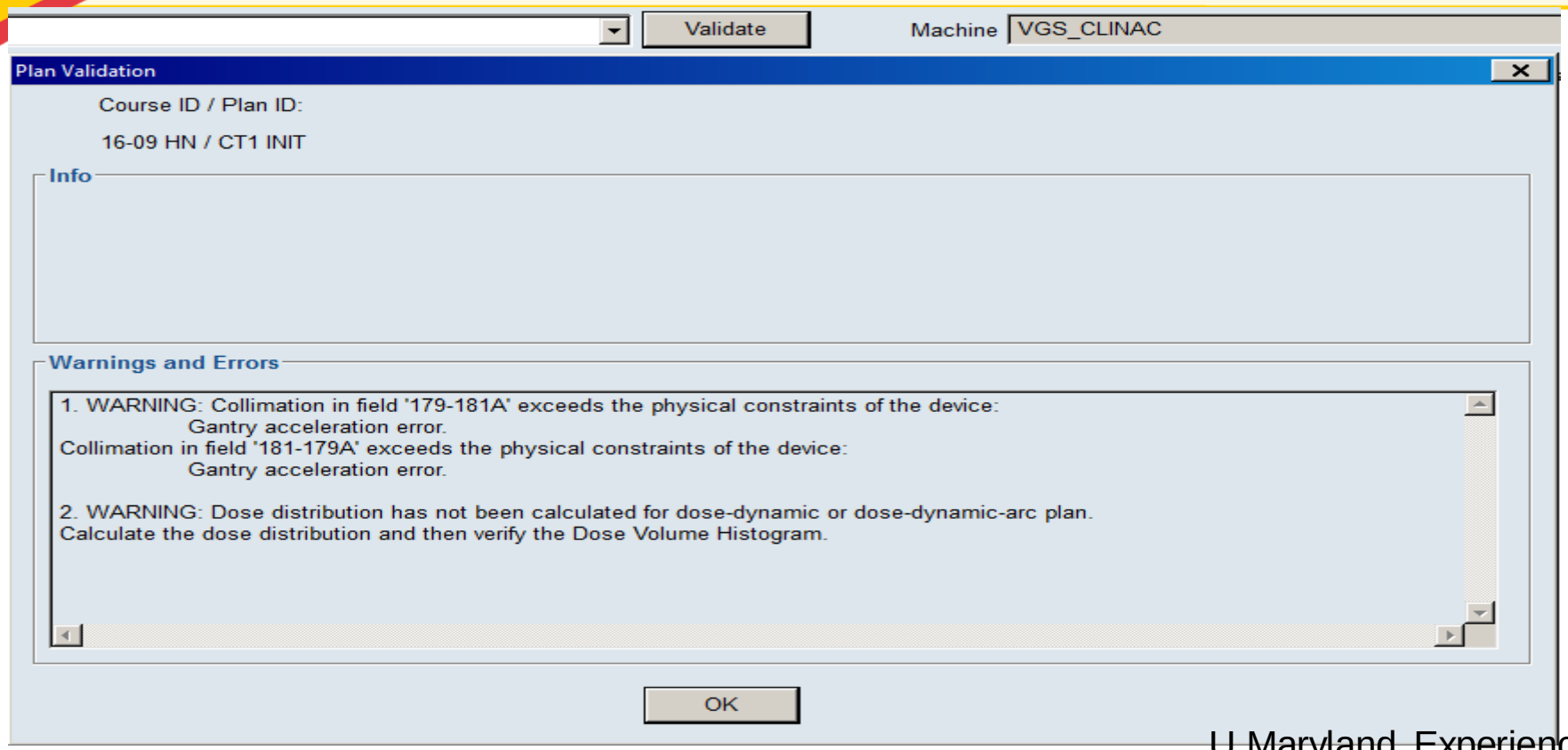
U Maryland Experience

Summary of Dose Comparison

Plan Name or Identifier:	CT1 Init		
Plan Type:	IMRT / VMAT		
Verification Method:	Mobius Secondary Calculation		
Target Volume:	**PTV1**		
Target Volume, Mean Dose:	**TPS (Gy)** **54.5**	**Mobius(Gy)** **54.1**	**Difference (%)** **0.7**
Target Volume, D95% Coverage:	**TPS (Gy)** **52.4**	**Mobius(Gy)** **51.5**	**Difference (%)** **-1.6**
Gamma Index Pass Percentage (%):	99.0	**DTA** = 3mm and **DIFF** =3%	
Point Dose (average over all beams) (%):	1.1		
All OAR Mean Dose Difference <5%	Pass	**Comments:**	
Plan parameter check:	Pass		
Required Modifications (if any):	N/A		
Dosimetric Feasibility of the Plan:	Plan is verified for treatment.		
Comments:			

U Maryland Experience

IMRT/VMAT Plan Deliverability

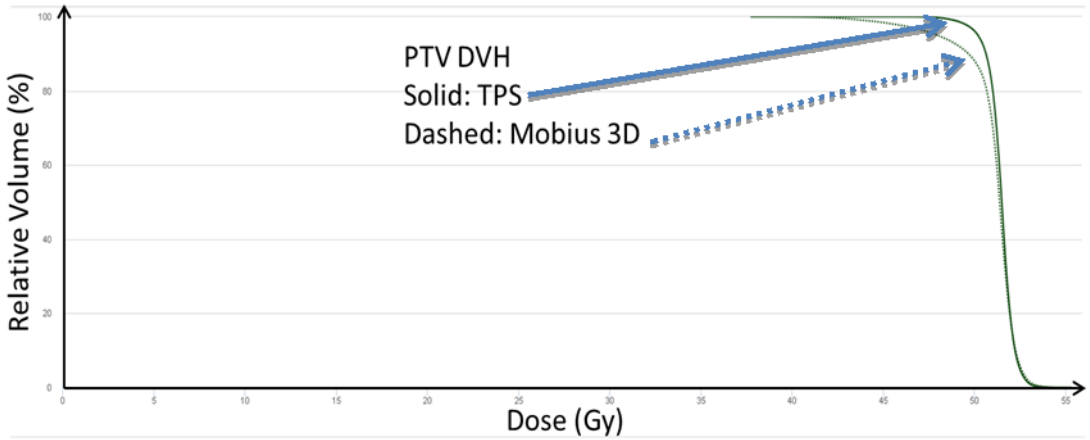


U Maryland Experience

Findings

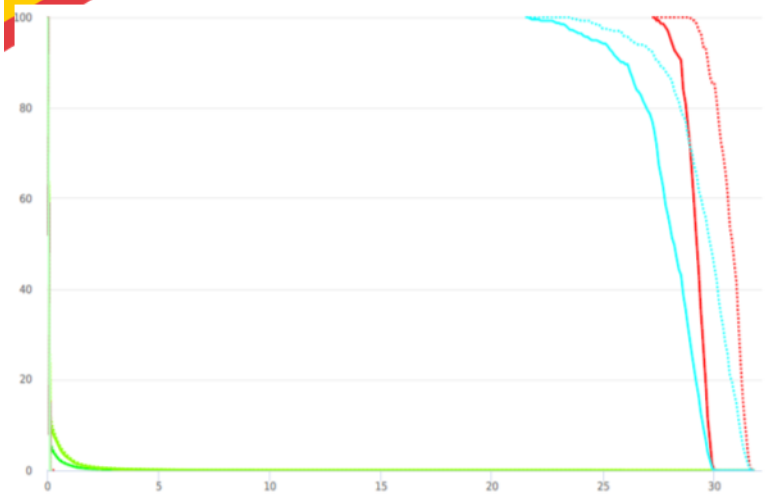
- 16 cases among 1200 cases of Calculation-based PSQA (1.3%)
  - *9 cases of suspicious calculations*
    - ❑ PTV extending to the skin
    - ❑ Heterogeneity correction
    - ❑ CT-to-density table
  - *5 cases of delivery failure*
    - ❑ Violate MLC leaf speed limitation
    - ❑ Violate MLC opposite leaf gap limitation
  - *2 data transfer errors*
    - ❑ Manually input of MU in ARIA
- No failure of weekly MLC QA

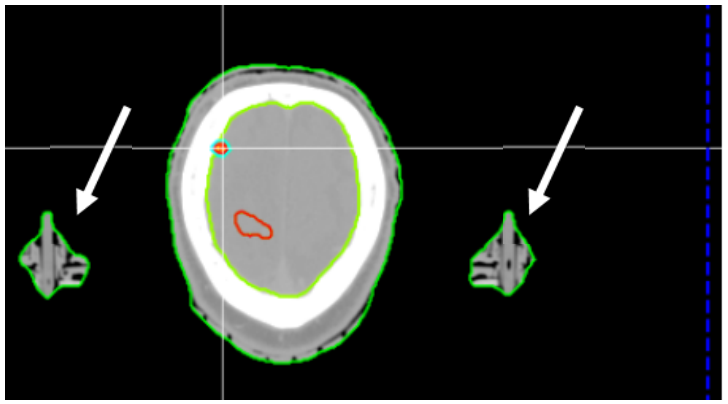
Cause of a failure of one case: bolus case

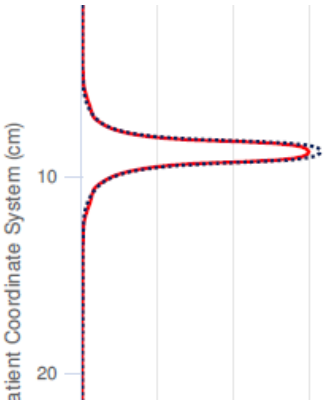


One example of DVH failure case: the difference between TPS and the verification calculation D95% (dose to 95% of PTV) was greater than 10%. The difference was due to the PTV extending to the skin but planned without bolus. The case was re-planned with bolus and QA was passed.

Cause of a failure :Cone SRS -I





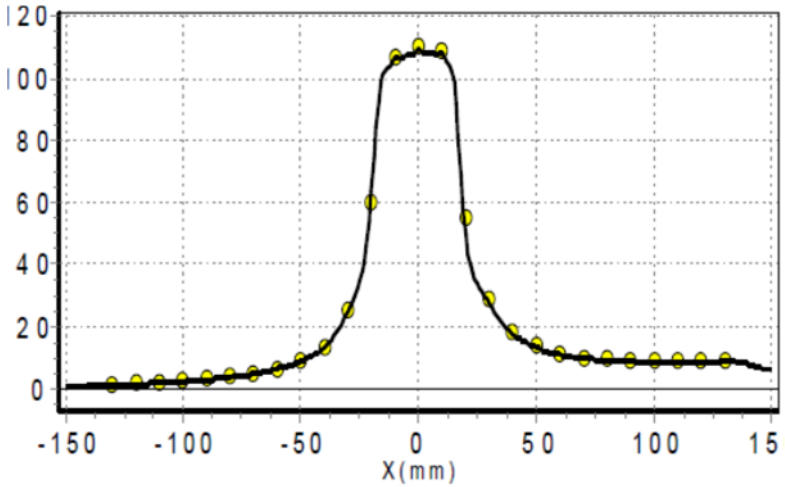


	PTV_RT_FRONT-ANT	27.9 Gy	29.5 Gy	5.38%	24.3 Gy	26.2 Gy	6.16%

Cone based SRS failed 5-6% . DVH

U Maryland Experience

Cause of a failure :Cone SRS -II



2D Measurements passed perfect. If it was the measurement-based, we may not have find the problem.

Cone-based planning system

Cone algorithm uses;

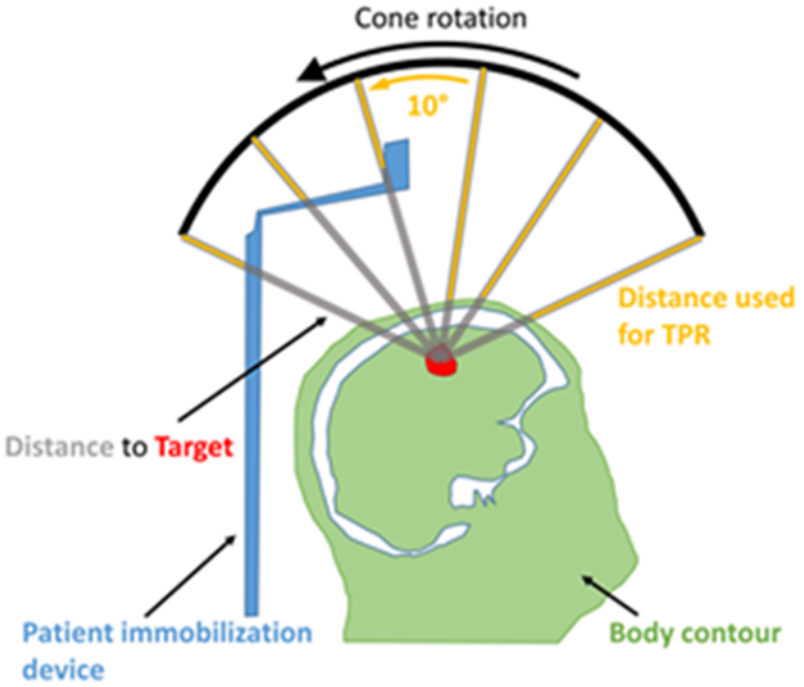
- TPR averaged per arc using first point of contact with body contour,

- No inhomogeneity correction,

- No HU override, either,

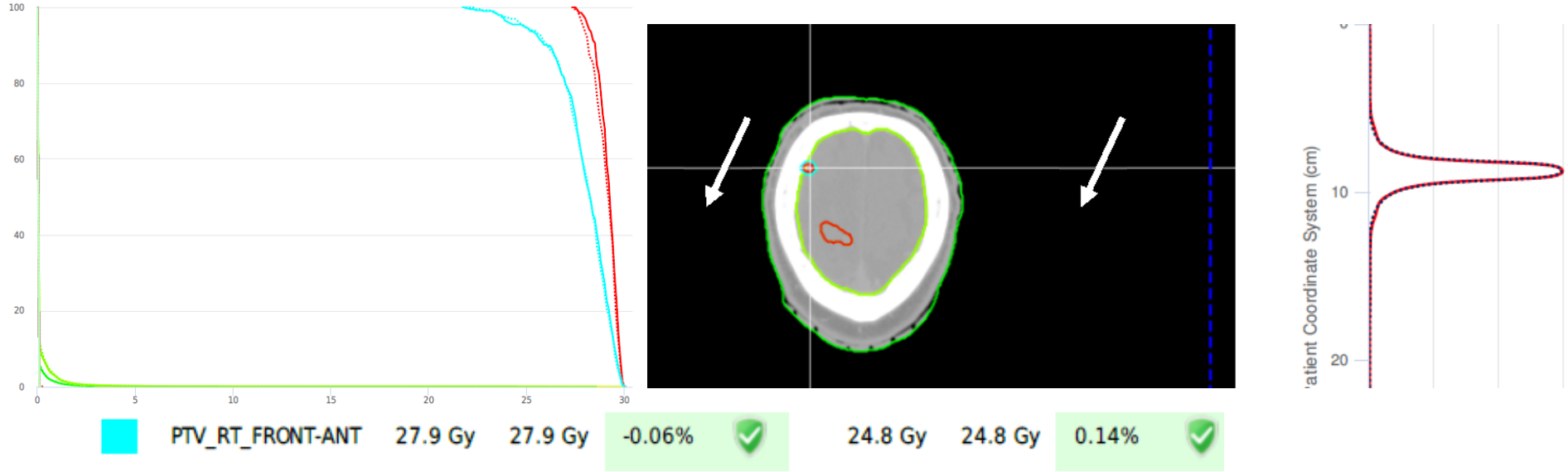
- No surface effect/angle of entry.

Patient support devices must be removed from the external contour, or calculated Mus aren't accurate.



U Maryland Experience

Cause of a failure :Cone SRS -III



Dose calculation after removing the patient support from the external contour.
Dose matched within clinically acceptable range.

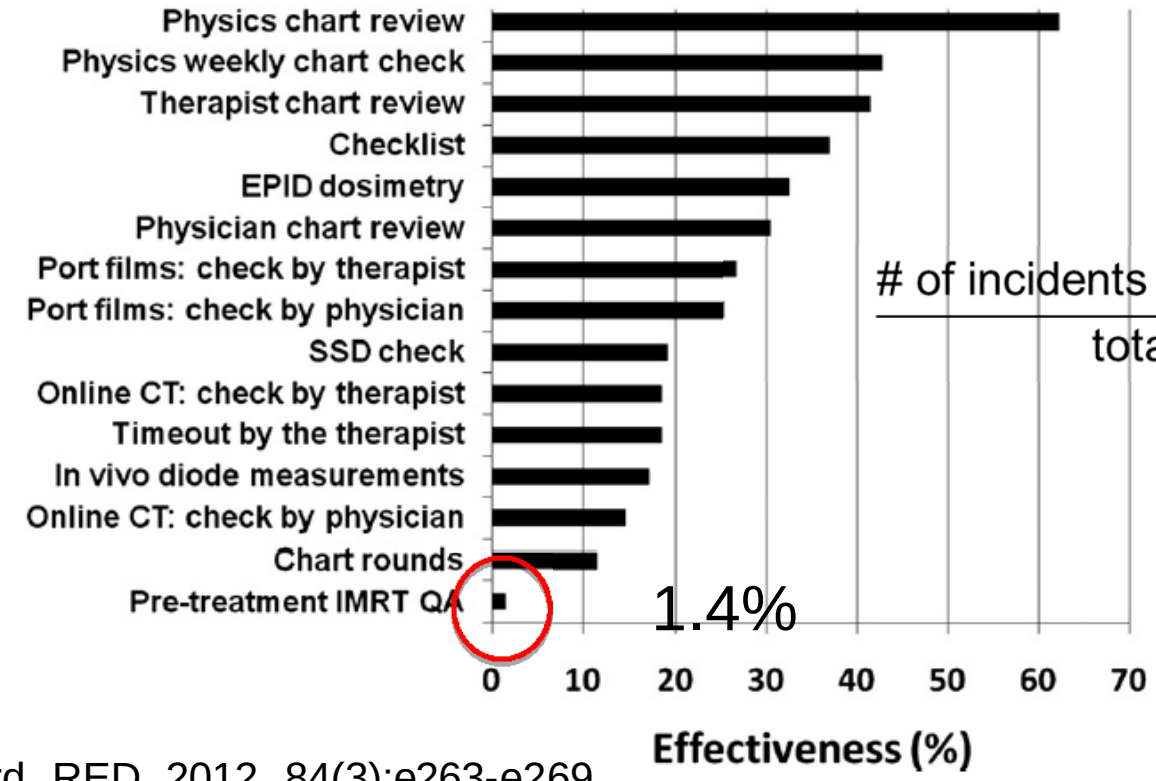
Comparisons: Measurement vs Calculation

- UM Experience: *1.3% Failure, no repeat but to try to find the cause(s) of the failure*
- Pulliam, 2012
  - *13,002 PSQA and 302 (2.3%) failure*
  - *222 cases passed after repeated measurements (1.7%)*
  - *Final failure: 0.6%*
- McKenzie, 2014
  - *Average of reproducibility of the PSQA is less than 2%*

→ *This explains that more than 2/3 of Pulliam's failed cases passed (above) after repeated measurements*

U Maryland Experience

Effectiveness of PSQA in detecting errors



$$\frac{\text{\# of incidents that each QC check could detect}}{\text{total \# of incident reports}}$$

4407 incidents

Ford, RED 2012, 84(3):e263-e269

Effectiveness of PSQA in detecting errors-Continured

Detecting power of PSQA is low: it detects only 1.4% of the cases

Physics chart review detects more than 60% of the potential errors

A physicist can distribute the time resources better by allocating more ont the items less emphasized, such as Rx'ed dose, beam geometries, MLC patterns using saved time from non-measurement-based PSQA

Ford, RED 2012, 84(3):e263-e269

Summary

Measurement-based PSQA has a few shortcomings,

We can achieve similar results, even better in some occasions, using calculation-based PSQA,

It picks up a few issues of various of clinical situations,

Before launch to the clinic, calculation-based PSQA requires;

Adequate commissioning of the dose calculation engine,

Comprehensive Machine QA, especially MLC QA,

A software tool which confirms the beam deliverability and compares the treatment parameters,

Measure when fails or questions.

University of Maryland experience shows that the calculation-based PSQA may be applied to the clinic safely, if above prerequisites are met.

Acknowledgement

Physics colleagues of University of Maryland

Special thanks to whom shared their slides;

Chai Kalavagunta, Ph.D.

Karl Prado, Ph.D.

Mariana Guerrero , Ph.D.

Shifeng Chen, Ph.D.

Yannick Poirier , Ph.D.

References quoted

AAPM Report 82, IMRT Guidance document 2003
AAPM TG 119 IMRT Commissioning, 2009
ASTRO and ACR Guideline for IMRT, RED. 2009;73(1):9-14
Rangel, Sensitivity of PSQA on MLC, Med Phys. 2010;37(7):3862-3867
ASTRO White Paper, Safety considerations for IMRT, PRO. 2011;1: 190-195
Smith, Dieterch, Measurement necessary before IMRT, Med Phys. 2011;38(2):553-555
Nelms, Per Beam, Planar PSQA, Med Phys. 2011;38(2):1037-1044
Ford, RED 2012, 84(3):e263-e269
Pulliam, 13,000 PSQA Results Med Phys. 2012;39(6):3866-3867
http://rpc.mdanderson.org/RPC/Publications/RPC_Presentations/2012%20AAPM%20Posters/AAPM%20Presentation_KPulliam.pptx
McKenzie, Reproducibility in PSQA, JACMP. 2014;15(3):241-251
Pulliam, Comparison of 2D and 3D gamma Med Phys.2014;41:021710
Siochi and Molineu, PSQA using Software, Med Phys.2013;40(7):0706011-070603
ACR, Practice Parameter for IMRT (Revised) 2016
AAPM TG 218 Tolerance for measurement-based IMRT PSQA , 2018