



Evaluation of the ability of three commercially available dosimeters to detect systematic delivery errors in step-and-shoot IMRT plans

Alison Gray¹⁻³, Omemh Bawazeer⁴, Sankar Arumugam¹⁻³, Philip Vial^{1-3,5}, Joseph Descallar^{2,3},
David Thwaites⁵, Lois Holloway^{1-3,5,6}

¹Liverpool and Macarthur Cancer Therapy Centres, South Western Sydney Local Health District, Sydney, NSW, Australia

²Ingham Institute for Applied Medical Research, Sydney, NSW, Australia

³South Western Sydney Clinical School, School of Medicine, University of New South Wales, Sydney, NSW, Australia

⁴Physics Department, Umm Al-Qura University, Mecca, Saudi Arabia

⁵Institute of Medical Physics, School of Physics, University of Sydney, Sydney, NSW, Australia

⁶Centre for Medical Radiation Physics, University of Wollongong, Wollongong, NSW, Australia

ABSTRACT

Background: There is limited data on error detectability for step-and-shoot intensity modulated radiotherapy (sIMRT) plans, despite significant work on dynamic methods. However, sIMRT treatments have an ongoing role in clinical practice. This study aimed to evaluate variations in the sensitivity of three patient-specific quality assurance (QA) devices to systematic delivery errors in sIMRT plans.

Materials and methods: Four clinical sIMRT plans (prostate and head and neck) were edited to introduce errors in: Multi-Leaf Collimator (MLC) position (increasing field size, leaf pairs offset (1–3 mm) in opposite directions; and field shift, all leaves offset (1–3 mm) in one direction); collimator rotation (1–3 degrees) and gantry rotation (0.5–2 degrees). The total dose for each plan was measured using an ArcCHECK diode array. Each field, excluding those with gantry offsets, was also measured using an Electronic Portal Imager and a MatriXX Evolution 2D ionisation chamber array. 132 plans (858 fields) were delivered, producing 572 measured dose distributions. Measured doses were compared to calculated doses for the no-error plan using Gamma analysis with 3%/3 mm, 3%/2 mm, and 2%/2 mm criteria (1716 analyses).

Results: Generally, pass rates decreased with increasing errors and/or stricter gamma criteria. Pass rate variations with detector and plan type were also observed. For a 3%/3 mm gamma criteria, none of the devices could reliably detect 1 mm MLC position errors or 1 degree collimator rotation errors.

Conclusions: This work has highlighted the need to adapt QA based on treatment plan type and the need for detector specific assessment criteria to detect clinically significant errors.

Key words: patient specific QA; IMRT; step-and-shoot; error detection

Rep Pract Oncol Radiother 2021;26(5):793–803

Address for correspondence: Alison Gray, Campbelltown Hospital, Macarthur Cancer Therapy Centre, Therry Road, Campbelltown, NSW 2560, Australia; e-mail: alison.gray@health.nsw.gov.au

This article is available in open access under Creative Common Attribution-Non-Commercial-No Derivatives 4.0 International (CC BY-NC-ND 4.0) license, allowing to download articles and share them with others as long as they credit the authors and the publisher, but without permission to change them in any way or use them commercially

Introduction

Intensity modulated radiotherapy (IMRT) and Volumetric Modulated Arc Therapy (VMAT) are used to achieve conformal dose distributions through motion of the Multi-Leaf Collimator (MLC), gantry and collimator. The position of the leaves as well as other delivery parameters, such as gantry and collimator angles, are susceptible to errors in modulated deliveries, so pre-treatment patient specific quality assurance (QA) is recommended [1]. This QA is commonly performed with 2D or 3D dosimeters, with measurements being compared to treatment planning system (TPS) calculations. The current standard for quantitative comparison is the gamma evaluation technique which combines both dose difference and distance to agreement criteria [2].

The American Association of Physicists in Medicine (AAPM) Task Group 218 [3] provided universal tolerance and action limits for IMRT measurement-based verification QA. However, they also recommended using statistical process control methods to determine local action limits and tracking gamma statistics across tumour sites. They discussed the benefits of different types of measurements, recommending the true composite method, or the perpendicular field-by-field method if this is not possible. Measurement-based plan verification is also recommended during the commissioning of an IMRT program [4]. Therefore, confidence in the equipment and procedures used is essential.

While VMAT is routinely used in many clinics, for some scenarios fixed gantry angle treatments with IMRT are preferred, e.g. for non-coplanar treatments to minimise the risk of collisions or for one-sided treatments such as breast, to reduce exit dose to normal tissue. IMRT can be delivered using a static step-and-shoot technique or a dynamic technique. The step-and-shoot IMRT (sIMRT) technique may be used for a variety of reasons. For example, clinics may choose to use static IMRT as a way to control the modulation when low complexity fields are desired, such as in free breathing breast treatments where the impact of patient motion may be unknown. However there is limited data available on the detectability of errors for static IMRT plans, with most research investigating dynamic IMRT or VMAT [5–30].

With respect to sIMRT plans, Sastre-Padro et al. [31] introduced systematic MLC leaf pair errors of 0.5–2 mm (0.25–1 mm per bank) and conducted QA measurements with radiographic film using a 2%/2 mm gamma criteria. They found that the dose difference increased with the error magnitude. However, for the worst case only a 1.5% median dose difference was observed, and the majority of QA measurements still obtained a pass rate over 95%. Létourneau et al. [32] expanded and contracted the MLC segments for a head and neck (H&N) plan by 1 and 2 mm and analysed MapCHECK measurements using a 3%/2 mm gamma criteria. They concluded that MapCHECK could detect systematic MLC position errors of the order of 1 mm with a properly designed set of QA measurements. Wang et al. [21] compared calculated fluence maps from plans with and without introduced MLC errors (± 0.5 , ± 1.0 and ± 1.5 mm) with a 1%/1 mm gamma criteria. They found that the gamma pass rate decreased with increased MLC leaf error, with their results indicating that MLC leaf pair errors of up to at least 1 mm were acceptable. They noted that the impact of MLC errors on planar QA is weak for low magnitude MLC errors but that the sensitivity of gamma pass rate is related to some beam complexity metrics. Yan et al. [33] found that both radiochromic film and the MapCHECK diode array could only detect systematic MLC position errors of around 2 mm and above when a 3%/3 mm criteria were applied. Gueorguiev et al. [34] assessed the ability of three dosimeters to identify errors which were introduced in energy, MLC, jaw position, MU, and collimator and gantry angle. The COMPASS system, which is a gantry mounted 2D array where the measured data is used to compute the dose on a patient CT, identified the most errors. This was followed by the IMRT MatriXX (2D array) using a 3%/3 mm gamma criteria, then the single ion chamber measurement. Shang et al. shifted the MLC leaves on one side of the bank by 1–4 mm and found that the MatriXX was able to detect a 2 mm shift with a 3%/3 mm gamma criteria and a 1 mm shift with a 2%/2 mm criteria [35]. Alaswad et al. introduced 2 mm and 3 mm systematic MLC errors to one plan and analysed measurements on the Seven29 2D array with a 3%/3 mm and 2%/2 mm gamma criteria. However, their focus was to assess the field-by-field versus single-gantry-angle composite methods of

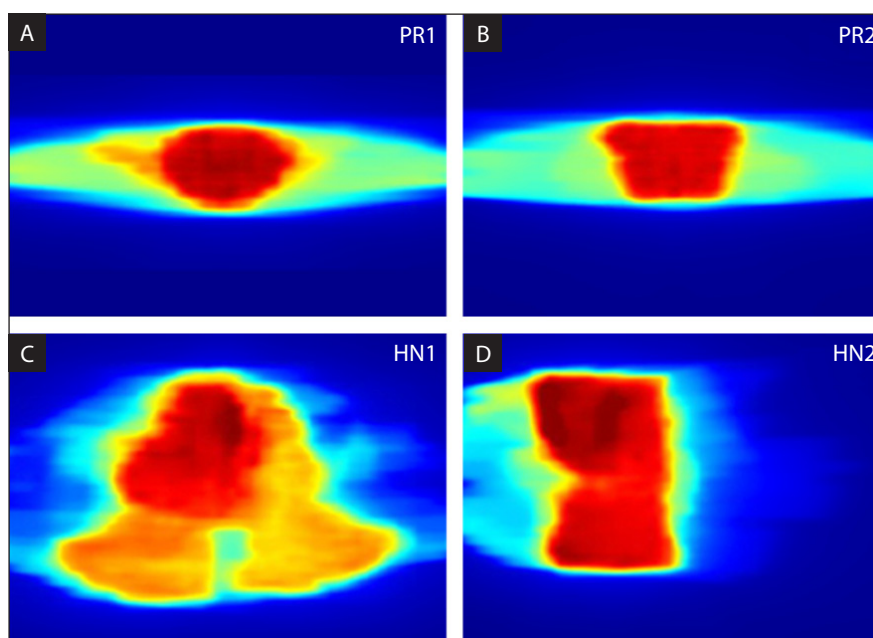


Figure 1. Representative dose planes from the no-error plans calculated in a cubic phantom. **A, B.** Single dose level prostate plans (PR1 & PR2); **C.** Three dose level head and neck plan (HN1); **D.** Two dose level head and neck plan (HN2)

analysis. The field-by-field method of analysis was found to be more sensitive [36].

The aim of this work was to investigate any variations in the ability of three dosimetry systems (ArcCHECK (3D), EPID (2D) and MatriXX Evolution (2D)) to detect systematic delivery errors in step-and-shoot IMRT. This work is part of a larger investigation by the local research group of the impact of uncertainties on advanced radiotherapy treatment techniques [18, 37–45].

Materials and methods

Four clinical 6 MV step-and-shoot IMRT plans, created in the Pinnacle³ (Version 9.0, Philips Healthcare, USA) TPS, were selected. Two of the plans (PR1 & PR2), were prostate treatments with a single high dose planning target volume (PTV). Each plan consisted of seven fields, with 50 segments in total, planned to treat 2 Gy per fraction. The third plan (HN1) was a head and neck plan utilising seven fields and 96 segments. It delivered dose to both sides (bilaterally) with three prescribed dose levels, with the highest dose level of 2 Gy per fraction. The fourth plan (HN2) was also a head and neck plan; however, it utilised five fields with 42 segments. The treatment was to one side of the neck with two prescribed dose levels, with the highest dose level

of 2.1 Gy per fraction. All plans were calculated with a grid size of 0.25 cm × 0.25 cm × 0.25 cm. The treatment machine was a Synergy linear accelerator (Elekta Ltd, UK). Dose planes from the plans can be seen in Figure 1.

Twelve plan variations with errors introduced were created for each plan as described in Table 1. MLC position errors were introduced using an in-house program written in the Python coding language. Two types of MLC position errors were applied to each MLC bank. The first MLC position error applied the same direction leaf pair shift, resulting in the field aperture maintaining the same size but being shifted in the X direction. The second MLC position error applied opposite direction shifts to each leaf in each leaf pair, resulting in a larger leaf gap or irradiated field/segment size. In

Table 1. Description of introduced errors

Introduced error	Magnitude
MLC Position: Increasing Field Size (Opposite direction leaf pair shifts)	1, 2, 3 mm
MLC Position: Field Shift (Same direction leaf pair shift)	1, 2, 3 mm
Collimator Rotation	1, 2, 3 degrees
Gantry Rotation (For ArcCHECK only)	0.5, 1, 2 degrees

both cases, the MLC position errors were applied to all segments and the jaws remained in the planned position. Collimator and gantry rotation errors were introduced manually by copying and editing plans in the record and verify system, MOSAIQ, V2.30.04D1 (Elekta Ltd, UK). This also allowed modifications for the field-by-field measurements, which were conducted with the gantry angle set to zero degrees.

The detectors investigated were the ArcCHECK (Sun Nuclear Corporation, USA), an Electronic Portal Imaging Device (EPID) (Elekta Limited, UK), and the MatriXX Evolution (IBA Dosimetry, Germany). The ArcCHECK is a 3D helical diode array with 10 mm detector spacing; however, the helical design results in offset detectors effectively doubling the detector density in the measurement field. The EPID is a 2D amorphous silicon panel with 0.4 mm detector spacing. The MatriXX evolution is a 2D ionisation chamber array with 7.6 mm detector spacing. The ArcCHECK uses a true composite measurement method, measuring the total plan at the clinical gantry angles. The EPID and MatriXX Evolution use a perpendicular field-by-field measurement method. Measurements for the EPID and MatriXX Evolution were conducted with the gantry at 0°.

Each original plan and plan variation was delivered using a Synergy linear accelerator (Elekta Ltd, UK) to each detector, except for the plans with gantry rotation errors, which were only delivered to the ArcCHECK device (132 plans, 858 fields). Field-by-field analyses were conducted for the EPID and MatriXX Evolution, while a composite (total plan) analysis was conducted for the ArcCHECK, reflecting their normal clinical use.

Repeat measurements for the PR1 no-error plan were also undertaken for each detector, three times on the same day and three times over several weeks, to provide an assessment of reproducibility. Gamma analysis with a 3%/3 mm gamma criteria was undertaken for all repeat measurements and the short and longer term reproducibility values were combined [46] to determine the uncertainty in gamma pass rates.

The measured dose was compared to the calculated dose for the no-error plan. It was assumed that the no-error plan was delivered without any errors. A variety of gamma analysis criteria, 3%/3 mm, 3%/2 mm and 2%/2 mm for dose difference/dis-

tance to agreement, with a 10% dose threshold, were used for the analysis (1716 gamma analyses). For the EPID and MatriXX the average gamma pass rate across all fields was used to assess the given error type.

The analysis was performed in the SNC Patient Software (Sun Nuclear Corporation, USA) for the ArcCHECK and in OmniPro I'mRT V1.4.0.1 (IBA Dosimetry, Germany) for the EPID and MatriXX measurements. To allow import into OmniPro-I'mRT, the EPID image was first converted to dose via in-house software written in Matlab (Version 7.11.0.584 (R20106); The MathWorks, Inc, USA). The conversion to dose was conducted via a pixel calibration factor, which links the measured pixel value of each element detector to the dose in water under reference conditions [47].

A renormalisation was applied to the measured EPID and MatriXX dose distributions and the calculated TPS dose distribution to allow a global percentage difference to be used in the gamma analysis. The maximum dose value in the TPS calculated dose distribution was used as the 100% normalisation point.

A correction was made for the linac output at the time of measurement if its value was greater than 1% from expected. Due to the MatriXX's inherent resolution of 7.6 mm, the grid size of the raw MatriXX data was converted to match that of the TPS calculated data using a cubic spline interpolation. The OmniPro-I'mRT option to shift the measured and planned dose distributions relative to each other prior to the analysis, to allow correction for setup errors or uncertainties, was not used, to avoid obscuring the impact of the introduced errors. The assumption was made that the detector was in all cases aligned correctly so the automatic alignment of the central pixel of the measurement image to the central pixel of the TPS data was used.

A z-test was used to identify if a significant difference in pass rate was observed between the no-error plan and plans with a 1 mm MLC or 1 degree rotation introduced error for the 3%/3 mm gamma analysis. This limit was selected as it relates to the monthly QA tolerances [48].

To assess the sensitivity of the dosimeters to errors, the gradient of the slope produced when the gamma pass rate was plotted against error magnitude. A linear regression test (SPSS statistics software package; IBM Corporation, USA)

Table 2. Gamma pass rates for the no-error plans (3%/3 mm gamma criteria)

	EPID (average for all fields)	MatriXX (average for all fields)	ArcCHECK
PR1	99.34	94.43	98.30
PR2	99.24	96.48	97.50
HN1	91.93	93.30	97.60
HN2	98.43	95.20	99.40

was used to determine if there was a difference in the slope for the different detectors. This test was performed for each error type and plan. Detector, error size (magnitudes of error type) and the interaction between detector and error size were used as variables in each of the multiple linear regression models. A Bonferroni corrected p-value of 0.0167 (0.05/3) was used as the significance level [49]. A linear regression test was also used to assess any differences in error detection sensitivity due to the gamma criteria (3%/3 mm, 3%/2 mm and 2%/2 mm).

Results

The short term reproducibility measurements taken for the PR1 no-error plan, resulted in a mean gamma value (± 1 SD) of $99.5 \pm 0.1\%$, $93.8 \pm 0.1\%$ and $95.8 \pm 0.5\%$ for the EPID, MatriXX and ArcCHECK, respectively. For the long term reproducibility, the results were $99.2 \pm 0.7\%$, $93.9 \pm 1.3\%$ and $97.4 \pm 1.1\%$. These resulted in a combined uncertainty of 0.7% for the EPID, 1.3% for the MatriXX and 1.2% for the ArcCHECK.

The gamma pass rates with a 3%/3 mm gamma criteria for the no-error plan for each detector and plan are presented in Table 2.

The trends in gamma values observed for introduced errors and different gamma criteria were similar across the different plan types. Results for the single level H&N plan (HN2) are presented in Figures 2–5. For the EPID and MatriXX (where a field-by-field analysis was performed), the plotted value represents the average pass rate of all the fields in the plan and the error bars represent \pm one standard deviation. For the ArcCHECK, where a composite (total plan) analysis was conducted, the results plotted are the HN2 plan pass rates.

The results of the z-test to identify which dosimeters can detect errors of 1 mm in MLC position and 1° rotation are presented in Table 3, where a p value of less than 0.05 indicates a significant result. The significant results have been highlighted using bold font.

The results of the linear regression test to assess if the sensitivity to errors varies between detectors are shown in Table 4. The gantry rotation errors are not included as they were only conducted on one

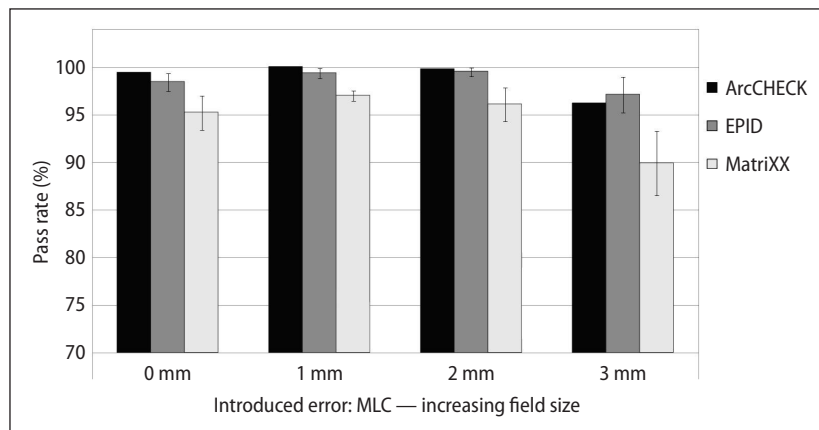


Figure 2. HN2 plan pass rates on ArcCHECK, EPID and MatriXX (3%/3 mm) gamma criteria for an introduced error: MLC — increasing field size. The error bars represent \pm one standard deviation

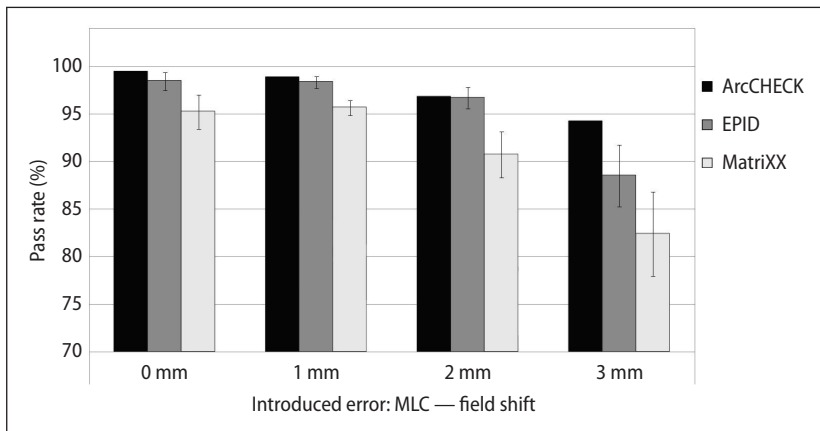


Figure 3. HN2 plan pass rates on ArcCHECK, EPID and MatriXX (3%/3 mm) gamma criteria for an introduced error: MLC — field shift

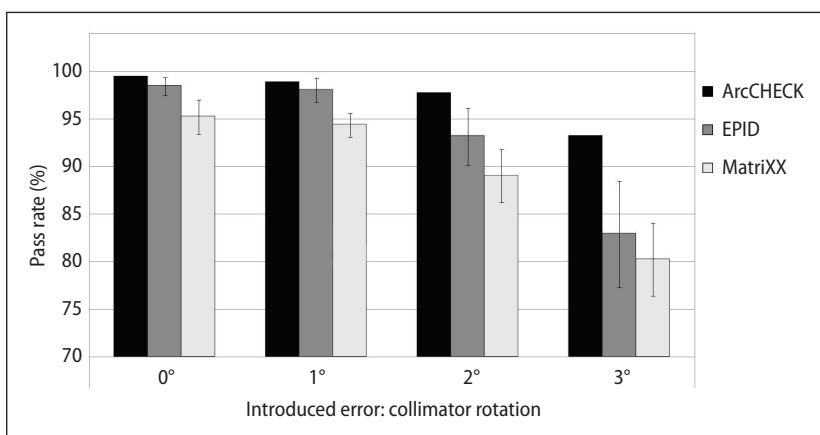


Figure 4. HN2 plan pass rates on ArcCHECK, EPID and MatriXX (3%/3 mm) gamma criteria for an introduced error: collimator rotation

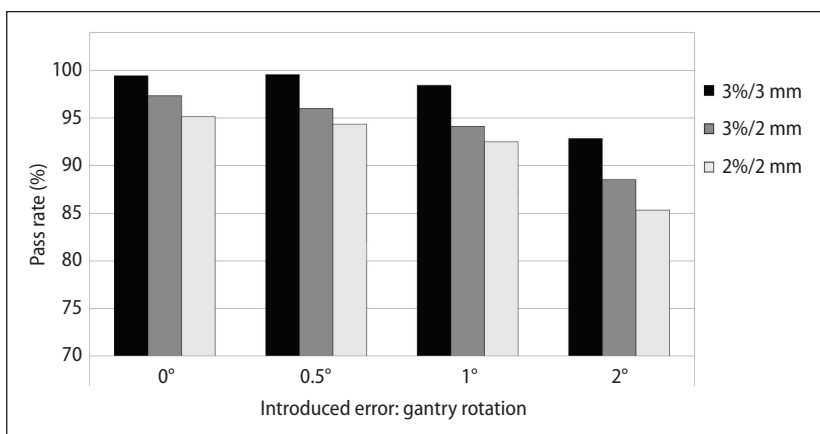


Figure 5. HN2 plan pass rates on ArcCHECK results for 3%/3 mm, 3%/2 mm and 2%/2 mm gamma criteria for an introduced error: gantry rotation

detector. The significant results are highlighted using bold font.

Table 5 illustrates the results of linear regression tests which were conducted for a subset of error

Table 3. The p-values from the z-test to evaluate the ability of detector systems to detect 1 mm MLC or 1° rotation introduced errors, using the 3%/3 mm gamma evaluation criteria

Plan	Error	EPID	MatriXX	ArcCheck
PR1	MLC Shift: Increasing field size	0.823	0.938	0.853
PR1	MLC: Field shift	0.159	0.917	0.098
PR1	Collimator Rotation	0.452	0.334	0.259
PR1	Gantry Rotation			0.000
PR2	MLC Shift: Increasing field size	0.843	0.719	0.947
PR2	MLC: Field shift	0.740	0.299	0.500
PR2	Collimator Rotation	0.618	0.399	0.596
PR2	Gantry Rotation			0.010
HN1	MLC Shift: Increasing field size	1.000	0.999	0.947
HN1	MLC: Field shift	1.000	0.999	0.937
HN1	Collimator Rotation	0.037	0.056	0.000
HN1	Gantry Rotation			0.000
HN2	MLC Shift: Increasing field size	0.918	0.914	0.686
HN2	MLC: Field shift	0.444	0.630	0.314
HN2	Collimator Rotation	0.276	0.266	0.314
HN2	Gantry Rotation			0.210

Table 4. The significance of the difference between the gamma pass rate vs error magnitude slopes of different detectors for each plan for the MLC and collimator errors, based on results of the multiple linear regression test

Plan	Detector	MLC: Increasing Field Size		MLC: Field Shift		Collimator Rotation	
		MatriXX	ArcCHECK	MatriXX	ArcCHECK	MatriXX	ArcCHECK
PR1	EPID	0.878	0.836	0.307	0.524	0.699	0.749
	MatriXX		0.958		0.097		0.480
PR2	EPID	0.710	0.641	0.463	0.600	0.585	0.789
	MatriXX		0.925		0.209		0.780
HN1	EPID	0.723	0.296	0.353	0.039	0.005	0.786
	MatriXX		0.491		0.259		0.010
HN2	EPID	0.366	0.675	0.624	0.003	0.941	0.000
	MatriXX		0.628		0.001		0.000

Table 5. Results (p-values) of the linear regression test evaluating the interaction between different gamma criteria for each error in PR1 and HN1 for the EPID and MatriXX

	EPID			MaxtriXX		
	MLC: Increasing Field Size	MLC: Field Shift	Collimator Rotation Error	MLC: Increasing Field Size	MLC: Field Shift	Collimator Rotation Error
PR1	0.869	0.286	0.442	0.938	0.937	0.830
HN1	0.356	0.756	0.306	0.878	0.232	0.757

types, plans and detectors to determine if a detector becomes more sensitive to errors if the gamma criteria were tightened (3%/3 mm, 3%/2 mm and

2%/2 mm). All of the results produced a p-value greater than the Bonferroni corrected p-value of 0.0167 (0.05/3), indicating that there were no dif-

ferences in the ability of the dosimetry systems to detect errors when the gamma analysis criteria were varied.

Discussion

While the no-error pass rate varied between detectors and plan type, in most cases pass rates decreased with increasing error magnitude and/or stricter gamma criteria as expected (Fig. 2–5). The pass rates for the prostate plans were also slightly higher than for the head and neck plans for the EPID and MatriXX (Tab. 2). This was attributed to the head and neck plans having larger fields with higher modulation and irregular field shapes. The differences between detectors may be attributed to factors such as the inherent variations between the detectors in terms of resolution, detector density, geometry (2D/3D) as well as potential differences in how gamma is calculated in the software [45].

It was assumed for the purpose of this study that the (no-error) plans were delivered accurately; however, in reality all measurements were subject to delivery uncertainties. Based on these results, the choice of acceptable pass rate may need to be reduced to less than 95% for complex plans for some detectors (Tab. 2).

A slight increase in pass rates was observed for the MLC field size (Fig. 2). As the tolerance for MLC position is 1 mm [48] it is possible that there may have been a sub-millimetre systematic calibration error (within clinical QA tolerances) which contributed to this.

It was not expected that a 1 mm error would be identified when a gamma criteria of 3%/3 mm was used. The z-test to identify if the dosimeters could detect errors of 1 mm in MLC position and 1° rotation (equivalent to the linear accelerator monthly QA tolerances) confirmed this, demonstrating that none of the devices could reliably detect systematic errors of this magnitude (Tab. 3). However both EPID and MatriXX were able to detect a 2° collimator error for the more elongated H&N plans (Fig. 4), an error which was found to result in relatively small dose errors, concentrated near the field edges [50]. The ArcCHECK was able to detect a 1° gantry angle error in 3 of the plans but a 2° error was the smallest detectable in the HN2 plan (Fig. 5).

The current results could not be compared to the study by Gueorguiev et al. [34] due to the MLC er-

rors introduced in that study being for a single leaf, and 1-3 leaf pairs rather than the whole bank. The magnitude of the collimator errors (5° and 10°) and gantry angle errors (5° and 15°) also differed substantially in magnitude to the errors introduced in this study. Where MLC bank errors were assessed in other studies, there were variations in the gamma criteria used 3%/3 mm [33], 3%/2 mm [32], 2%/2 mm [31] and 1%/1 mm [21]; however, similar to this study, the results generally indicate that whole bank MLC errors of 1 mm are not identifiable in a clinical treatment plan. For example, Sastre-Padro et al., who investigated MLC leaf errors of up to 1 mm per bank (2 mm per leaf pair), found that a 2%/2 mm gamma criteria only identified a 2 mm leaf pair error in one of the four patient plans assessed. However, they found that the median dose difference was typically 1.0% (1.5% for the worst case). Based on this they suggested that a leaf positioning accuracy of at least 1.0 mm on each leaf bank may be essential for correct IMRT delivery and that this could be achieved with precise calibration and quality assurance procedures [31].

Due to the different detectors used and the different methods for introducing MLC errors not all of the reported research on dynamic IMRT and VMAT could be directly compared to this study. Where a comparison could be made, the results of the current study are generally consistent with published results for both dynamic IMRT and VMAT. Previous work has indicated that 1 mm MLC errors cannot be consistently detected using a 3%/3 mm criteria for dynamic IMRT plans measured with ArcCHECK [7] or EPID [20], or for VMAT plans measured with ArcCHECK [16] or EPID [9, 16]. The ability to detect errors in collimator angle have also been observed to vary with VMAT treatment site, with smaller magnitude errors being detected in H&N plans measured with an EPID [9] and ArcCHECK [24]. Another study observed that collimator errors larger than 2 degrees could be detected in VMAT plans for a variety of treatment sites using the ArcCHECK and EPID [16].

García-Vicente et al. found that the ArcCHECK can detect gantry rotations of 1 degree with a 3%/3 mm analysis criteria for dynamic IMRT [7], which is consistent with what was observed in this work. However, studies investigating the ability of the ArcCHECK to identify gantry angle errors in VMAT plans have resulted in a larger error detec-

tion limit, with a 3 degree error being detectable with a 3%/3 mm criteria in one study [16] and a 2 degree error for 2%/2 mm in another [22]. This may be related to the detectors which are positioned at the periphery of the cylindrical device, measuring dose wash in VMAT plans and fixed entry and exit dose regions in IMRT plans.

The results of the assessment to identify if the sensitivity to errors varies between detectors indicate that there were no consistent differences in the ability of the dosimetry systems to detect introduced errors. However, significant differences were observed for the HN2 plan when a MLC field shift or collimator error was introduced, where the EPID and MatriXX were found to be more sensitive (Tab. 4). This may be related to the elongated shape of the dose distribution for the plan and the fact that the ArcCHECK data was for a composite measurement rather than field-by-field. Previous studies which compared detectors assessed different devices: the COM-PASS system, MatriXX and single ion chamber measurements [34]; and radiochromic film and the MapCHECK diode array [33].

While only a subset of results was analysed with the linear regression test to determine if a detector becomes more sensitive to errors if the gamma criteria are tightened (Tab. 5), based on the similarities of the results across the plans and the large p-values obtained for the current data set, it is not expected that any significant difference would be obtained for the remaining data.

Conclusions

For the detectors assessed in this study on error detectability for sIMRT, the gamma pass rates varied with detector, and generally decreased with increasing errors and/or stricter gamma criteria. The results observed were similar to those in studies on dynamic IMRT and VMAT, with the exception of the magnitude of gantry angle errors being detected by the ArcCHECK for VMAT plans. Based on the analysis, there were no consistent differences in the ability of the three investigated devices to detect systematic delivery errors. The relationship between the gamma pass rate and the dosimetric impact of introduced errors remains unclear and may be affected by the plan complexity, measurement type and the combined dose difference and

distance to agreement gamma analysis method. This work has highlighted the need to establish assessment criteria that can detect clinically significant errors for a given detector system. It has also highlighted the potential need to adapt QA systems, based on particular treatment plan types. The use of process-based tolerances and action limits, as recommended in the AAPM Task Group 218 report [3] may facilitate this.

Conflict of interest

None declared.

Funding

The Saudi Arabian Cultural Mission funded Omemh Bawazeer's work on this project, in association with her studies in Australia.

References

1. Mu G, Ludlum E, Xia P. Impact of MLC leaf position errors on simple and complex IMRT plans for head and neck cancer. *Phys Med Biol*. 2008; 53(1): 77–88, doi: [10.1088/0031-9155/53/1/005](https://doi.org/10.1088/0031-9155/53/1/005), indexed in Pubmed: [18182688](https://pubmed.ncbi.nlm.nih.gov/18182688/).
2. Low DA, Moran JM, Dempsey JF, et al. Dosimetry tools and techniques for IMRT. *Med Phys*. 2011; 38(3): 1313–1338, doi: [10.1118/1.3514120](https://doi.org/10.1118/1.3514120), indexed in Pubmed: [21520843](https://pubmed.ncbi.nlm.nih.gov/21520843/).
3. Miften M, Olch A, Mihailidis D, et al. Tolerance limits and methodologies for IMRT measurement-based verification QA: Recommendations of AAPM Task Group No. 218. *Med Phys*. 2018; 45(4): e53–e83, doi: [10.1002/mp.12810](https://doi.org/10.1002/mp.12810), indexed in Pubmed: [29443390](https://pubmed.ncbi.nlm.nih.gov/29443390/).
4. 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*. 2009; 36(11): 5359–5373, doi: [10.1118/1.3238104](https://doi.org/10.1118/1.3238104), indexed in Pubmed: [19994544](https://pubmed.ncbi.nlm.nih.gov/19994544/).
5. Masi L, Casamassima F, Doro R, et al. Quality assurance of volumetric modulated arc therapy: evaluation and comparison of different dosimetric systems. *Med Phys*. 2011; 38(2): 612–621, doi: [10.1118/1.3533900](https://doi.org/10.1118/1.3533900), indexed in Pubmed: [21452698](https://pubmed.ncbi.nlm.nih.gov/21452698/).
6. Bailey DW, Kumaraswamy L, Bakhtiari M, et al. EPID dosimetry for pretreatment quality assurance with two commercial systems. *J Appl Clin Med Phys*. 2012; 13(4): 3736, doi: [10.1120/jacmp.v13i4.3736](https://doi.org/10.1120/jacmp.v13i4.3736), indexed in Pubmed: [22766944](https://pubmed.ncbi.nlm.nih.gov/22766944/).
7. García-Vicente F, Fernández V, Bermúdez R, et al. Sensitivity of a helical diode array device to delivery errors in IMRT treatment and establishment of tolerance level for pretreatment QA. *J Appl Clin Med Phys*. 2012; 13(1): 3660, doi: [10.1120/jacmp.v13i1.3660](https://doi.org/10.1120/jacmp.v13i1.3660), indexed in Pubmed: [22231218](https://pubmed.ncbi.nlm.nih.gov/22231218/).
8. Heilemann G, Poppe B, Laub W. On the sensitivity of common gamma-index evaluation methods to MLC misalignments in Rapidarc quality assurance. *Med Phys*. 2013; 40(3): 031702, doi: [10.1118/1.4789580](https://doi.org/10.1118/1.4789580), indexed in Pubmed: [23464297](https://pubmed.ncbi.nlm.nih.gov/23464297/).

9. Fredh A, Scherman JB, Fog LS, et al. Patient QA systems for rotational radiation therapy: a comparative experimental study with intentional errors. *Med Phys*. 2013; 40(3): 031716, doi: [10.1118/1.4788645](https://doi.org/10.1118/1.4788645), indexed in Pubmed: [23464311](https://pubmed.ncbi.nlm.nih.gov/23464311/).
10. Carlone M, Cruje C, Rangel A, et al. Poster - Thur Eve - 21: ROC analysis in patient specific quality assurance. *Med Phys*. 2012; 39(7Part3): 4628, doi: [10.1118/1.4740129](https://doi.org/10.1118/1.4740129), indexed in Pubmed: [28516687](https://pubmed.ncbi.nlm.nih.gov/28516687/).
11. Li G, Bai S, Chen N, et al. Evaluation of the sensitivity of two 3D diode array dosimetry systems to setup error for quality assurance (QA) of volumetric-modulated arc therapy (VMAT). *J Appl Clin Med Phys*. 2013; 14(5): 13–24, doi: [10.1120/jacmp.v14i5.3828](https://doi.org/10.1120/jacmp.v14i5.3828), indexed in Pubmed: [24036856](https://pubmed.ncbi.nlm.nih.gov/24036856/).
12. Hussein M, Rowshanfarzad P, Ebert MA, et al. A comparison of the gamma index analysis in various commercial IMRT/VMAT QA systems. *Radiother Oncol*. 2013; 109(3): 370–376, doi: [10.1016/j.radonc.2013.08.048](https://doi.org/10.1016/j.radonc.2013.08.048), indexed in Pubmed: [24100148](https://pubmed.ncbi.nlm.nih.gov/24100148/).
13. Hussein M, Adams EJ, Jordan TJ, et al. A critical evaluation of the PTW 2D-ARRAY seven29 and OCTAVIUS II phantom for IMRT and VMAT verification. *J Appl Clin Med Phys*. 2013; 14(6): 4460, doi: [10.1120/jacmp.v14i6.4460](https://doi.org/10.1120/jacmp.v14i6.4460), indexed in Pubmed: [24257288](https://pubmed.ncbi.nlm.nih.gov/24257288/).
14. Coleman L, Skourou C. Sensitivity of volumetric modulated arc therapy patient specific QA results to multileaf collimator errors and correlation to dose volume histogram based metrics. *Med Phys*. 2013; 40(11): 111715, doi: [10.1118/1.4824433](https://doi.org/10.1118/1.4824433), indexed in Pubmed: [24320423](https://pubmed.ncbi.nlm.nih.gov/24320423/).
15. Kim JI, Park SY, Kim HJ, et al. The sensitivity of gamma-index method to the positioning errors of high-definition MLC in patient-specific VMAT QA for SBRT. *Radiat Oncol*. 2014; 9: 167, doi: [10.1186/1748-717X-9-167](https://doi.org/10.1186/1748-717X-9-167), indexed in Pubmed: [25070065](https://pubmed.ncbi.nlm.nih.gov/25070065/).
16. Vieilleveigne L, Molinier J, Brun T, et al. Gamma index comparison of three VMAT QA systems and evaluation of their sensitivity to delivery errors. *Phys Med*. 2015; 31(7): 720–725, doi: [10.1016/j.ejmp.2015.05.016](https://doi.org/10.1016/j.ejmp.2015.05.016), indexed in Pubmed: [26095758](https://pubmed.ncbi.nlm.nih.gov/26095758/).
17. Kadoya N, Saito M, Ogasawara M, et al. Evaluation of patient DVH-based QA metrics for prostate VMAT: correlation between accuracy of estimated 3D patient dose and magnitude of MLC misalignment. *J Appl Clin Med Phys*. 2015; 16(3): 5251, doi: [10.1120/jacmp.v16i3.5251](https://doi.org/10.1120/jacmp.v16i3.5251), indexed in Pubmed: [26103486](https://pubmed.ncbi.nlm.nih.gov/26103486/).
18. Arumugam S, Xing A, Young T, et al. Sensitivity of a helical diode array dosimeter to Volumetric Modulated Arc Therapy delivery errors. *Phys Med*. 2015; 31(8): 1043–1054, doi: [10.1016/j.ejmp.2015.08.011](https://doi.org/10.1016/j.ejmp.2015.08.011), indexed in Pubmed: [26429384](https://pubmed.ncbi.nlm.nih.gov/26429384/).
19. Nithiyantham K, Mani GK, Subramani V, et al. Analysis of direct clinical consequences of MLC positional errors in volumetric-modulated arc therapy using 3D dosimetry system. *J Appl Clin Med Phys*. 2015; 16(5): 296–305, doi: [10.1120/jacmp.v16i5.5515](https://doi.org/10.1120/jacmp.v16i5.5515), indexed in Pubmed: [26699311](https://pubmed.ncbi.nlm.nih.gov/26699311/).
20. Lat SZ, Suriyapee S, Sanghangthum T. Sensitivity in error detection of patient specific QA tools for IMRT plans. *J Phys Conf Ser*. 2016; 694: 012008, doi: [10.1088/1742-6596/694/1/012008](https://doi.org/10.1088/1742-6596/694/1/012008).
21. Wang J, Jin X, Peng J, et al. Are simple IMRT beams more robust against MLC error? Exploring the impact of MLC errors on planar quality assurance and plan quality for different complexity beams. *J Appl Clin Med Phys*. 2016; 17(3): 147–157, doi: [10.1120/jacmp.v17i3.6022](https://doi.org/10.1120/jacmp.v17i3.6022), indexed in Pubmed: [27167272](https://pubmed.ncbi.nlm.nih.gov/27167272/).
22. Liang B, Liu Bo, Zhou F, et al. Comparisons of volumetric modulated arc therapy (VMAT) quality assurance (QA) systems: sensitivity analysis to machine errors. *Radiat Oncol*. 2016; 11(1): 146, doi: [10.1186/s13014-016-0725-4](https://doi.org/10.1186/s13014-016-0725-4), indexed in Pubmed: [27821135](https://pubmed.ncbi.nlm.nih.gov/27821135/).
23. Defoor DL, Stathakis S, Roring JE, et al. Investigation of error detection capabilities of phantom, EPID and MLC log file based IMRT QA methods. *J Appl Clin Med Phys*. 2017; 18(4): 172–179, doi: [10.1002/acm2.12114](https://doi.org/10.1002/acm2.12114), indexed in Pubmed: [28585300](https://pubmed.ncbi.nlm.nih.gov/28585300/).
24. Au IW, Ciurlionis L, Campbell N, et al. Validation of the Mobius system for patient-specific quality assurance using introduced intentional errors. *Australas Phys Eng Sci Med*. 2017; 40(1): 181–189, doi: [10.1007/s13246-016-0503-5](https://doi.org/10.1007/s13246-016-0503-5), indexed in Pubmed: [27943009](https://pubmed.ncbi.nlm.nih.gov/27943009/).
25. Woon W, Ravindran PB, Ekayanake P, et al. A study on the effect of detector resolution on gamma index passing rate for VMAT and IMRT QA. *J Appl Clin Med Phys*. 2018; 19(2): 230–248, doi: [10.1002/acm2.12285](https://doi.org/10.1002/acm2.12285), indexed in Pubmed: [29460465](https://pubmed.ncbi.nlm.nih.gov/29460465/).
26. Thongsawad S, Srisatit S, Fuangrod T. Error detection sensitivity test using complex errors on three patient-specific VMAT QA systems. *J Phys Conf Ser*. 2019; 1285: 012030, doi: [10.1088/1742-6596/1285/1/012030](https://doi.org/10.1088/1742-6596/1285/1/012030).
27. Gay SS, Netherton TJ, Cardenas CE, et al. Dosimetric impact and detectability of multi-leaf collimator positioning errors on Varian Halcyon. *J Appl Clin Med Phys*. 2019; 20(8): 47–55, doi: [10.1002/acm2.12677](https://doi.org/10.1002/acm2.12677), indexed in Pubmed: [31294923](https://pubmed.ncbi.nlm.nih.gov/31294923/).
28. Sanghangthum T, Lat S, Suriyapee S. Investigation of Error Detection Capabilities of Various Patient-Specific Intensity Modulated Radiotherapy Quality Assurance Devices. *Int J Med Phys Clin Eng Radiat Oncol*. 2019; 08(01): 21–31, doi: [10.4236/ijmpcero.2019.81003](https://doi.org/10.4236/ijmpcero.2019.81003).
29. Agarwal A, Rastogi N, Das KJM, et al. Evaluating the dosimetric consequences of MLC leaf positioning errors in dynamic IMRT treatments. *J Radiother Pract*. 2019; 18(03): 225–231, doi: [10.1017/s1460396918000705](https://doi.org/10.1017/s1460396918000705).
30. Ghafarian M, Price M, Morales-Paliza M. Comparison of pretreatment VMAT quality assurance with the integral quality monitor (IQM) and electronic portal imaging device (EPID). *J Appl Clin Med Phys*. 2021; 22(3): 166–175, doi: [10.1002/acm2.13201](https://doi.org/10.1002/acm2.13201), indexed in Pubmed: [33596336](https://pubmed.ncbi.nlm.nih.gov/33596336/).
31. Sastre-Padro M, Welleweerd J, Malinen E, et al. Consequences of leaf calibration errors on IMRT delivery. *Phys Med Biol*. 2007; 52(4): 1147–1156, doi: [10.1088/0031-9155/52/4/019](https://doi.org/10.1088/0031-9155/52/4/019), indexed in Pubmed: [17264376](https://pubmed.ncbi.nlm.nih.gov/17264376/).
32. Létourneau D, Gulam M, Yan Di, et al. Evaluation of a 2D diode array for IMRT quality assurance. *Radiother Oncol*. 2004; 70(2): 199–206, doi: [10.1016/j.radonc.2003.10.014](https://doi.org/10.1016/j.radonc.2003.10.014), indexed in Pubmed: [15028408](https://pubmed.ncbi.nlm.nih.gov/15028408/).
33. Yan G, Liu C, Simon TA, et al. On the sensitivity of patient-specific IMRT QA to MLC positioning errors. *J Appl Clin Med Phys*. 2009; 10(1): 120–128, doi: [10.1120/jacmp.v10i1.2915](https://doi.org/10.1120/jacmp.v10i1.2915), indexed in Pubmed: [19223841](https://pubmed.ncbi.nlm.nih.gov/19223841/).
34. Gueorguiev G, Cotter C, Turcotte JC, et al. Clinical implementation and error sensitivity of a 3D quality assurance protocol for prostate and thoracic IMRT. *J Appl Clin*

- Med Phys. 2015; 16(5): 179–192, doi: [10.1120/jacmp.v16i5.5392](https://doi.org/10.1120/jacmp.v16i5.5392), indexed in Pubmed: [26699299](https://pubmed.ncbi.nlm.nih.gov/26699299/).
35. Shang Q, Godley A, Huang L, et al. Sensitivity of array detector measurements in determining shifts of MLC leaf positions. *J Appl Clin Med Phys*. 2017; 18(5): 80–88, doi: [10.1002/acm2.12148](https://doi.org/10.1002/acm2.12148), indexed in Pubmed: [28799273](https://pubmed.ncbi.nlm.nih.gov/28799273/).
 36. Alaswad M, Coleman L. The sensitivity of patient-specific IMRT QA methods in detecting systematic errors: field-by-field versus single-gantry-angle composite. *J Phys Conf Ser*. 2019; 1248: 012063, doi: [10.1088/1742-6596/1248/1/012063](https://doi.org/10.1088/1742-6596/1248/1/012063).
 37. Bawazeer O, Gray A, Arumugam S, et al. Evaluation of the ability of a 2D ionisation chamber array and an EPID to detect systematic delivery errors in IMRT plans. *J Phys Conf Ser*. 2014; 489: 012071, doi: [10.1088/1742-6596/489/1/012071](https://doi.org/10.1088/1742-6596/489/1/012071).
 38. Arumugam S, Xing A, Young T, et al. Comparison of three commercial dosimetric systems in detecting clinically significant VMAT delivery errors. *Phys Med*. 2016; 32(10): 1238–1244, doi: [10.1016/j.ejmp.2016.09.016](https://doi.org/10.1016/j.ejmp.2016.09.016), indexed in Pubmed: [27717741](https://pubmed.ncbi.nlm.nih.gov/27717741/).
 39. Deshpande S, Geurts M, Vial P, et al. Clinical significance of treatment delivery errors for helical TomoTherapy nasopharyngeal plans - A dosimetric simulation study. *Phys Med*. 2017; 33: 159–169, doi: [10.1016/j.ejmp.2017.01.006](https://doi.org/10.1016/j.ejmp.2017.01.006), indexed in Pubmed: [28110824](https://pubmed.ncbi.nlm.nih.gov/28110824/).
 40. Blake SJ, Arumugam S, Holloway L, et al. Investigating the impact of treatment delivery uncertainties on treatment effectiveness for lung SABR. *Australas Phys Eng Sci Med*. 2017; 40(4): 823–829, doi: [10.1007/s13246-017-0591-x](https://doi.org/10.1007/s13246-017-0591-x), indexed in Pubmed: [29086179](https://pubmed.ncbi.nlm.nih.gov/29086179/).
 41. Pogson EM, Aruguman S, Hansen CR, et al. Multi-institutional comparison of simulated treatment delivery errors in ssIMRT, manually planned VMAT and autoplan-VMAT plans for nasopharyngeal radiotherapy. *Phys Med*. 2017; 42: 55–66, doi: [10.1016/j.ejmp.2017.08.008](https://doi.org/10.1016/j.ejmp.2017.08.008), indexed in Pubmed: [29173921](https://pubmed.ncbi.nlm.nih.gov/29173921/).
 42. Alharthi T, Pogson EM, Arumugam S, et al. Pre-treatment verification of lung SBRT VMAT plans with delivery errors: Toward a better understanding of the gamma index analysis. *Phys Med*. 2018; 49: 119–128, doi: [10.1016/j.ejmp.2018.04.005](https://doi.org/10.1016/j.ejmp.2018.04.005), indexed in Pubmed: [29685425](https://pubmed.ncbi.nlm.nih.gov/29685425/).
 43. Pogson EM, Arumugam S, Hansen CR, et al. Comparison of multi-institutional pre-treatment verification for VMAT of nasopharynx with delivery errors. *Phys Med*. 2018; 53: 25–31, doi: [10.1016/j.ejmp.2018.07.007](https://doi.org/10.1016/j.ejmp.2018.07.007), indexed in Pubmed: [30241751](https://pubmed.ncbi.nlm.nih.gov/30241751/).
 44. Alharthi T, Arumugam S, Vial P, et al. Pre-treatment verification of lung SBRT VMAT plans with delivery errors: Toward a better understanding of the gamma index analysis. *Phys Med*. 2018; 49: 119–128, doi: [10.1016/j.ejmp.2018.04.005](https://doi.org/10.1016/j.ejmp.2018.04.005), indexed in Pubmed: [29685425](https://pubmed.ncbi.nlm.nih.gov/29685425/).
 45. Alharthi T, Vial P, Holloway L, et al. Intrinsic detector sensitivity analysis as a tool to characterize ArcCHECK and EPID sensitivity to variations in delivery for lung SBRT VMAT plans. *J Appl Clin Med Phys*. 2021; 22(6): 229–240, doi: [10.1002/acm2.13221](https://doi.org/10.1002/acm2.13221), indexed in Pubmed: [33949087](https://pubmed.ncbi.nlm.nih.gov/33949087/).
 46. Gregory K, Bibbo G, Pattison JE. A standard approach to measurement uncertainties for scientists and engineers in medicine. *Australas Phys Eng Sci Med*. 2005; 28(2): 131–139, doi: [10.1007/BF03178705](https://doi.org/10.1007/BF03178705), indexed in Pubmed: [16060321](https://pubmed.ncbi.nlm.nih.gov/16060321/).
 47. Xing A, Arumugam S, Deshpande S, et al. Streamlining EPID-based IMRT quality assurance: auto-analysis and auto-report generation. *J Phys Conf Ser*. 2014; 489: 012084, doi: [10.1088/1742-6596/489/1/012084](https://doi.org/10.1088/1742-6596/489/1/012084).
 48. Hanley J, Dresser S, Simon W, et al. Task Group 142, American Association of Physicists in Medicine. Task Group 142 report: quality assurance of medical accelerators. *Med Phys*. 2009; 36(9): 4197–4212, doi: [10.1118/1.3190392](https://doi.org/10.1118/1.3190392), indexed in Pubmed: [19810494](https://pubmed.ncbi.nlm.nih.gov/19810494/).
 49. Bonferroni Method. <http://www.ma.utexas.edu/users/mks/384E07/multcomp>. (7th October 2019).
 50. Low DA, Zhu XR, Purdy JA, et al. The influence of angular misalignment on fixed-portal intensity modulated radiation therapy. *Med Phys*. 1997; 24(7): 1123–1139, doi: [10.1118/1.598015](https://doi.org/10.1118/1.598015), indexed in Pubmed: [9243475](https://pubmed.ncbi.nlm.nih.gov/9243475/).