



Considerations for the verification of volumetric modulated arc therapy-planned dose distributions

**Authors:**Lourens Strauss¹ William Shaw¹ **Affiliations:**

¹Department of Medical Physics, University of the Free State, Bloemfontein, South Africa

Corresponding author:

Lourens Strauss,
strausslj@ufs.ac.za

Dates:

Received: 23 May 2019

Accepted: 16 June 2019

Published: 02 Sept. 2019

How to cite this article:

Strauss L, Shaw W.
Considerations for the verification of volumetric modulated arc therapy-planned dose distributions. S. Afr. j. oncol. 2019;3(0), a86. <https://doi.org/10.4102/sajo.v3i0.86>

Copyright:

© 2019. The Authors.
Licensee: AOSIS. This work is licensed under the Creative Commons Attribution License.

Background: Volumetric modulated arc therapy (VMAT) is the standard of care for many clinical indications, but should only be considered with proper technical support and quality assurance (QA) in place. Despite the high accuracy of VMAT systems, errors can be present and adequate verification is required. Dosimetric VMAT verification systems have a broadly similar analysis philosophy. However, many factors influence the analyses and the subsequent QA outcome, based on which the plan will pass or fail.

Aim: This study investigated various factors that influence the dosimetric impact and detectability of known linac component deviations on VMAT QA, including geometries, tissue densities, gamma criteria and dose-volume differences.

Setting: Universitas Hospital (Annex), Bloemfontein, South Africa.

Methods: Deliberate multi-leaf collimator (MLC)-bank offsets were introduced on four different VMAT plans of the prostate, nasopharynx and brain. Measured reference dose sets were compared to measured QA results, using the IBA Dolphin® detector and Compass® software for three dosimetric scenarios. Gamma pass rates over a range of criteria from 1%/2-mm to 4%/4-mm in the total volumes and per structure, as well as dose-volume differences were studied.

Results: Gamma tests in the total patient/phantom did not sufficiently detect errors. The calculation media did not influence the QA outcome greatly. However, the detection geometry affected the results. Per structure gamma analyses provided superior error detection, although still missed some clinically relevant differences. The addition of dose-volume analyses highlighted several important errors.

Conclusion: Volumetric modulated arc therapy using only total volume gamma analyses can easily overlook clinically relevant errors. The choice of gamma criterion is crucial. Verification with at least a per structure gamma test in combination with dose-volume checks is recommended, especially in small target volume cases.

Keywords: QA; VMAT; gamma; DVH; verification.

Introduction

The complexity of volumetric modulated arc therapy (VMAT) requires considerable linear accelerator (linac) component analysis and verification to guarantee the integrity of treatment plan quality and treatment delivery accuracy. The treatment is executed using a non-static arc with dynamic dose rate and aperture variation allowing significant normal tissue dose sparing while maintaining tumour coverage, with the added advantage of a greatly reduced treatment time.^{1,2} Volumetric modulated arc therapy and intensity modulated radiotherapy (IMRT), the precursor to VMAT, offer several advantages over conventional radiotherapy treatment and have been promulgated as a standard of care for a large selection of cancer treatment sites.^{3,4,5,6} However, to maintain treatment integrity, adequate emphasis should be placed on the execution and interpretation of quality assurance (QA) to be aligned with clinical and technical requirements.

Although component malfunctions and deviations from planned treatment positions or values may occur, VMAT delivery with linacs from various manufacturers has been found to be dosimetrically accurate.^{7,8,9} To identify treatment delivery errors, deviations or miscalibrations of linac components, patient-specific QA via dosimetric plan verification and dose recalculation^{10,11,12} generally follows the acceptance of a clinical treatment plan. These can be performed with a wide range of dosimetry systems, ranging from radiation detector arrays^{13,14,15} to linac logfiles.^{16,17}

Read online:

Scan this QR code with your smart phone or mobile device to read online.

Irrespective of the dosimetry system, the analysis philosophy is to dosimetrically compare the measurement or calculation to the original treatment plan. Most systems employ gamma analyses¹⁸ to perform this comparison of recalculated dose in an irradiated phantom, or recalculated in the patient computed tomography (CT) geometry, sometimes utilising each treatment plan structure as a region-of-interest (ROI). Gamma values can be calculated in either 2D (planar) or 3D, which may yield different results.¹⁹ Such results can be influenced by the user's settings, for example the choice of low dose thresholds, dose-difference (DD) and distance-to-agreement (DTA) criteria. Depending on the dosimetry system, other available metrics include 2D dose differences and 3D dose-volume histogram (DVH) comparisons.

Depending on various input settings, inherent features, as well as the data sets, the analysis methods may lead to subjective results. Based on these results, the user must employ pass or fail criteria and action levels to decide if a plan will be acceptable for treatment execution or not. We investigated several analyses metrics, phantom geometries and dosimetry settings often used for VMAT QA to study the dosimetric impact and detectability of known linac component deviations, in particular the multi-leaf collimator (MLC).

Methods

Systematic deviations, such as miscalibration, or random deviations of the MLC from planned positions and movements, lead to errors in patient dose. Although deviations of other linac components also occur, the dosimetric impact of MLC miscalibrations are significantly larger.^{20,21} To evaluate and compare typical treatment verification metrics, complete MLC-bank offsets were introduced into existing VMAT plans to represent a range of MLC miscalibrations. The composite effect of the inherent random linac deviations and the introduced systematic errors thus produced measurable dosimetric differences. Patient-specific verification was subsequently performed with the IBATM Dolphin[®] transmission detector and Compass[®] software package (v2018, IBA Dosimetry GmbH, Schwarzenbruck, Germany).

Planning and patient data

Clinical VMAT plans for three anatomical sites were included in this study: prostate, nasopharynx and brain. Treatment plans were created for the ElektaTM Synergy linac with Agility[®] MLC (80 leaf pairs, 5 mm wide with positioning resolution of 0.1 mm), using the Monaco[®] Treatment Planning System (TPS) (v5.11, Elekta, Stockholm, Sweden).

The prostate and nasopharynx plans consisted of a 10 MV single beam dual-arc rotation, while the two brain-lesion treatment plans consisted of four 6 MV non-coplanar arcs at couch angles of 90°, 45°, 0° and 315°. The prostate plan was a hypo-fractionated treatment with an integrated boost, while the nasopharynx plan was a conventional treatment with a high degree of modulation. The brain plans involved small target volumes with very steep dose gradients. Therefore, the selection of plans covered a range of treatment sites, fractionation schemes and dose gradients. The details, including the target volume of the lesions or planning target volume (PTV), fractionation and total dose, are summarised in Table 1.

Multi-leaf collimator miscalibrations of various magnitudes were introduced into each of the four plans, thus leading to plans with intentional MLC-bank offsets. This was performed using an in-house software package developed in IDL[®] (v8.7, Harris Geospatial Solutions, Inc.). The program reads the initial DICOM RT plan and applies the desired offset to each leaf pair in the open part of the segment (i.e. not the leaf pairs under the Jaws), for each control point. The introduced MLC-bank offsets were applied for both a larger aperture (LA) and smaller aperture (SA), with magnitudes of 1.0 mm, 0.5 mm and 0.2 mm for prostate, nasopharynx and the brain stereotactic radiotherapy (SRT). The brain stereotactic radiosurgery (SRS) plan had only 0.5 mm and 0.2 mm offsets, since larger offsets are unacceptable for SRS treatments. The original clinical treatment plans were used as reference data sets, leading to a total of 26 VMAT plans to be analysed.

Plan verification measurements

Treatment plan verification measurements were performed with the IBATM Dolphin[®] transmission detector, a 2D array of ionisation chambers used for pre-treatment QA or during actual treatment.^{15,22} The Dolphin[®] is attached securely to the linac head and consists of 1513 air-vented plain parallel ionisation chambers covering a full 40 × 40 cm² field size. Each chamber has a diameter of 3.2 mm and a height of 2.0 mm, with a volume of 0.016 cm³. The detectors have a 5.0 mm spacing in the high-resolution central area, extending on the central axes and diagonally, with a lower resolution in the outer section.

Treatment plans, original and modified, were exported in DICOM format to the linac via the Mosaic[®] record-and-verify (R&V) system and measured on the same day in succession. Apart from well-known small random deviations in linac components,²³ this ensured that the differences detected would be from the introduced MLC-bank offsets.

TABLE 1: Plan details of cases investigated

Site	Target volume	Energy	Beam arrangement	Treatment fractions	Total dose
Prostate	176.09 cm ³	10 MV	360° dual arc	20	60 Gy
Nasopharynx	313.09 cm ³	10 MV	360° dual arc	33	70 Gy
Brain SRT	13.24 cm ³	6 MV	Four non-coplanar single arcs	5	35 Gy
Brain SRS	2.08 cm ³	6 MV	Four non-coplanar single arcs	1	18 Gy

SRT, stereotactic radiotherapy; SRS, stereotactic radiosurgery

Measurements and subsequent QA analyses were performed using the Compass© software, which utilises a convolution superposition algorithm for dose recalculation. The Dolphin© samples fluence measurements at a rate of 5 Hz during the treatment delivery. In Compass© these samples represent the fluence for each segment from the treatment plan followed by dose recalculation on the planning CT data set. The DICOM structure set containing the contour data of anatomical structures was available for analyses as well. The dose was calculated on a $2.0 \times 2.0 \times 3.0$ mm grid. All results compare the MLC-bank offsets to the original measured plans, and therefore the effects of TPS/Compass beam models and calculation algorithms were avoided.

Dosimetric scenarios

The measured data was used in three different scenarios to investigate different dosimetric aspects. The first scenario involves dose comparisons in a cylindrical phantom filled with water, denoted as Cyl_w . The phantom had a diameter of 32 cm and a length of 34 cm and was CT-scanned and imported to Compass, with all densities inside the phantom set to water. This scenario only included a ROI for the complete phantom outline. In the second scenario, the dose comparisons were performed on the actual patient planning CT, including all structures, while still overriding all the densities with water (Pt_w). Thus, the actual patient geometry effects and availability of analyses in clinical structures were explored while maintaining the densities as in the phantom. Lastly, the patient geometry was used with the actual electron densities of all media, Pt_m , which is the closest to the real patient's treatment. The measurements were applied separately to each scenario, and the dose subsequently reconstructed. In each scenario, the reconstructed dose from the measurements was compared to the reference dose also recalculated in the same conditions.

Data analyses

The Compass© software was used for all the QA analyses. The software uses a global gamma calculation method in 3D. All plans were evaluated with 12 different gamma criteria settings by using combinations of a DD of 1/2/3/4% and DTA of 2/3/4 mm. Dose normalisation was set to the maximum dose of the reference set, and a low dose exclusion threshold of 10% of the maximum dose was applied throughout for the gamma calculations. Gamma pass rates, which are calculated from the percentage of points in the evaluated region with a gamma value below 1, were determined and reported. Gamma analyses were performed in the total patient/phantom for the Cyl_w , Pt_w and Pt_m scenarios. For both the patient geometry scenarios, gamma analyses were also performed for each target and organ ROI.

Various DVH metric differences for all ROIs were analysed in the patient geometry scenarios. The DVH metrics chosen are taken from our clinical protocol aims and QUANTEC recommendations.²⁴ The DVH metrics that were investigated

are given in Table 2, and include the mean dose (D_{mean}) for all ROIs, minimum dose to 98% of the target volume (D_{98}) as well as the near-maximum dose (D_2). DVH metrics for the organs-at-risk (OARs) include the dose to specific volumes (shown as the volume V to which a dose of x Gy is given, V_x), and the highest dose to 1 cm³ (D_{1cc}).

Action levels and significant dose differences

To evaluate the gamma pass rate, results were grouped to show either above 95% (a pass result), above 88% (marginal failure) or values lower than 88% (failure). This is in accordance with the American Association of Physicists in Medicine (AAPM) Task Group 119 and Palta et al.^{25,26} Setting a specific pass/fail limit on DVH-based differences requires further qualitative interpretation. These should include the radiobiology of fractionation, organ architecture and population averaging effects, as well as clinical endpoints and are discussed elsewhere.²⁷ Although it has been suggested that uncertainties in absorbed dose delivery could be in the order of 3.5%–7.0% depending on the complexity of the treatment,^{28,29} we considered values below 5% to fall within the range of the uncertainties expected and values greater than 5% to be clinically relevant. We present DVH differences in levels of less than 5%, 5%–10% and 10%–20%.

Ethical considerations

Ethical approval was obtained from the University of the Free State (reference number: UFS-HSD2016/1424).

Results

Gamma pass rates for only selected gamma criteria choices of 1%/2-mm, 2%/2-mm and 3%/3-mm are given here for the sake of clarity and are considered to be the most relevant results. The interpretation and discussion pertain to the complete set of findings. The full set, showing all analysed criteria, can be provided by the authors upon request.

TABLE 2: Dose–volume histogram metrics investigated

Variable	Organ	Metric
Prostate	Seminal Vesicles, Prostate	D_{mean} , D_{98}
	PTV 1 (60 Gy)	D_{mean} , D_{98} , D_2
	Bladder, Small Bowel, Rectum	D_{mean} , V_{45} , V_{50} , V_{65} , D_{1cc}
Nasopharynx	PTV 1 (70 Gy)	D_{mean} , D_{98} , D_2
	PTV 2 (59.4 Gy)	D_{mean} , D_{98}
	Brainstem	D_{mean} , V_{54} , V_{60} , D_{1cc}
	Spine	D_{mean} , V_{45} , V_{50} , D_{1cc}
Brain SRT	Nidus	D_{mean} , D_{98}
	PTV (35 Gy)	D_{mean} , D_{98} , D_2
	Brainstem, Optic chiasm, Optic nerve, Lens	D_{mean} , D_{1cc}
	Brain	V_{10} , V_{12} , D_{mean} , D_{1cc}
Brain SRS	Nidus (18 Gy)	D_{mean} , D_{98} , D_2
	Brainstem	D_{mean} , D_{1cc}
	Brain	V_{10} , V_{12} , D_{mean} , D_{1cc}

PTV, planned target volume; SRT, stereotactic radiotherapy; SRS, stereotactic radiosurgery; V_x , volume receiving a dose of x Gy.

Total volume comparison

Gamma pass rates calculated in the total phantom or patient volume for the different dosimetric scenarios are shown in Figure 1, for all MLC-bank offsets introduced. A colour scale is used to indicate the range of pass rates, while the actual pass rate is displayed as percentage numerically. Blue colour represents the pass rate above 95% and indicates an acceptable QA outcome, while red indicates a fail result. Pass rates between 88 and 95% are shown as light blue and could still be acceptable, but mandate further investigation.

Differences in calculated gamma pass rates between the three dosimetric scenarios show a higher sensitivity to dose differences in the cylindrical phantom than the patient geometries. The patient geometry scenarios have very similar results. Gamma criteria less strict than 2%/2-mm

show little sensitivity to the errors introduced here. Pass rates below 95% were only found in the 1.0 mm MLC-bank offset cases. This might be related to the geometric differences between the phantom and the patient. An example of the gamma maps where the pass rate is below 95% is shown in Figure 2 for the nasopharynx case with 1.0 mm MLC-bank offsets for the different dosimetric scenarios and a gamma criterion of 1%/2-mm. The evaluated volumes, that is, volume with dose values above the low dose threshold used for gamma calculation, are also shown below each scenario.

In this case, the evaluated volumes in the phantom scenario were much larger than the patient geometry. The effect of this on the gamma calculation and subsequently the gamma pass rate can be seen in these figures.

		Cyl _w			Pt _w			Pt _m			Cyl _w			Pt _w			Pt _m			Cyl _w			Pt _w			Pt _m			Cyl _w			Pt _w			Pt _m		
		0.2mm SA			0.5mm SA			1.0mm SA			0.2mm LA			0.5mm LA			1.0mm LA																				
Prostate	1%/2mm	100%	100%	100%	99.6%	99.6%	99.6%	94.5%	95.6%	95.7%	99.6%	99.8%	99.8%	97.1%	98%	98.1%	85.1%	88.3%	88.9%																		
	2%/2mm	100%	100%	100%	99.9%	99.9%	99.9%	99.1%	99.2%	99.2%	100%	100%	100%	99.7%	99.7%	99.7%	96.8%	97.3%	99.6%																		
	3%/3mm	100%	100%	100%	100%	100%	100%	99.8%	99.8%	99.8%	100%	100%	100%	100%	100%	100%	99.6%	99.6%	95.7%																		
Nasopharynx	1%/2mm	99.8%	100%	100%	99.2%	99.7%	99.7%	80.4%	91.8%	92.2%	99.9%	100%	100%	96.9%	98%	98%	70.2%	89.6%	90.6%																		
	2%/2mm	100%	100%	100%	100%	100%	100%	96.9%	97.5%	97.3%	100%	100%	100%	99.9%	99.9%	99.9%	93.8%	95.7%	99.6%																		
	3%/3mm	100%	100%	100%	100%	100%	100%	99.8%	99.8%	99.8%	100%	100%	100%	100%	100%	100%	99.7%	99.7%	100%																		
Brain SRT	1%/2mm	100%	100%	100%	100%	100%	100%	99.6%	99.9%	99.9%	100%	100%	100%	100%	100%	100%	99.7%	100%	100%																		
	2%/2mm	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%																		
	3%/3mm	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%																		
Brain SRS	1%/2mm	100%	100%	100%	100%	100%	100%				100%	100%	100%	100%	100%	100%																					
	2%/2mm	100%	100%	100%	100%	100%	100%				100%	100%	100%	100%	100%	100%																					
	3%/3mm	100%	100%	100%	100%	100%	100%				100%	100%	100%	100%	100%	100%																					

SA, smaller aperture; LA, larger aperture.

FIGURE 1: Gamma pass rates in total calculation geometry for each clinical case with all multi-leaf collimator-bank offsets.

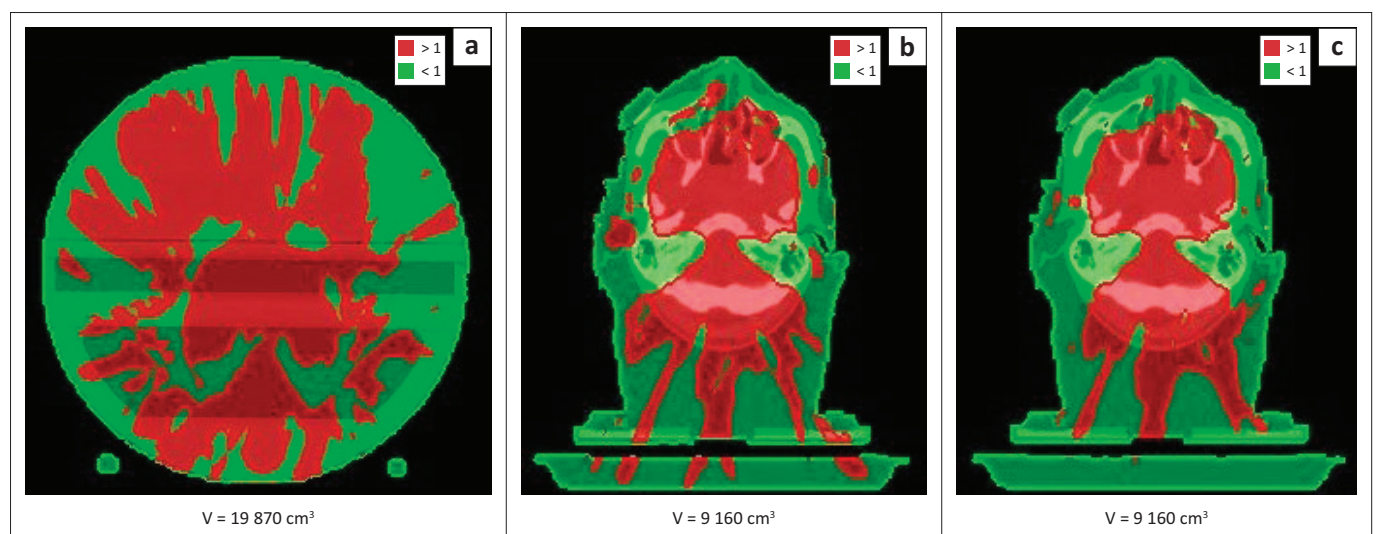


FIGURE 2: Gamma maps and evaluated volumes for nasopharynx case with 1.0 mm larger aperture multi-leaf collimator-bank offset with a 1%/2-mm criterion for (a) Cyl_w, (b) Pt_w and (c) Pt_m scenarios

Region-of-interest comparison

Water versus medium

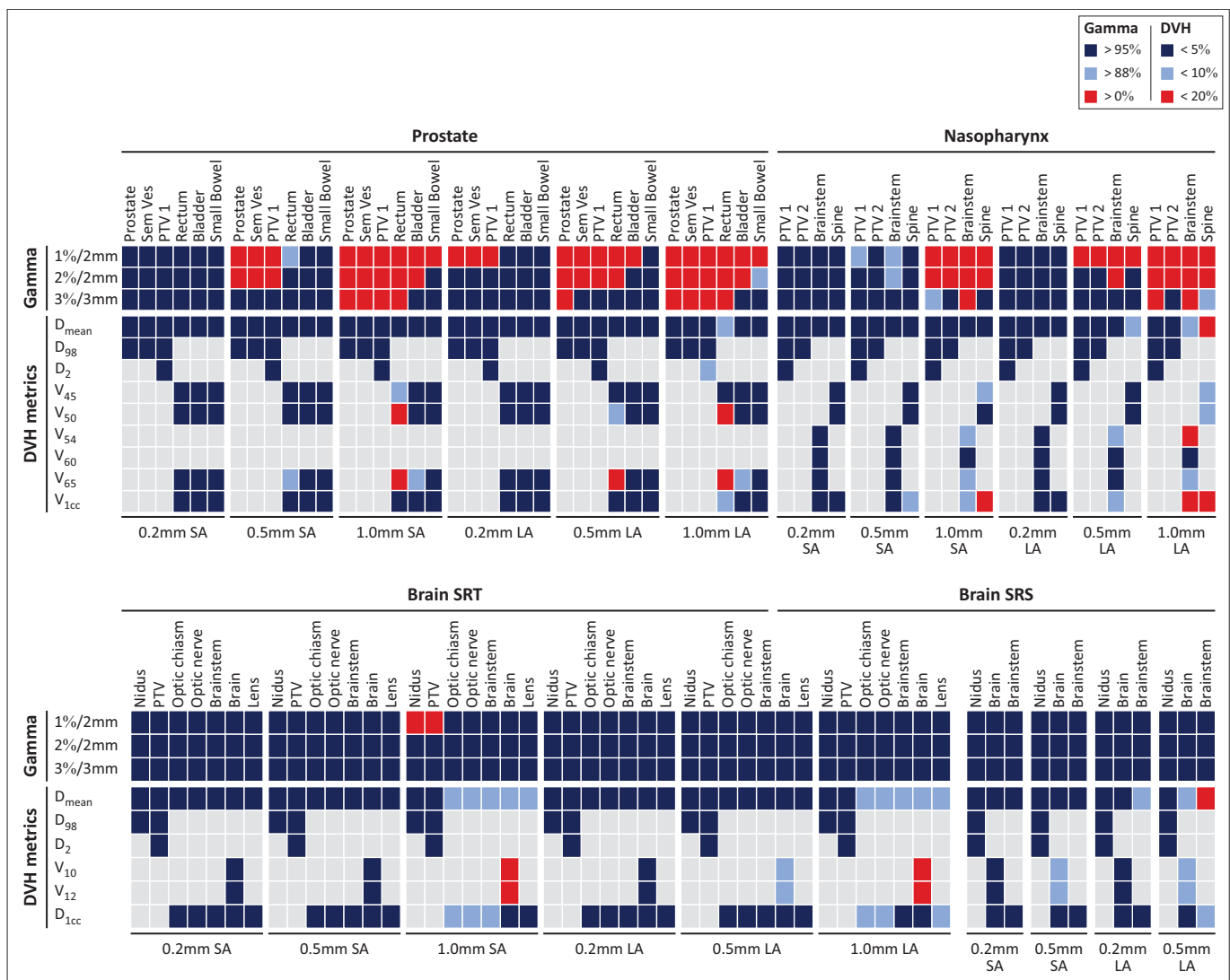
Gamma pass rates calculated for each ROI in the patient geometry with water densities applied (Pt_w) yielded very similar results to the scenario with the actual densities of the medium (Pt_m). The gamma pass rates in the Pt_w scenario were within 3% of Pt_m results in 98% of all evaluated values, with the Pt_w pass rates either equal to or marginally higher than the Pt_m pass rates for the same ROI. This was found even in the nasopharynx case where nasal cavities with very heterogeneous media are found. Only the Pt_m scenario will, therefore, be discussed in further detail, but the same conclusions can be made where densities are set to water.

Pt_m analyses

Region-of-interest gamma pass rates for all the planned sites are shown in Figure 3 for the different MLC-bank offsets introduced, with the DVH differences indicated in the same figure. The gamma pass rate colour scale is similar to the

previous section. The DVH differences are presented with a different scale, but with the same colours as the gamma pass rates, to aid in the interpretation and interrelating of the results. Dose-volume histogram differences below 5% are indicated in blue, differences of 5% – 10% which can become clinically relevant are shown in light blue, and the large differences with values between 10% and 20% are shown in red. Only absolute differences are used; therefore, DVH differences in the LA results are of values lower than the reference, while in the SA results the differences are higher than the reference.

In the prostate and nasopharynx cases, the gamma pass rates per ROI fail in many of the targets and OARs for the 0.5 mm and 1.0 mm MLC-bank offsets for the stricter gamma criteria. The DVH differences are in agreement with this. In the nasopharynx case, where some organ architecture types are serial, larger DVH differences occur than in the prostate plan with organs with a parallel architecture. In the brain cases, gamma pass rates below 95% are only seen in the 1.0 mm SA



Dose-volume histogram, DVH; LA, larger aperture; PTV, planning target volume; SA, smaller aperture; SRS, stereotactic radiosurgery; SRT, stereotactic radiotherapy.

FIGURE 3: Gamma pass rates and dose-volume histogram differences per region-of-interest for all plans investigated, with different colour scales for gamma and dose-volume histogram as indicated in the legend

offset on the SRT plan, while dose differences above 5% are found in both SA and LA offsets, starting at offset values of 0.5 mm. Large DVH differences are seen for the brainstem and brain ROIs, which are not identified with the gamma pass rate results. Gamma tests in both brain cases detect virtually no errors, while DVH data show significant differences at V_{10} and V_{12} .

The DVH of the OAR ROIs and gamma maps for selected MLC-bank offsets in the prostate and brain SRT plans are shown in Figure 4.

The DVHs for the targets and OARs (e.g. rectum and optic chiasm) are shifted by more than 3%, but also differ in slope, whereas differences in the bladder, small bowel and brainstem are very small. Due to the proximity of the optic chiasm and brainstem to the target volume, dose differences above 5% are present but not identified by gamma because of a small number of voxels (small distance) being affected. Gamma maps display virtually no failing points in these volumes.

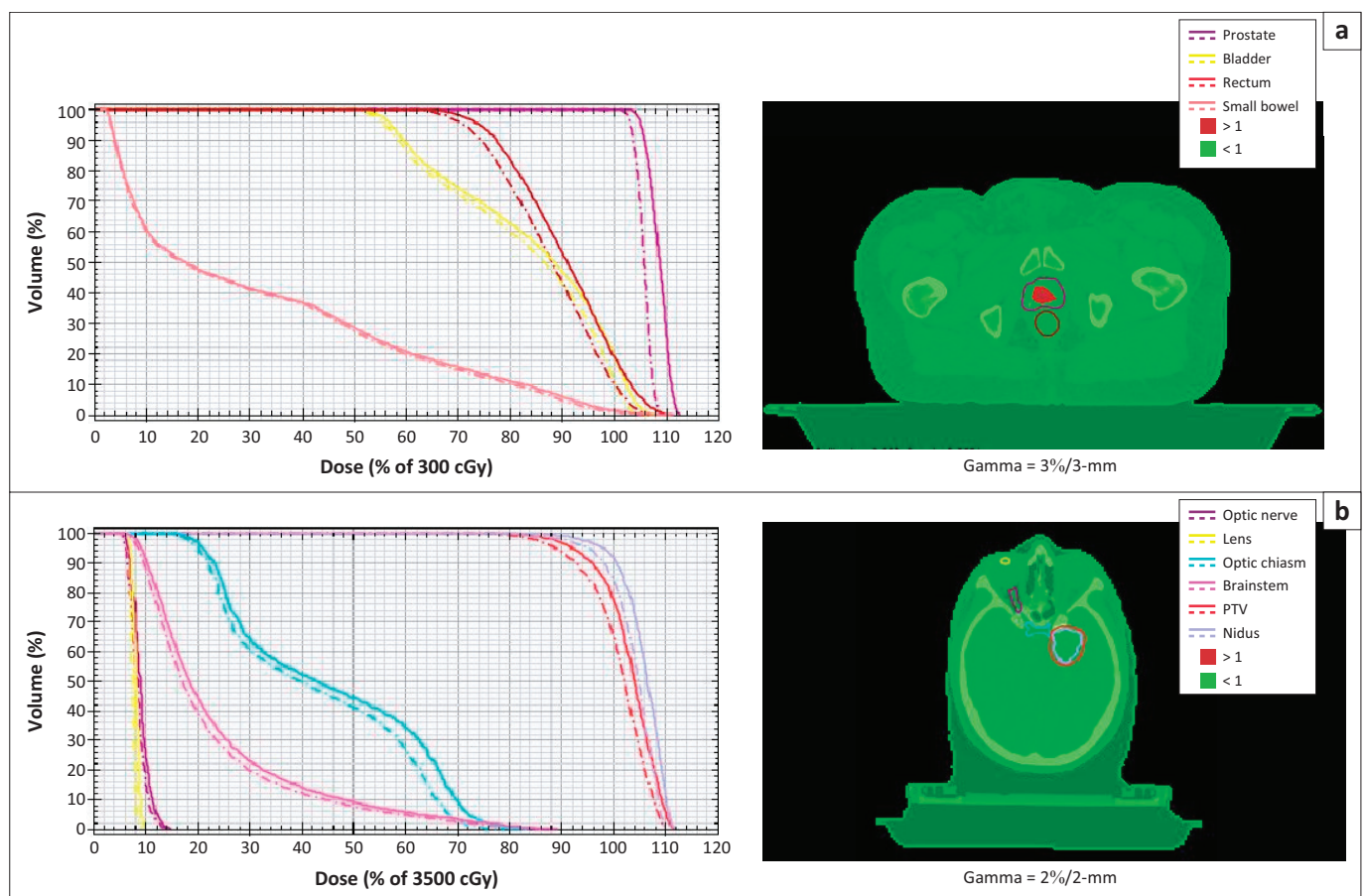
Discussion

Volumetric modulated arc therapy QA results for different treatment sites and MLC-bank offsets were compared to the original dose distributions and dose differences evaluated.

Other types of linac component errors are known to occur; MLC-bank offsets were used in this study to create measurable dosimetric differences. Measured reference dose sets were compared to measured QA results, and thus two recalculated data sets were compared. In effect, any potential dose calculation algorithm bias is eliminated as the same algorithm was used in all cases. This methodology highlights the sensitivity of the evaluation tools. Gamma analyses for the total calculation geometry or per ROI as well as DVH metrics were studied for different dosimetric scenarios.

Only a few treatment sites that have been treated at our clinic were considered and are in general representative of many other similar sites. Therefore, we anticipate that our results are generally applicable elsewhere. The total calculation geometry gamma tests investigated on the treatment sites of this study proved to be insensitive to the introduced errors unless a criterion with at most a DD of 1% was used. A pass rate below 95% was observed only in MLC-bank offsets of 1.0 mm with a gamma criterion stricter than 1%/3-mm, and thus QA results using the gamma test alone on the complete geometries would have incorrectly passed nearly all plans. Recommended gamma criteria are much more lenient than these.

In the brain cases, where the target volumes are very small and extremely high doses may be delivered in a small number



PTV, planning target volume.

FIGURE 4: Dose-volume histogram (left) and gamma maps (right) for the 0.5 mm larger aperture case for (a) prostate and (b) brain stereotactic radiotherapy.

of treatment fractions, the gamma test provided no indication whatsoever of errors even when using the strictest gamma criterion. The use of a total phantom or patient gamma test alone should, therefore, be used with caution. The choice of minimum dose threshold in combination with gamma criterion should be carefully considered to ensure that errors are not overlooked.

In all dosimetric scenarios, the cylindrical phantom gamma test showed a higher sensitivity to the errors than the patient geometries. The irradiated volume over which the gamma was calculated may be the largest contributor to this difference. In the nasopharynx case, the same dose was distributed in a smaller volume in the patient geometry than the phantom (see Figure 2), and therefore the dose distribution is smeared out more in the larger volume. This leads to more voxels considered in the gamma calculation and subsequently a different ratio of pass or fail voxels. The relative volume difference between the patient and phantom geometries may potentially influence the final verification results. Phantoms usually have larger volumes than the actual patient geometry, and therefore phantom-based QA can mask real differences to some extent. Using a stricter distance parameter in the gamma test may be necessary, or higher required pass rates. Nonetheless, using the gamma over a total volume alone is not recommended.

Using the patient geometry with uniform water density may be beneficial when considering the limitations of dose calculation algorithms or models. We could not identify any significant differences in the evaluated results compared to the scenario with actual densities. Therefore, the advantage of using actual densities does not relate to QA analyses, but rather allows determination of the actual patient dose distribution.

The gamma tests performed on the patient geometries where ROIs could be used provided superior error detection ability. A criterion of 2%/2-mm already indicated dosimetric differences with 0.5 mm MLC offsets in some OARs for the prostate and nasopharynx cases. In none of the 0.2 mm MLC offset cases were any differences found, which is expected since these errors are small. Also, none of the introduced offsets led to a significant change in target doses. In the brain cases, however, the gamma again showed virtually no errors for either the 0.2 mm or 0.5 mm MLC-bank offsets. Only the 1.0 mm SA MLC offset case when using a gamma criterion with DD of 1% indicated a pass rate below 95% on the target structures. At the same time, the DVH metrics indicate substantial differences in almost all 1.0 mm MLC-bank offset cases in at least one of the OARs. This is critically important when considering, for example, the V_{10} and V_{12} normal brain volumes. Differences in the order of 5% – 10% were found in some of the 0.5 mm MLC-bank offsets already and should be cause for investigation to determine the magnitude of the clinical effect this would have. An increase in the V_{12} could lead to a significantly increased probability of cyst formation when normal brain tolerance volumes are violated. The detector used here was not ideally suited for SRS, but has

been found in our experience to detect small differences very accurately. However, as the detector was constant for all cases, the results can be consistently interpreted.

The DVH metrics and gamma analyses per ROI agreed very well in the OARs in the prostate case, as well as for the nasopharynx when a gamma criterion of 2%/2-mm or stricter was used. This is similar to the suggestions of Sdrolia et al.³⁰ and Heileman et al.²¹. The contrary, however, can also be said: gamma tests with criterion less strict than 2%/2-mm overlooked errors that are detected in the DVH differences. This can have a major impact in cases where serial organs, such as the spine or brainstem, are very sensitive to maximum dose violations. In the brain cases, DVH differences were seen for structures where the gamma pass rates are above 95%, regardless of the criterion used. Other studies have investigated the relationship between gamma pass rate and DVH differences, and the overall conclusion has been that the correlation is poor.^{31,32,33,34} In our study, the range of gamma criteria evaluated overall as well as per ROI and including DVH parameters provides a comprehensive overview of these QA metrics and their limitations.

There have been many studies on the appropriate acceptance/action levels for gamma analyses. The findings of the AAPM Task Group 119 on IMRT dosimetry, in accordance with the conclusions of Palta et al., suggests an action level at 88% – 90% gamma pass rate for a 3%/3-mm criterion.^{25,26} Nelms and Simon found that many institutions assumed that a 95% pass rate was acceptable.³⁵ A recent study assessed radiobiological metrics used in prostate VMAT QA, in which generalised Equivalent Uniform Dose, Normal Tissue Complication Probability and Tumour Control Probability calculations were performed for the AAPM TG-116 test cases.³⁶ Therein was highlighted that the characteristics of the PTV or OAR as well as its location can have an impact on setting tolerance levels. We strongly support the suggestion by Cozzolino et al.³⁷ that at least the gamma pass rate per structure should be analysed when evaluating VMAT QA results.

The recommended gamma criterion of 3%/3-mm was found in this study to be relatively insensitive to non-negligible dose errors, and therefore proper verification requires the inclusion of supplementary DVH-analyses. Whereas MLC-bank miscalibration, or for that matter any other linac component deviations, may be identified with conventional QA procedures, dosimetric deviations from the intended treatment are not easily identified when using gamma analyses only.

Conclusion

The use of the gamma test to evaluate QA results using only a total geometry comparison should be used with caution and could easily overlook significant dosimetric deviations. Region-of-interest specific gamma comparison provides a more effective solution to finding differences with potential clinical relevance. Treatment sites with small volumes, like in

SRT/SRS, should place a higher emphasis on DVH difference analyses than gamma tests. The choice of appropriate gamma criterion for evaluation is an important consideration and should be carried out carefully. Even a 3%/3-mm gamma criterion is insensitive to dosimetric differences, with potential clinical impact. Dose-volume histogram metric analyses in combination with appropriate ROI gamma tests can greatly improve error detection and should be the standard approach to VMAT plan verification.

Acknowledgements

The authors would like to acknowledge IBA Dosimetry GmbH for providing the Dolphin© detector and Compass© software used in this study.

Competing interests

The authors declare that they have no financial or personal relationships that may have inappropriately influenced them in writing this article.

Authors' contributions

L.S. was responsible for the measurements, data analyses and writing of the manuscript. W.S. was the supervisor for this project, and provided guidance, aided in concept and study design, and was co-writer of the manuscript.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Data availability statement

Data sharing is not applicable to this article as no new data were created in this study.

Disclaimer

The views expressed in this article are those of the authors and not an official position of the institution or funder.

References

- Otto K. Volumetric modulated arc therapy: IMRT in a single gantry arc. *Med Phys* [serial online]. John Wiley & Sons, Ltd; 2007 Dec 26 [cited 2019 Feb 19];35(1):310–317. Available from: <https://doi.wiley.com/10.1118/1.2818738>
- Teoh M, Clark CH, Wood K, Whitaker S, Nisbet A. Volumetric modulated arc therapy: A review of current literature and clinical use in practice. *Br J Radiol* [serial online]. British Institute of Radiology; 2011 Nov [cited 2019 Feb 19];84(1007):967–996. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/22011829>.
- Guadagnolo BA. IMRT should be considered a standard-of-care approach for radiation therapy for soft tissue sarcoma of the extremity. *Ann Surg Oncol* [serial online]. Springer International Publishing; 2019 Jan 31 [cited 2019 Apr 04];26(5):1186–1187. Available from: <https://link.springer.com/10.1245/s10434-019-07192-3>.
- Fischer-Valuck BW, Rao YJ, Michalski JM. Intensity-modulated radiotherapy for prostate cancer. *Transl Androl Urol* [serial online]. AME Publications; 2018 Jun [cited 2019 Apr 04];7(3):297–307. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/30050791>.
- Samuels MA, Freedman LM, Elsayyad N. Intensity-modulated radiotherapy for early glottic cancer: Transition to a new standard of care?. *Futur Oncol* [serial online]. Future Medicine Ltd London, UK; 2016 Nov 09 [cited 2019 Apr 04];12(22):2615–2630. Available from: <https://www.futuremedicine.com/doi/10.2217/fo-2016-0156>.
- Palma D, Vollans E, James K, et al. Volumetric modulated arc therapy for delivery of prostate radiotherapy: Comparison with intensity-modulated radiotherapy and three-dimensional conformal radiotherapy. *Int J Radiat Oncol Biol Phys*. 2008;72(4):996–1001. <https://doi.org/10.1016/j.ijrobp.2008.02.047>
- Haga A, Nakagawa K, Shiraishi K, et al. Quality assurance of volumetric modulated arc therapy using Elekta Synergy. *Acta Oncol (Madr)* [serial online]. Taylor & Francis; 2009 Jan 28 [cited 2019 Feb 26];48(8):1193–1197. Available from: <https://www.tandfonline.com/doi/full/10.3109/02841860903081905>.
- Ling CC, Zhang P, Archambault Y, Bocanek J, Tang G, LoSasso T. Commissioning and quality assurance of rapid arc radiotherapy delivery system. *Int J Radiat Oncol Biol Phys*. 2008;72(2):575–581. <https://doi.org/10.1016/j.ijrobp.2008.05.060>
- Mayo C, Fong de los Santos L, Kruse J, et al. Report on use of a methodology for commissioning and quality assurance of a VMAT system. Chen C-T, editor. *PLoS One* [serial online]. Public Library of Science; 2013 Mar 15 [cited 2019 Feb 26];8(3):e58877. Available from: <https://plos.org/10.1371/journal.pone.0058877>.
- Ezzell GA, Galvin JM, Low D, et al. Guidance document on delivery, treatment planning, and clinical implementation of IMRT: Report of the IMRT subcommittee of the AAPM radiation therapy committee. *Med Phys* [serial online]. John Wiley & Sons, Ltd; 2003 Jul 24 [cited 2019 Apr 09];30(8):2089–2115. Available from: <https://doi.wiley.com/10.1118/1.1591194>
- LoSasso T, Chui C-S, Ling CC. Comprehensive quality assurance for the delivery of intensity modulated radiotherapy with a multileaf collimator used in the dynamic mode. *Med Phys* [serial online]. John Wiley & Sons, Ltd; 2001 Nov 01 [cited 2019 Jan 18];28(11):2209–2219. Available from: <https://doi.wiley.com/10.1118/1.1410123>
- LoSasso T, Chui CSCS, Ling CCC. Physical and dosimetric aspects of a multileaf collimation system used in the dynamic mode for implementing intensity modulated radiotherapy. *Med Phys*. 1998;25(10):1919–1927. <https://doi.org/10.1118/1.598381>
- Masi L, Casamassima F, Doro R, Francescon P. Quality assurance of volumetric modulated arc therapy: Evaluation and comparison of different dosimetric systems. *Med Phys* [serial online]. John Wiley & Sons, Ltd; 2011 Jan 10 [cited 2019 Feb 26];38(2):612–621. Available from: <https://doi.wiley.com/10.1118/1.3533900>
- Arumugam S, Xing A, Young T, Thwaites D, Holloway L. Comparison of three commercial dosimetric systems in detecting clinically significant VMAT delivery errors. *Phys Medica* [serial online]. Associazione Italiana di Fisica Medica; 2016;32(10):1238–1244. Available from: <https://doi.org/10.1016/j.ejmp.2016.09.016>
- Thoelking J, Fleckenstein J, Sekar Y, et al. Patient-specific online dose verification based on transmission detector measurements. *Radiother Oncol* [serial online]. Elsevier Ireland Ltd; 2015;119(2):351–356. Available from: <https://doi.org/10.1016/j.radonc.2016.04.003>
- Luo W, Li J, Price RA, et al. Monte Carlo based IMRT dose verification using MLC log files and R/V outputs. *Med Phys* [serial online]. John Wiley & Sons, Ltd; 2006 Jun 26 [cited 2019 Jan 18];33(7Part1):2557–2564. Available from: <https://doi.wiley.com/10.1118/1.2208916>
- Narayanan G, Zalman T, Ha CS, Papanikolaou N, Stathakis S. Evaluation of Dosimetry Check software for IMRT patient-specific quality assurance. *J Appl Clin Med Phys*. 2015;16(3):329–338. <https://doi.org/10.1120/jacmp.v16i3.5427>
- Low DA, Harms WB, Mutic S, Purdy JA. A technique for the quantitative evaluation of dose distributions. *Med Phys* [serial online]. John Wiley & Sons, Ltd; 1998 May 01 [cited 2019 Jan 18];25(5):656–661. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/9608475>.
- Rajasekaran D, Jeevanandam P, Sukumar P, Ranganathan A, Johnjothi S, Nagarajan V. A study on correlation between 2D and 3D gamma evaluation metrics in patient-specific quality assurance for VMAT. *Med Dosim* [serial online]. Elsevier; 2014;39(4):300–308. Available from: <https://doi.org/10.1016/j.meddos.2014.05.002>
- Oliver M, Gagne I, Bush K, Zavgorodni S, Ansbacher W, Beckham W. Clinical significance of multi-leaf collimator positional errors for volumetric modulated arc therapy. *Radiother Oncol* [serial online]. Elsevier Ireland Ltd; 2010 Dec 01 [cited 2019 Jan 21];97(3):554–560. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/20817291>.
- Heilemann G, Poppe B, Laub W. On the sensitivity of common gamma-index evaluation methods to MLC misalignments in Rapidarc quality assurance. *Med Phys* [serial online]. John Wiley & Sons, Ltd; 2013 Feb 08 [cited 2019 Jan 18];40(3):031702. Available from: <https://doi.wiley.com/10.1118/1.4789580>
- Thoelking J, Sekar Y, Fleckenstein J, Lohr F, Wenz F, Wertz H. Charakterisierung eines neuen Transmissionsdetektors für die patientenindividualisierte Online-Planverifikation und der Einfluss des Detektors auf die Strahlcharakteristik eines 6MV-Röntgentherapiestrahls. *Z Med Phys* [serial online]. Elsevier B.V.; 2016;26(3):200–208. Available from: <https://doi.org/10.1016/j.zemedi.2015.08.001>
- Kaurin DGL, Sweeney LE, Marshall EI. VMAT testing for an Elekta accelerator. *J Appl Clin Med Phys*. 2012;13(2):55–72. <https://doi.org/10.1120/jacmp.v13i2.3725>
- Marks LB, Yorke ED, Jackson A, et al. Use of normal tissue complication probability models in the clinic. *Int J Radiat Oncol* [serial online]. Elsevier; 2010 Mar 01 [cited 2019 Mar 04];76(3):S10–S19. Available from: <https://www.sciencedirect.com/science/article/pii/S036030160903288X?via%3Dihub>.
- Ezzell GA, Burmeister JW, Dogan N, et al. IMRT commissioning: Multiple institution planning and dosimetry comparisons, a report from AAPM Task Group 119. *Med Phys* [serial online]. John Wiley & Sons, Ltd; 2009 Nov 01 [cited 2019 Feb 26];36(11):5359–5373. Available from: <https://aapm.onlinelibrary.wiley.com/doi/full/10.1118/1.3238104>.
- Palta JR, Mackie TR, Chen Z. Intensity-modulated radiation therapy-the state of the art. *Med Phys* [serial online]. John Wiley & Sons, Ltd; 2003 Dec 01 [cited 2019 Mar 01];30(12):3265–3265. Available from: <https://doi.wiley.com/10.1118/1.1628279>
- Marks LB, Yorke ED, Jackson A, et al. Use of normal tissue complication probability models in the clinic. *Int J Radiat Oncol Biol Phys* [serial online]. 2010 Mar [cited 2014 Oct 02];76(3):S10–S19. Available from: <https://www.redjournal.org/article/S036030160903288X/abstract>.

28. Mijnheer BJ, Battermann JJ, Wambersie A. What degree of accuracy is required and can be achieved in photon and neutron therapy? *Radiother Oncol* [serial online]. Elsevier; 1987 Mar 01 [cited 2019 Apr 12];8(3):237–252. Available from: <https://www.sciencedirect.com/science/article/pii/S0167814087802475?via%3Dihub>.
29. Thwaites D. Accuracy required and achievable in radiotherapy dosimetry: Have modern technology and techniques changed our views? *J Phys Conf Ser*. 2013;444(1):012006. <https://doi.org/10.1088/1742-6596/444/1/012006>
30. Sdrolia A, Brownsword KM, Marsden JE, Alty KT, Moore CS, Beavis AW. Retrospective review of locally set tolerances for VMAT prostate patient specific QA using the COMPASS® system. *Phys Medica* [serial online]. Elsevier Ltd; 2015;31(7):792–797. Available from: <https://doi.org/10.1016/j.ejmp.2015.03.017>
31. Stasi M, Bresciani S, Miranti A, Maggio A, Sapino V, Gabriele P. Pretreatment patient-specific IMRT quality assurance: A correlation study between gamma index and patient clinical dose volume histogram. *Med Phys* [serial online]. John Wiley & Sons, Ltd; 2012 Nov 30 [cited 2019 Mar 05];39(12):7626–7634. Available from: <https://doi.wiley.com/10.1118/1.4767763>
32. Zhen H, Nelms BE, Tomé WA. Moving from gamma passing rates to patient DVH-based QA metrics in pretreatment dose QA. *Med Phys* [serial online]. John Wiley & Sons, Ltd; 2011 Sep 19 [cited 2019 Feb 26];38(10):5477–5489. Available from: <https://doi.wiley.com/10.1118/1.3633904>
33. Nelms BE, Zhen H, Tomé WA. Per-beam, planar IMRT QA passing rates do not predict clinically relevant patient dose errors. *Med Phys* [serial online]. John Wiley & Sons, Ltd; 2011 Jan 31 [cited 2019 Feb 26];38(2):1037–1044. Available from: <https://doi.wiley.com/10.1118/1.3544657>
34. Fredh A, Scherman JB, Fog LS, Munck af Rosenschöld P. Patient QA systems for rotational radiation therapy: A comparative experimental study with intentional errors. *Med Phys* [serial online]. John Wiley & Sons, Ltd; 2013 Feb 28 [cited 2019 Mar 05];40(3):031716. Available from: <https://doi.wiley.com/10.1118/1.4788645>
35. Nelms BE, Simon JA. A survey on planar IMRT QA analysis. *J Appl Clin Med Phys* [serial online]. 2007 Jul 17 [cited 2019 Mar 05];8(3):2448. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/17712302>.
36. Clemente-Gutierrez F, Perez-Vara C, Clavo-Herranz MH, Lepez-Carrizosa C, Perez-Regadera J, Ibanez-Villoslada C. Assessment of radiobiological metrics applied to patient-specific QA process of VMAT prostate treatments. *J Appl Clin Med Phys*. 2016;17(2):341–367. <https://doi.org/10.1120/jacmp.v17i2.5783>
37. Cozzolino M, Oliviero C, Califano G, et al. Clinically relevant quality assurance (QA) for prostate RapidArc plans: Gamma maps and DVH-based evaluation. *Phys Medica* [serial online]. Elsevier Ltd; 2014;30(4):462–472. Available from: <https://doi.org/10.1016/j.ejmp.2014.01.003>