Treatment plan complexity metrics for predicting IMRT pre-treatment quality assurance results

Scott Crowe, QUT

Pre-treatment QA

Verification of fluence delivery to ensure that dose can be delivered accurately

To identify:

- if leaf sequence doesn't produce expected fluence
- if plan isn't being correctly transferred to linac
- if treatment isn't being delivered correctly

What's popular for contemporary treatments?

- comparing 2D/3D measurements against predictions

Genesis QLD?

Current procedure involves evaluation of predicted portal dose against EPID images

- Therapists responsible for portal image acquisition (after RO approval of plan)
- Physicists responsible for analysis of results, where:
  - pass for IMRT = mean GAI > 90% and no beam GAI < 85%
  - pass for RapidArc = no beam GAI < 90%
  - both calculated using 3%, 3mm
  - done within the Epiqa software

GAI = gamma agreement index, % of (sig.) points in portal image that agree with predicted dose

Past procedures included MapCheck system

"Deliverability"

What are some possible limitations could prevent the linac from delivering the calculated dose?

- limitations in the machine – leaf movement, dose rate, gantry movement, etc.
- limitations in the software – fluence optimisation, dose calculation, etc.

Deliverability is related to complexity/modulation

Cost functions are used to control complexity in inverse planning algorithms

Potential for time saving

Since pre-treatment QA results correlate with complexity …

Identifying overly complex plans before QA could potentially save time.

Why bother with the QA if you expect the plan will be sent back for adjusting anyway?

Established measures

Modulation index – deviations between adjacent bixels against standard deviation of fluence map

Fluence map complexity – deviations between adjacent bixels against sum of each bixel

Modulation complexity score – indication of planned mobility of leaves in the beam (aperture area variations between control points and differences between adjacent leaf positions)

Novel measures

Novel metrics added to TADA code:

- Mean field area – segment weighted by MU
- Mean aperture displacement – mean of difference between centre of beam and centre of leaf apertures
- Cross axis score – % of leaf pairs blocking central axis
- Closed leaf score – % of leaf pairs in field closed
  - could indicate problem with transmission modelling
- Small aperture score – percentage of leaf pair apertures below a threshold (i.e. 1mm, 5mm)
  - could indicate limits of dose calculation accuracy
- and others

MapCheck results

Analysis of 122 prostate beams, planned in BrainLab

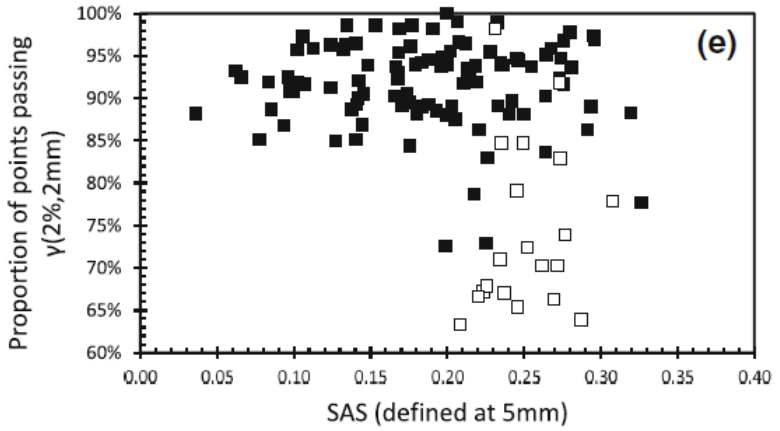
Modulation index and small aperture score significantly correlated to QA pass rates

SAS had useful threshold (beams with SAS(5mm)<0.27 all passed)

Table 1 Summary of mean plan complexity metrics for plans that passed ($Mean_{pass}$) and failed ($Mean_{fail}$) QA testing

Metric	$Mean_{pass}$	$Mean_{fail}$	F value	p value
MCS	0.37 ± 0.05	0.35 ± 0.05	0.140	0.646
AAV	0.49 ± 0.06	0.47 ± 0.06	0.097	0.756
LSV	0.75 ± 0.02	0.75 ± 0.02	0.506	0.478
MI	0.012 ± 0.002	0.015 ± 0.002	**11.397**	**0.001**
FMC	0.008 ± 0.001	0.008 ± 0.001	0.002	0.965
MFA	1300 ± 200	1200 ± 200	5.439	0.021
SAS (2 mm)	0.17 ± 0.06	0.21 ± 0.03	3.117	0.080
SAS (5 mm)	0.19 ± 0.06	0.25 ± 0.03	**9.918**	**0.002**
SAS (10 mm)	0.26 ± 0.07	0.32 ± 0.05	6.163	0.014
SAS (20 mm)	0.50 ± 0.09	0.56 ± 0.07	2.984	0.087
CLS	0.10 ± 0.04	0.12 ± 0.05	0.346	0.558
CAS	0.65 ± 0.09	0.68 ± 0.08	1.818	0.180
MAD	20 ± 4	21 ± 4	0.859	0.356

F statistic tests for significant linear relationship between the metric and γ pass rates. Critical F statistic values were 3.92 for $\alpha = 0.05$, 6.85 for $\alpha = 0.01$, and 11.38 for $\alpha = 0.001$. The p value represents the probability of a larger F statistic occurring by chance. Significant F statistic values ($\alpha = 0.0025$, after Šidák correction to experiment-wide $\bar{\alpha} = 0.05$) are shown in boldface type



Epiqa results

Analysis of IMRT and Rapidarc Plans over multiple treatment sites.

Table 1. Plans & sites.

Treatment Site	# IMRT plans	# IMRT beams	# VMAT plans	# VMAT beams
Prostate	30	210	40	80
Head & Neck	3	22	11	24
Anus/Rectum	6	79	5	12
Endometrium	3	35	2	4
Brain	10	56	12	24
Total	52	402	70	144

Modulation index distinguished between treatments for different sites

- more complex: anus/rectum, endometrium
- less complex: prostate and brain

Epiqa results

Table 4. Summary of F statistic tests for significant linear relationships between metric values and GAI for beams. Significance is evaluated against Šidák corrected α values of approximately 0.004 (*), 0.0008 (**) and 0.00008 (***), corresponding to experiment-wide α values of 0.05, 0.01 and 0.001 respectively.

Metric	VMAT 3%,3mm	VMAT 2%,2mm	Sub-arc 3%,3mm	Sub-arc 2%,2mm	IMRT 3%,3mm	IMRT 2%,2mm
# Plans	72	72	19	19	52	52
# Beams	150	150	207	207	402	402
MU	ns	ns	ns	ns	***	***
MFA	ns	ns	ns	ns	ns	ns
AAV	*	ns	ns	ns	***	***
CLS	ns	ns	ns	ns	ns	ns
CAS	*	ns	ns	ns	***	***
FMC	ns	ns	***	**	ns	ns
MAD	ns	*	ns	ns	***	***
MCS	***	**	***	***	***	**
MI	ns	ns	ns	ns	***	***
SAS1	ns	ns	*	*	***	***
SAS5	ns	ns	ns	ns	***	***
SAS10	**	ns	ns	ns	***	***

Strong correlation for many metrics for IMRT treatments

Weak correlation for some metrics for RapidArc treatments, even when split up into sub-arcs.

CAS / MAD: interleaf leakage not modelled in Eclipse, which has a greater effect for IMRT, as collimator angles chosen for RapidArc to minimise effect

Other observations – arcs

Representative sample of arcs re-QA'd after beams split into 6× 60° sub arcs: sub-arc GAI values significantly lower

Table 2. Significance of difference is evaluated between whole and divided sub-arc pairings using a two-tailed Welch's t-test.

	3%,3mm	2%,2mm
Whole arc (360°) GAI	95.9%	90.0%
Sub-arc (60°) maximum GAI	96.0%	89.2%
Significance	ns	ns
Sub-arc (60°) mean GAI	93.5%	85.4%
Significance	**	***
Sub-arc (60°) minimum GAI	90.6%	80.8%
Significance	***	***

Other observations – GAI

2%,2mm & 3%,3mm were both used in Epiqa

GAI(2%,2mm) ≈ 1.08 × GAI(3%,3mm)

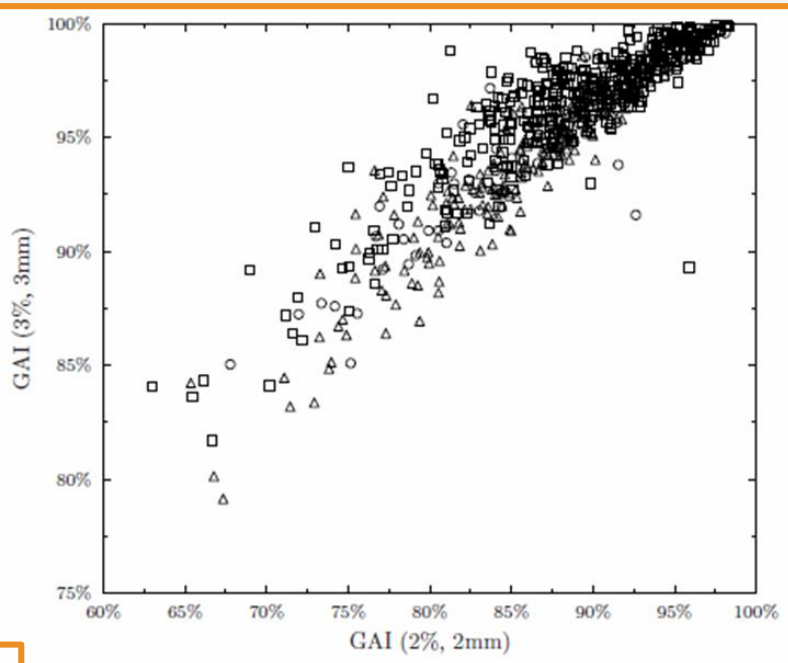


Figure 1. Relationship between GAI calculated using acceptance criteria of 3%,3mm and 2%,2mm. Linear coefficient for all data is 1.085 ± 0.002, with a coefficient of determination of 0.998.

Table 3. m is the linear coefficient and r^2 is the coefficient of determination.

Treatment	3%,3mm Mean GAI	2%,2mm Mean GAI	Linear Coefficient	r^2
IMRT	96.48%	88.91%	1.082 ± 0.002	0.998
Whole arc (360°)	95.34%	87.78%	1.083 ± 0.005	0.998
Sub-arc (60°)	93.49%	85.32%	1.094 ± 0.003	0.999
Total	95.44%	87.70%	1.085 ± 0.002	0.998

References

- Crowe et al. 2014. Treatment plan complexity metrics for predicting IMRT pre-treatment quality assurance results. *Australas Phys Eng Sci Med* (in press)
- Kairn et al. 2014. Predicting the likelihood of QA failure using treatment plan accuracy metrics. *J Phys Conf Ser* 489(1), 012051

- Kairn et al. 2014. Experienced-based management of IMRT quality assurance. ***Upcoming CSM poster!***