Tolerance levels of EPID-based quality control for volumetric modulated arc therapy[a)]

M. K. Jørgensen,[b)] L. Hoffmann, J. B. B. Petersen, L. H. Præstegaard, and R. Hansen
Department of Medical Physics, Aarhus University Hospital, Nørrebrogade 44, Building 5, DK-8000 Aarhus C, Denmark

L. P. Muren
Department of Medical Physics, Aarhus University Hospital, Nørrebrogade 44, Building 5, DK-8000 Aarhus C, Denmark and Department of Oncology, Aarhus University Hospital, Aarhus, Nørrebrogade 44, 8000 Aarhus C, Denmark

(Received 30 June 2010; revised 13 January 2011; accepted for publication 17 January 2011; published 18 February 2011)

Purpose: Volumetric modulated arc therapy (VMAT) includes features such as a variable dose rate and gantry speed in addition to the beam modulation achieved with multileaf collimator (MLC) motion patterns employed in intensity modulated radiotherapy. Three tests have previously been proposed for the evaluation of the performance of VMAT delivery. In order to enable a convenient and accurate routine machine quality control (QC) program, the present study proposes tolerance levels for these tests based on a department-wide implementation of an electronic portal imaging device (EPID)-based QC.

Methods: Three different VMAT tests—a picket fence (PF) test, a dose rate versus gantry speed (DRGS) test, and a dose rate versus MLC leaf speed (DRMLC) test—were performed on nine accelerators using two different EPIDs (aS1000 and aS500, Varian Medical Systems). All tests were repeated six times for each accelerator. The images were analyzed using an in-house-developed software. For the PF test, the positions and widths of individual MLC leaf gaps were compared to the mean value. In the DRGS and DRMLC tests, different combinations of dose rate, gantry speed, and MLC leaf speed were used to deliver identical doses to separate parts of the EPID. The tests were evaluated by looking for deviations in the constancy of the measured dose for the preset combinations of dose rate, gantry speed, and MLC leaf speed.

Results: For the PF test, a 0.3 mm tolerance level was suggested for the positioning of the MLC leaves. The tolerance level for the gap width was 0.5 mm. For the DRGS and DRMLC tests, a 3% tolerance level was proposed.

Conclusions: With the adapted levels of tolerance for an EPID-based approach, the PF, the DRGS, and the DRMLC tests offer a convenient and accurate machine QC program for linear accelerators used for VMAT.
[DOI: 10.1118/1.3552922]

Key words: VMAT, QC, EPID

I. INTRODUCTION

By maximizing dose to target and minimizing dose to organs at risk, intensity modulated radiotherapy (IMRT) has been established as an accurate, reliable, efficient, and highly conformal radiotherapy (RT) technique. The IMRT delivery techniques used in clinical practice are the step-and-shoot technique and the sliding window technique.[1,2] In comparison to conventional RT, IMRT increases the treatment delivery time as a result of a larger number of beam directions and monitor units (MUs). This increases the risk of intrafraction organ motion and thereby the risk of missing the target.[3]

Intensity modulated arc therapy is an alternative IMRT method, combining gantry rotation and beam aperture changes.[4] Initially, clinical implementation was complicated by the fact that the optimization algorithm generated plans that could not be delivered due to limitations of the accelerator. Subsequent research led to optimization algorithms that dealt with the constraints imposed by the accelerator, but not necessarily reduced treatment time; e.g., one approach led to several superimposed gantry rotations.[5,6] However, the more recent approach proposed by Otto, termed volumetric modulated arc therapy (VMAT), uses only one gantry rotation.[7]

VMAT includes features such as variable dose rate and gantry speed in addition to the beam modulation achieved with complex patterns of multileaf collimator (MLC) motion in IMRT. Although VMAT is an extension of the IMRT-DMLC technique, additional quality control (QC) is required.[8–14] A large number of studies have addressed IMRT machine QC,[15–20] but only a few studies have been published on machine QC for VMAT.[21–24] For commissioning and QC, Ling *et al.*[22] proposed three test plans: (1) A picket fence (PF) test, (2) a dose rate versus gantry speed (DRGS) test, and (3) a dose rate versus MLC leaf speed (DRMLC) test. These tests evaluate the VMAT performance in a stepwise manner, so, for example, if the PF test or the DRGS test fails, the reason should be found and solved before evaluating the

following DRMLC test. Ling *et al.*[22] performed their measurements with films, whereas Fredh *et al.*[23] used an electronic portal imaging device (EPID).

The purpose of this study was to create a convenient and accurate machine QC program covering the critical elements of VMAT: The precision of the DMLC, the variable dose rate, and the MLC and gantry speed. The test plans were based on the three tests originally suggested by Ling *et al.*,[22] but all measurements were performed with an EPID. The tests were performed at regular intervals as a part of our routine machine QC program in order to ensure the accurate delivery of VMAT. The study led to the establishment of tolerance levels for all tests in order to indicate when machine adjustments are required.

II. METHODS AND MATERIALS

The PF (test 1), the DRGS (test 2), and the DRMLC (test 3) tests provided by Varian Medical Systems (Palo Alto, CA) were implemented as separate treatment plans in the record and verify system, Aria (Varian Medical Systems, Palo Alto, CA). All plans were exported in DICOM format and treated in DICOM RT mode at the accelerator using a 6 MV photon beam (Varian Clinac, Varian Medical Systems, Palo Alto, CA). The resulting images were acquired using the amorphous silicon (aS) EPIDs available on the treatment units, as described in the following sections. In general, the 16-bit grayscale images have a resolution of 1024×768 pixels when acquired with an aS1000 EPID and a resolution of 512×384 pixels when acquired with an aS500 EPID. In a 16-bit grayscale image, black is assigned the grayscale value 0 and white the grayscale value 2^{16} (i.e., 65 536). For all measurements, the treatment couch was positioned outside the treatment field. The tests were performed six times on each accelerator during a period of at least 4 weeks (the implementation period).

In addition to the PF, DRGS, and DRMLC tests, two additional tests were performed; test 4 detected the relative movement of the EPID and accelerator head during rotation of the gantry and test 5 investigated the problem of EPID saturation as reported by Van Esch *et al.*[25]

II.A. Test 1: PF test

In the PF test,[15,22] the 60 MLC leaf pairs were swept across a 200×390 mm^2 ($X\times Y$) field collimated by the jaws. The MLC stopped every 20 mm, irradiating a 1 mm gap while rotating the gantry 28.6°, resulting in ten strips of narrow gaps acquired during a 352° gantry rotation [see Fig. 1(a)]. The test is a control of the leaf position and the distance between two opposing leaves. The PF test was performed at a collimator angle of 90° in order to fit all the MLC pairs within the rectangular field of view of the EPID. The source to detector distance (SDD) was 1000 mm.

The PF test was analyzed using the IMAGEJ (Ref. 26) software and an image of the test is shown in Fig. 1(a). First, the background level was subtracted, defined as the mean of the pixel values encompassed by the rectangle drawn in Fig. 1(a). This subtraction was performed in order to remove the low dose contribution emerging from the closed MLCs sweeping across the EPID; hence producing better contrast. Afterward, the image was inverted so that areas receiving a high dose showed high pixel values and therefore MLC gaps showed high dose regions. Finally, beam flatness was accounted for by the "rolling ball" algorithm,[27] which removed smooth continuous background variations from the image. The result of this background removal is illustrated in Fig. 1(b).

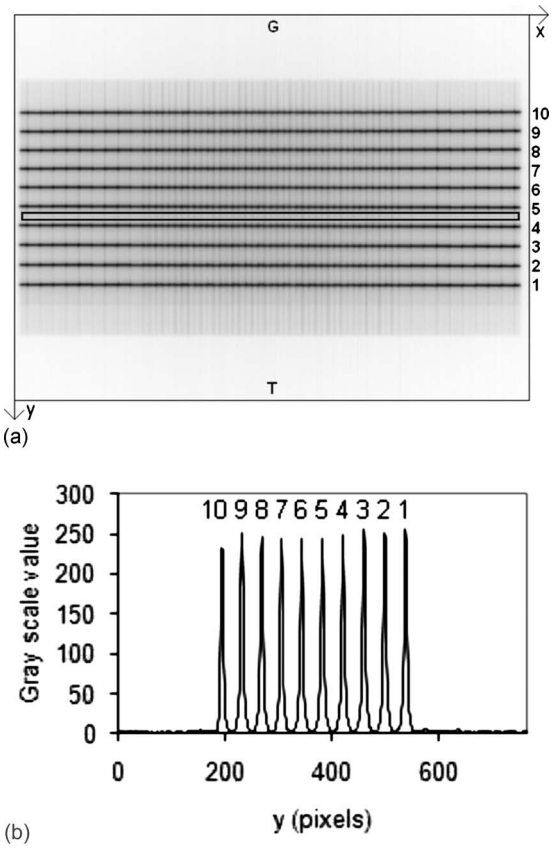

FIG. 1. (a) Grayscale image of the PF test acquired with the EPID. The strips of gaps are numbered 1–10 and the MLC leaves are numbered 1–60 from left to right. During the PF test, the MLC leaves moved in the target-gun direction, from target, T to gun, G. (b) Grayscale values along a central MLC leaf (leaf 30) after the background removal from the PF tests.

In the analysis, the x and y axes are defined as shown in Fig. 1(a). The central x position of each MLC leaf pair was calculated from the position of the vertical lines caused by the tongue and groove effect [see Fig. 1(a)]. The grayscale value at the central position of each MLC leaf pair was denoted by $I(x_{\mathrm{MLC}},y)$. It was determined as the average grayscale value along the x axis over the four central pixels (i.e., 1.5 mm) for an aS1000 EPID and over the two central pixels for an aS500 EPID. The position of the ten successive gaps for each MLC leaf pair in the y direction in Fig. 1(a) was determined from $I(x_{\mathrm{MLC}},y)$: Three successive grayscale values on the y axis having a value above a threshold of 25 is taken to represent the detection of a gap [see Fig. 1(b)]. From inspection of the calculations, it was found that a gap covered approximately 13 pixels for an aS1000 EPID and 7

pixels for an aS500 EPID. For every gap, the position, $p(x_{\mathrm{MLC}},\mathrm{gap})$, is calculated as the weighted average and the intensity, $I_{\mathrm{sum}}(x_{\mathrm{MLC}},\mathrm{gap})$, as the sum of the grayscale values of the gap,

$$p(x_{\mathrm{MLC}},\mathrm{gap}) = \frac{\sum_{I(x_{\mathrm{MLC}},y)>\mathrm{Threshold}} y \cdot I(x_{\mathrm{MLC}},y)}{\sum_{I(x_{\mathrm{MLC}},y)>\mathrm{Threshold}} y},$$

$$I_{\mathrm{sum}}(x_{\mathrm{MLC}},\mathrm{gap}) = \sum_{I(x_{\mathrm{MLC}},y)>\mathrm{Threshold}} I(x_{\mathrm{MLC}},y).$$

The intensity $I_{\mathrm{sum}}(x_{\mathrm{MLC}},\mathrm{gap})$ is an estimate of the width of the gap as, for example; a broadening of the gap would lead to a higher value of the intensity. For each MLC leaf pair, the deviation of the gap intensity with respect to the mean gap intensity over all 60 MLC gaps in the associated strip was calculated. For each PF test, the mean, maximum, and standard deviation (SD) of the gap intensity deviations were calculated using the absolute values. The deviation of the gap intensity is not sensitive to an overall error in the gap width for all MLC leaf pairs and is this why the minimum and maximum gap intensities were also recorded.

The position of the MLC leaf pairs on the x axis and the gap positions on the y axis were correlated due to a possible rotation of the EPID relative to the accelerator head. The rotation was assumed to be constant and small during each test. The extent of the rotation was determined for each strip of gaps by performing a least-squares straight line fit to the gap positions versus the positions of the MLC leaf pairs and gives an estimation of the average gap position. The relative rotation was calculated from the mean slope of the line. For example, a relative rotation of 0.3° would correspond to a 2 mm difference in the position of the first and last MLC leaf pair. The gap position deviation was determined as the deviation from the average gap position. For each PF test, the mean, maximum, and SD of the gap position deviations were calculated.

The sensitivity of the PF test was evaluated by delivering a plan provided by Varian Medical Systems with an intentional inherent positional error of 0.5 mm and width error of 0.5 mm. There were six strips of gaps in the test, which was delivered twice on three different accelerators. This test has approximately 48 MU/strip, whereas the PF test without inherent errors has 30 MU/strip. Thus, the gap intensities had to be scaled by the ratio of the MUs in order to compare the two PF tests.

II.B. Test 2: DRGS test

The DRGS test[22] assessed the control of dose rate and gantry speed. The test consisted of seven segments with different combinations of dose rate and gantry speed chosen such that the delivered dose was identical for all segments (see Table I). The MLC banks formed an open static 200 $\times$20 mm^2 field irradiated during a gantry rotation of 102.6° for the first segment. The MLC banks move simultaneously during beam hold so that a different part of the EPID is irradiated during the gantry rotation forming segment 2. This is the same for the rest of the segments shown in Table I and the resulting image is shown in Fig. 2(a). During the test, the collimator rotation was 90° and the SDD was 1500 mm. For comparison, an open field image was acquired for a fixed gantry angle with a dose rate of 300 MU/min.

TABLE I. Technical details of the seven segments used in the DRGS test. Figure 2(a) displays an EPID image of the DRGS test along with the position of the segments.

Segment No.	Gantry start angle (deg)	Covered angle (deg)	Dose rate (MU/min)	Gantry speed (deg/s)
1	179.0	102.5	105	4.8
2	76.5	51.2	210	4.8
3	25.3	37.5	314	4.8
4	347.8	30.7	417	4.8
5	317.1	26.5	524	4.8
6	290.6	24.6	592	4.8
7	266.0	23.2	600	4.4

The DRGS test was analyzed using the IMAGEJ (Ref. 26) software. As for test 1, the image was inverted, the central

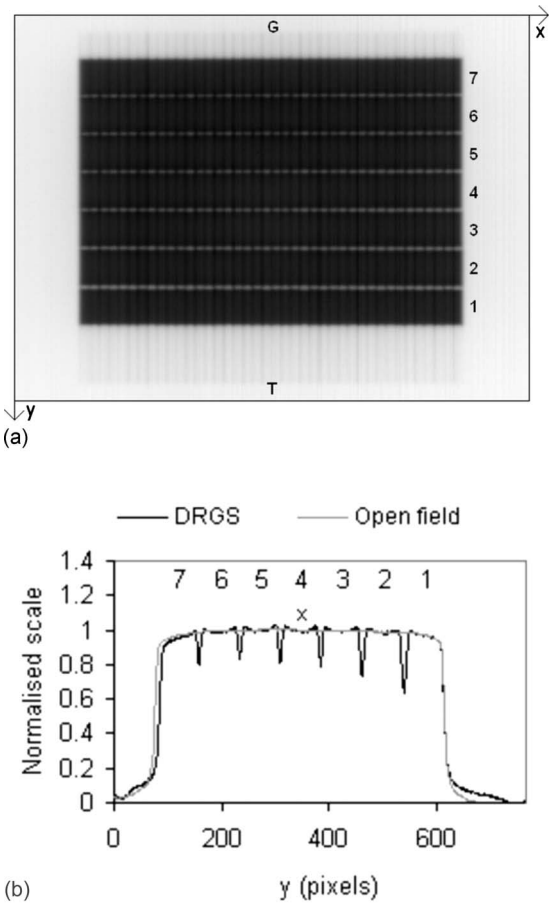

FIG. 2. (a) Grayscale image of the DRGS test acquired with the EPID. During the DRGS test the MLC leaves moved from target, T to gun, G. Technical details of the DRGS segments can be found in Table I. Only the 40 central narrow leaves (MLC leaf 11–50) were investigated in the DRGS test. (b) The test and open profile along a central MLC leaf after normalization to the segment marked by the cross.

TABLE II. Technical details of the four segments used in the DRMLC test. Figure 3(a) displays an EPID image of the DRMLC test along with the position of the segments.

Segment No.	Gantry start angle (deg)	Covered angle (deg)	Dose rate (MU/min)	MLC leaf speed (cm/s)	Gantry speed (deg/s)
1	170	18	480	1.6	4.8
2	152	12	600	2.4	4.0
3	140	32	240	0.8	4.8
4	108	76	120	0.4	4.8

position of each MLC leaf pair was located, and the grayscale value at the central position of each MLC leaf pair, $I(x_{\mathrm{MLC}},y)$, was determined. There is a linear relationship between grayscale value and dose for an amorphous silicon detector.[21] In order to correct the DRGS image for daily output variations, the data were normalized to the central grayscale value of segment 4 at a central MLC leaf pair, i.e., approximately in the middle of the image [see Table I and Figs. 2(a) and 2(b)]. The corresponding location was found for normalization of the open field image. Individual normalization of the two tests were performed since the number of MUs delivered in each segment in the DRGS image and in the open field image differed.

In order to correct the DRGS image for daily flatness and symmetry variations, the ratio $I_{\mathrm{corr}}(x_{\mathrm{MLC}},y)$ between the $I(x_{\mathrm{MLC}},y)$ values measured for the DRGS and the open field image was determined for each MLC pair,

$$I_{\mathrm{corr}}(x_{\mathrm{MLC}},y)=\frac{I_{\mathrm{DRGS}}(x_{\mathrm{MLC}},y)}{I_{\mathrm{Open}}(x_{\mathrm{MLC}},y)}\cdot 100\%.$$

The deviation of the dose in each segment compared to the mean dose value in all segments for each MLC pair was defined as follows:

$$\mathrm{Deviation}(x_{\mathrm{MLC}})_j=\frac{\langle I_{\mathrm{corr}}(x_{\mathrm{MLC}},y)\rangle_j}{\langle I_{\mathrm{corr}}(x_{\mathrm{MLC}},y)\rangle_{j\in[1,m]}}-1,$$

where j refers to the segment number and m is the largest segment number (see Table I). $\langle I_{\mathrm{corr}}(x_{\mathrm{MLC}},y)\rangle$ was determined as the mean value of I_{corr} in 21 pixels (8.2 mm) located around the central position of the segment for the aS1000 EPID (and 11 pixels for the aS500 EPID). Grayscale values close to or in the gap between two segments were excluded. Due to penumbra effects, we also excluded the results from the outermost MLC leaf pairs. The deviations were evaluated for all MLC leaf pairs and segments by calculating the global minimum and maximum deviation, $\Delta_{\mathrm{Min}}^{\mathrm{DRGS}}$ and $\Delta_{\mathrm{Max}}^{\mathrm{DRGS}}$. Furthermore, the mean deviation, $\Delta_{\mathrm{Mean}}^{\mathrm{DRGS}}$, and SD, $\Delta_{\mathrm{SD}}^{\mathrm{DRGS}}$, were calculated by using the absolute values of the total data set.

II.C. Test 3: DRMLC test

The DRMLC test[22] assessed the control of dose rate and MLC leaf speed. It consisted of four segments with different combinations of dose rate and MLC leaf speed chosen so that the delivered dose to each segment was identical (see Table II). To test the MLC leaf speed, the MLC banks were initially closed. One bank was then retracted to produce a 200×20 mm^2 MLC collimated field, after which the other bank moved to close the field. This motion was repeated until all four segments had been delivered to different parts of the EPID [see Fig. 3(a)]. In order to keep the dose constant across the field while reaching the maximum MLC leaf speed of 2.4 cm/s, it was necessary to reduce the gantry speed for segment 2 to 4.0 deg/s (see Table II)). For this test, the collimator rotation was 90° and the SDD was 1500 mm.

For normalization, an open field image was acquired at a fixed gantry angle with a dose rate of 300 MU/min. The

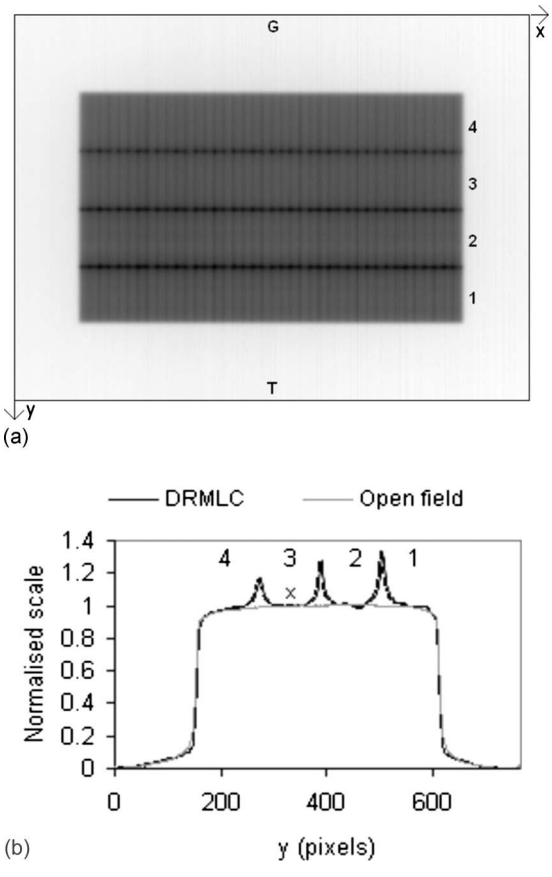

FIG. 3. (a) Grayscale image of the DRMLC test acquired with the EPID. During the DRMLC test, the MLC leaves moved from target, T to gun, G. Technical details of the DRMLC segments can be found in Table II. Only the 40 central narrow leaves (MLC leaf 11–50) were investigated in the DRMLC test. (b) The test and open profile along a central MLC leaf after normalization to the segment marked by the cross.

DRMLC test was analyzed in the same way as test 2 except that the images were normalized to the central grayscale value of segment 3 [see Fig. 3(b)].

II.D. Test 4: EPID movement test

The movement of the EPID relative to the accelerator head could be caused by a mechanical instability of the support arm or by gantry sag. To avoid effects related to the movement of the jaws during gantry rotation, a homemade fixed aperture (3.5 cm square opening) was mounted firmly on the accelerator head.

A total of 68 images was acquired using the EPID in the continuous image acquisition mode during a full gantry rotation. For all images, the center of the field was recorded and differences in its position were calculated relative to that at gantry angle 0°. Using beam's eye view, the differences in position may be separated into two directional shifts; the gun-target direction (i.e., longitudinal) or perpendicular to the gun-target direction (left-right). This test was performed on two accelerators.

II.E. Test 5: EPID saturation test

Van Esch *et al.*[25] described saturation problems for an aS500 EPID operating with the image acquisition unit IAS2. Using a 6 MV photon beam, they reported a dose error of approximately 0.8% at the isocenter for a dose rate of 240 MU/min, which increased to 25% for a dose rate of 600 MU/min. The aS1000 EPIDs and aS500 EPIDs used in the present study all operate with image acquisition unit IAS3.

To investigate the possibility of saturation effects being present, the DRMLC test was performed at the following SDDs: 1000, 1200, 1500, 1600, 1700, and 1800 mm. The test was performed at two aS1000 EPIDs and one aS500 EPID. As for test 3, the images were inverted and normalized to segment 3 delivered with a dose rate of 240 MU/min. For simplicity, only the central MLC leaf pair (leaf 30) was investigated. Segment 2 was delivered with the highest dose rate of 600 MU/min (Table II) and therefore saturation effects should be most pronounced at that segment. If present, the mean value of $I(x_{\mathrm{MLC}}, y)$ should increase with increasing SDD for that segment.

III. RESULTS

Each of the 54 PF, DRGS, and DRMLC tests was carefully examined both visually and analytically as described above. A few of the tests showed large deviations from the typical results achieved for a particular accelerator or when compared to another accelerator equipped with the same type of EPID. These tests were marked as outliers and were investigated separately in order to identify possible technical explanations. For the remainder of the six tests performed on each accelerator, average values of the test results were calculated and are presented as follows.

III.A. Test 1: The PF test

In Table III, the mean value of the gap position deviation is given for all accelerators. On the average, it was found to be 0.04 ± 0.03 mm, with a maximum deviation of 0.21 mm (quoted only for aS1000 EPID). The relative rotation between the gantry head and the EPID was found to be 0.8 mm with a maximum value of 1.3 mm (quoted only for aS1000 EPID) (see Table III). For accelerator 2, the relative rotation between the gantry head and the EPID was 3.0 mm. A visual comparison between the cross hair and the lasers confirmed the misalignment of the gantry head. It was corrected by calibrating the collimator rotation after the implementation period.

Table III also presents the gap intensity deviation for all accelerators. On the average, it was found to be 70 grayscale values with a standard deviation of 51 and a maximum deviation of 278 for an aS1000 EPID. For an aS500 EPID, it was found to be 29 with a standard deviation of 20 and a maximum deviation of 104. The range of gap intensities (from minimum to maximum values) is also given in Table III. The gap intensities are accelerator-specific due to different sensitivities of the EPIDs.

Inspection of outlier A (with a maximum gap intensity deviation of 333 grayscale values) revealed only one MLC gap with a deviation above 290 grayscale values. Inspection of outlier B (with a maximum gap intensity deviation of 412 grayscale values) revealed that one MLC pair produced gap openings that were too small: Approximately 0.5 mm as estimated from the gap position deviation. This was not observed for the following measurement at the accelerator. Both outliers may disappear if the MLC motors, T-nuts, and/or the soft pots are replaced.

Based on the maximum deviations in Table III, the tolerance for the gap position deviation was suggested to be 0.3 mm for all accelerators. Similarly, based on the results for the maximum gap intensity deviations, the tolerances for the gap intensity deviation were proposed to be 300 grayscale values for an aS1000 EPID and 120 grayscale values for an aS500 EPID. As a result of the different sensitivities of the EPIDs, the minimum and maximum gap intensities were assigned accelerator-specific tolerances. These were established as the values in Table III rounded off to the nearest hundred either below or above the minimum and maximum values.

The Varian supplied PF test with inherent errors was delivered and analyzed to test the sensitivity of these tolerances (see Fig. 4). For leaf 31, the mean gap position deviation was 0.48 mm, well in excess of the proposed tolerance of 0.3 mm. For leaf 26, the mean gap position deviation was 0.27 mm, i.e., below the tolerance. The mean gap intensity deviation is 349 grayscale values for an aS1000 EPID and 144 grayscale values for an aS500 EPID. Both results are above the proposed tolerance. The gap position deviation at leaf 26 indicated that the inherent error is due to the misplacement of one of the MLC leaves in the MLC leaf pair by 0.5 mm. If it had been a symmetrical displacement, the gap position deviation for that leaf would be much smaller.

TABLE III. The test results of the PF test for the aS1000 EPID (upper) and the aS500 EPID (lower). For all accelerators, the tables list the relative rotation, the gap position deviation, the gap intensity deviation, and the gap intensity. Furthermore, the mean value and the standard deviation of the gap position deviation and gap intensity deviation are given. For clarity, the relative rotation, gap position deviation, and gap intensity deviations are averaged (mean) for a fast comparison with the outliers (indicated in bold). The aS500 EPID on accelerators 7 and 8 has a lower sensitivity than the aS1000 EPID with which the other accelerators are equipped. (Acc. 3 is currently under implementation and therefore no data are reported.)

		Gap position deviation (mm)			Gap intensity deviation (grayscale value)			Gap intensity, strips 1–9[a] (grayscale value)	
aS1000 EPID	Relative rotation (mm)	Mean	SD	Max	Mean	SD	Max	Min	Max
Acc. 1	0.7	0.04	0.03	0.15	71	53	264	1334	1873
Acc. 4	1.3	0.04	0.03	0.14	60	45	224	1503	1960
Acc. 5	0.0	0.04	0.03	0.21	61	46	212	1163	1609
Acc. 6	0.5	0.04	0.03	0.13	91	63	278	1354	1931
Acc. 9	1.3	0.04	0.03	0.16	62	45	238	1326	1835
Acc. 10	1.1	0.04	0.03	0.12	72	52	238	1450	1934
Mean	0.8	0.04	0.03	0.15	70	51	242		
Acc. 2	**3.0**	0.06	0.04	0.27	89	66	359	1420	2139
Outlier A[b]	0.7	0.04	0.03	0.13	92	65	**333**	1335	1969
Outlier B[c]	1.2	0.04	0.04	0.25	66	60	**412**	1129	1820

		Gap position deviation (mm)			Gap intensity deviation (grayscale value)			Gap intensity, strips 1–9[a] (grayscale value)	
aS500 EPID	Relative rotation (mm)	Mean	SD	Max	Mean	SD	Max	Min	Max
Acc. 7	0.1	0.04	0.03	0.20	29	20	103	531	730
Acc. 8	1.8	0.05	0.03	0.19	28	20	104	505	757
Mean	1.0	0.05	0.03	0.20	29	20	104		

[a]The intensity of strip 10 consistently fell below the intensity of the other strips and therefore no data are reported.
[b]Reported as a mean of two consecutive measurements.
[c]Only one measurement and not observed for the following measurement.

III.B. Test 2: The DRGS test

In Table IV, the mean deviation, $\Delta_{\text{Mean}}^{\text{DRGS}}$, is shown for all accelerators. On the average, it was found to be $0.6\% \pm 0.4\%$, with a minimum deviation of -2.3% and a maximum deviation of 2.0%.

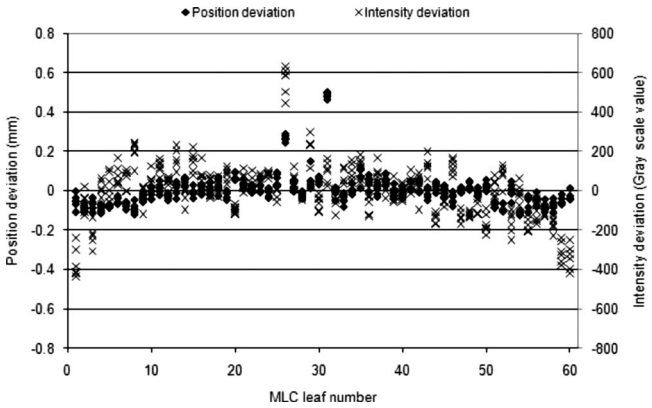

FIG. 4. The result of an analysis of the PF test with inherent errors in position and width. For each MLC, ◆ shows the gap position deviation for each of the six strips. Likewise, × shows the gap intensity deviation for each of the six strips. For leaf 31, the mean gap position deviation determined as the mean value of the gap position deviations in the six strips is 0.48 mm. For leaf 26, the mean gap position deviation is 0.27 mm. In addition, the mean intensity deviation prior to rescaling has a grayscale value of 559 for leaf 26. The errors can also be visually identified in the grayscale image. Similar results were obtained for all PF test with inherent errors.

Outlier C, with a minimum deviation of -49.7% and a maximum deviation of 8.7%, was caused by a loss of EPID communication around the gantry angle of the last segment [see Fig. 5(a)]. Such communication problems may result in various deviations, which may be misinterpreted as problems in the link between dose rate and gantry speed [see outlier α

TABLE IV. The minimum and maximum deviations for the DRGS test on each accelerator, $\Delta_{\text{Min}}^{\text{DRGS}}$ and $\Delta_{\text{Max}}^{\text{DRGS}}$. The typical deviations are given by the mean deviation, $\Delta_{\text{Mean}}^{\text{DRGS}}$, with related SD, $\Delta_{\text{SD}}^{\text{DRGS}}$. For clarity, the values have been averaged for a fast comparison to the outliers (indicated in bold).

	Mean deviation (%)	Mean SD (%)	Deviation min (%)	Deviation max (%)
	$\Delta_{\text{Mean}}^{\text{DRGS}}$	$\Delta_{\text{SD}}^{\text{DRGS}}$	$\Delta_{\text{Min}}^{\text{DRGS}}$	$\Delta_{\text{Max}}^{\text{DRGS}}$
Acc. 1	0.6	0.5	−1.9	1.9
Acc. 2	0.5	0.4	−1.4	1.9
Acc. 4	0.4	0.4	−1.6	1.5
Acc. 5	0.6	0.6	−2.3	1.7
Acc. 6	0.9	0.4	−2.0	1.7
Acc. 7	0.7	0.4	−1.7	1.9
Acc. 8	0.6	0.5	−1.7	2.0
Acc. 9	0.3	0.3	−1.2	0.9
Acc. 10	0.5	0.4	−1.8	1.9
Mean	0.6	0.4	−1.7	1.7
Outlier C[a]	13.8	14.2	**−49.7**	**8.7**

[a]Reported as a mean of three consecutive measurements.

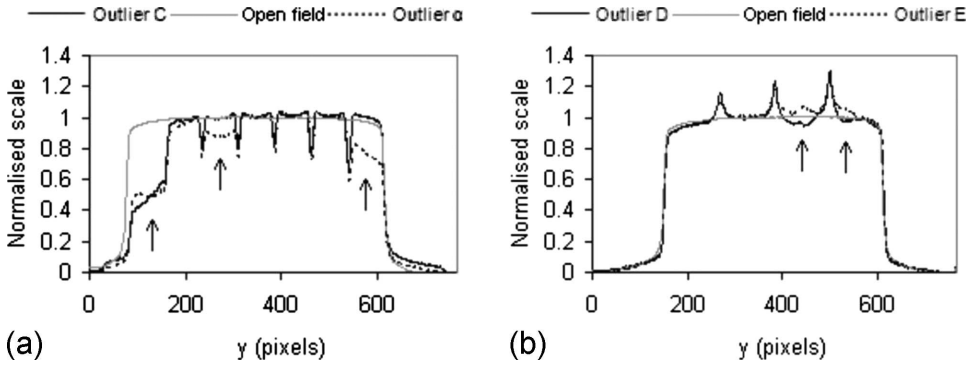

FIG. 5. Observed outliers in the DRGS test (a) and the DRMLC test (b) are indicated by the arrows. Numerical values are reported in Tables IV and V. Outliers C and α, the latter observed after the implementation period had ended, were caused by loss of EPID communication. They disappeared after replacement of cables. Outliers D and E are assumed to be repairable by adjusting the accelerator.

in Fig. 5(a)]. After replacement of cables, the results of the analysis were similar to the previous results obtained from that accelerator. Based on the typical minimum and maximum deviations in Table IV, the tolerance for the minimum and maximum deviations was suggested to be ±3%.

III.C. Test 3: The DRMLC test

Table V gives the mean deviation, $\Delta_{\text{Mean}}^{\text{DRMLC}}$, for all accelerators. On the average, it was found to be 0.9% ± 0.5%, with a minimum deviation of −2.5% and a maximum deviation of 2.0%. Outlier D, with a minimum deviation of −3.4%, and outlier E, with a minimum deviation of −3.8% and a maximum deviation of 3.6%, were both due to an unstable dose rate [see Fig. 5(b)]. Based on the outliers and the typical minimum and maximum deviations in Table V, the tolerance for the minimum and maximum deviations was proposed to be ±3%.

TABLE V. The minimum and maximum deviations for the DRMLC test on each accelerator, $\Delta_{\text{Min}}^{\text{DRMLC}}$ and $\Delta_{\text{Max}}^{\text{DRMLC}}$. The typical deviations are given by the mean deviation, $\Delta_{\text{Mean}}^{\text{DRMLC}}$, with related SD, $\Delta_{\text{SD}}^{\text{DRMLC}}$. For clarity, the values have been averaged for a fast comparison with the outliers (indicated in bold).

	Mean deviation (%)	Mean SD (%)	Deviation min (%)	Deviation max (%)
	$\Delta_{\text{Mean}}^{\text{DRMLC}}$	$\Delta_{\text{SD}}^{\text{DRMLC}}$	$\Delta_{\text{Min}}^{\text{DRMLC}}$	$\Delta_{\text{Max}}^{\text{DRMLC}}$
Acc. 1	1.0	0.6	−1.8	1.9
Acc. 2	1.2	0.8	−2.5	2.0
Acc. 4	0.8	0.5	−1.4	1.7
Acc. 5	1.0	0.4	−2.0	1.6
Acc. 6	0.7	0.4	−1.7	1.3
Acc. 7	0.7	0.5	−1.6	1.7
Acc. 8	0.8	0.5	−1.8	1.4
Acc. 9	0.8	0.4	−1.6	1.4
Acc. 10	1.1	0.6	−2.2	1.8
Mean	0.9	0.5	−1.8	1.6
Outlier D[a]	1.6	1.1	**−3.4**	1.6
Outlier E[a]	2.1	1.1	**−3.8**	**3.6**

[a]Only one measurement and not observed for the following measurement.

III.D. Test 4: The EPID movement test

For a full rotation of the gantry, the maximum relative movement of the EPID and the accelerator head was found to be 1.6 mm in the gun-target direction and 1.2 mm in the left-right direction for both accelerators. At most, the movement of the EPID and the accelerator head for a 30° gantry rotation is approximately 0.4 mm in the gun-target direction.

III.E. Test 5: The EPID saturation test

Figure 6 shows segment 3 and segment 2 of the DRMLC tests performed at different SDD. The variations within segment 2 are much more variable than within segment 3. These variations are assumed to be related to the increase in MLC leaf speed between the two segments. For segment 2 in Fig. 6, the average values of $I(x_{\text{MLC}}, y)$ were 1.017, 1.008, 1.015, 1.018, 1.021, and 1.017 for increasing SDD (1000, 1200, 1500, 1600, 1700, and 1800 mm). It was concluded that saturation is not an issue because the average values were independent of the SDD.

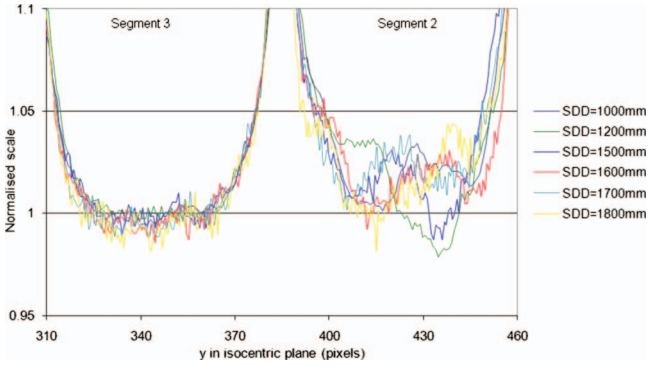

FIG. 6. Normalized profiles along the y axis for a central MLC leaf pair obtained with an aS1000 EPID at different SDDs. For clarity, the profiles have been calculated on the isocentric plane. Segment 3 appears from approximately pixel 330–360 and segment 2 from approximately pixel 410–440. A similar result was obtained for another aS1000 EPID and for the aS500 EPID.

IV. DISCUSSION

A large number of repeated QC tests performed across a group of treatment units have established tolerances of ±0.3 mm for the MLC leaf gap position deviation, ±300 grayscale values for the MLC leaf gap intensity deviation for an aS1000 EPID, and ±120 grayscale values for the MLC leaf gap intensity deviation for an aS500 EPID. The provided PF test with inherent errors showed that the gap intensity deviation is a reliable measure of the gap width. Outlier B (Table III) indicated that width deviations of 0.5 mm are detectable.

The MLC positional tolerance is within the AAPM TG 142 recommendations of ±1 mm and the ±0.5 mm proposed in the ESTRO guidelines (see Table III).[28,29] The maximum gap position deviation found by Fredh *et al.*[23] was 0.48 mm (SD of 0.13 mm), which is larger than the observations in the present work (see Table III). Fredh *et al.*[23] detected a width error as an error in the positioning of one of the MLC leaves, which will lead to a shift in the detected position of the gap and will be observable as half the value of the actual misalignment. This shows that width errors are more difficult to detect than positional errors and is therefore why the intensity was taken as an estimator of the gap width (see Fig. 4, leaf 26). It should be noted that Ling *et al.*[22] only evaluated their measurements visually. Van Esch *et al.*[20] reported a positional reliability better than 0.2 mm for PF tests performed at fixed gantry angles. The tolerance suggested from the present study is a compromise between clinical relevance and the mechanical stability of the accelerators.

If the relative rotation between the EPID and the accelerator head becomes too large, the analysis will be performed at an inclined angle to the gap and that will result in an apparent increase in the gap intensity. Furthermore, the determination of the gap position is affected because the analysis at the outermost strips will be performed close to the edge of the leaves where the intensity varies due to the tongue and groove effect [see Fig. 1(a)]. Accelerator 2 in Table III indicates that this may lead to the misinterpretation that the gap position deviation, gap intensity deviation, and gap intensity are all too high. Therefore, all data collected for accelerator 2 were omitted due to the erroneous collimator rotation. After that was corrected, the results for this accelerator became similar to the other accelerators equipped with an aS1000 EPID.

For both the DRGS test and the DRMLC test, a tolerance for the minimum and maximum deviations of ±3% was established for the deviation between the dose delivered in the test segments and an open field. This agrees with the ±3% confidence limit proposed in the ESTRO guideline for a field with a high dose and a low dose gradient.[29] In portal dose predictions used for patient specific QC of IMRT, a 3%/3 mm criterion is used for the gamma evaluation, indicating that a ±3% tolerance on the minimum and maximum dose deviation is appropriate.[22,25,31] The mean deviation for the DRGS test is approximately 2% (range of 0.9%–2.3%). This is about 1% higher than that found in the studies by Ling *et al.*[22] and Fredh *et al.*[23] The difference could be caused by noise, as indicated by a 0.1% higher standard deviation than that reported by Fredh *et al.*[23] Our results could be improved by averaging the corrected pixel value readings ($I_{corr}(x_{MLC}, y)$) for all MLC leaves, as the leaves are stationary in the DRGS test and therefore do not influence the result. The DRGS test evaluates the ability of the accelerator to adjust the gantry speed when the dose rate changes so that the final delivered dose is correct. The tolerance is the limit within which the link between the dose rate and the gantry speed is considered to be acceptable.

Our mean deviation for the DRMLC test [approximately 2% (range of 1.3%–2.2%)] is 1.5% higher than the value found by Ling *et al.*[22] Fredh *et al.*[23] found deviations between 1.2% and 5.4% with unexplained additional outliers of up to 17%. The latter group therefore rejected this test for QC of VMAT. This study observed two outliers deviating by approximately 3.5% for the DRMLC test, but both were observed at accelerators incapable of maintaining a stable dose rate. Furthermore, it was seen that cable communication problems might lead to large deviations. According to Ling *et al.*,[22] the MLC leaf position error increases linearly with the leaf speed. The present study confirms this result, suggesting that it is the reason for the larger variations within segment 2 (MLC leaf speed of 2.4 cm/s) compared to segment 3 (MLC leaf speed of 0.8 cm/s), as shown in Fig. 6. Test 5, the EPID saturation test, strongly indicates that saturation is not an issue for the image acquisition unit, IAS3, and can be ignored. This was confirmed by observations by Nicolini *et al.*[30] Overall, the ±3% tolerance for the DRMLC test is found to be acceptable from the data collected and from the 3%/3 mm criterion used for the gamma evaluation of portal dose predictions in patient-specific QC.[22,25,31] The DRMLC test evaluates the ability of the accelerator to adjust the MLC leaf speed when the dose rate changes so that the dose delivered is constant. The tolerance is the limit within which the link between the dose rate and the MLC leaf speed is considered to be adequate.

The reliability of the analysis of the PF, DRGS, and DRMLC tests depends on the relative movement between the EPID and the accelerator head being negligible. In the present study, it was found that the EPID was able to move at most 1.6 mm in the gun-target direction and this confirms the observations of Nicolini *et al.*[21] They observed a left-right movement of only 0.3 mm for a full gantry rotation, which is significantly less than the finding of 1.2 mm in the present study. Their finding may be due to the movement of the jaws in the same direction as the EPID.

It should be pointed out that the narrowest MLC leaves are 5 mm wide and the analysis in this study is performed centrally (±0.8 mm) relative to the leaves in order to stay clear of the intensity variation caused by the tongue and groove effect [see Figs. 1(a), 2(a), and 3(a)]. Thus, a 1.2 mm left-right movement would not influence the results of the PF, DRGS, and DRMLC tests.

Similarly, for the PF test, a strip of gaps is acquired simultaneously during a 28.6° gantry rotation, for which the

movement is approximately 0.4 mm in the gun-target direction. In the analysis of the PF test, the position and intensity of each MLC leaf gap were compared to mean values calculated from all gaps within the same strip and therefore the 0.4 mm gun-target movement has no influence on the analysis. However, it could influence the relative position of the strips by up to 0.4 mm. This is not considered to be relevant.

The MLC leaves are stationary during delivery of a segment in the DRGS test. Therefore, a movement in the gun-target direction will broaden the edges of the segment in question [see Fig. 2(a)]. As the evaluation is performed at the center of the segments, the movement in the gun-target direction has no influence on the test result.

Because the MLC leaves are moving in the segments of a DRMLC test, a relative movement of the EPID and the accelerator head causes blurring in a segment—i.e., intensity variation. If a large deviation is observed in a DRMLC test and it occurs for all MLC leaves, it should be considered whether the relative movement of the EPID and the accelerator head causes the deviation.

The PF, the DRGS, and the DRMLC tests primarily focus on the dynamic performance of the accelerator. These tests do not replace standard geometric machine QC of, e.g., the gantry angle or the collimator angle that might influence the results of the PF, DRGS, and DRMLC tests. Moreover, neither of the tests evaluates the absolute dosimetry of the accelerator or the absolute position of the MLCs.

In conclusion, from the typical results of the PF, DRGS, and DRMLC tests, it was possible to establish a level of tolerance within which the MLC positioning, and the links between the dose rate and the gantry speed, and the dose rate and MLC speed were considered to be adequate. From the levels of tolerance determined, it could be concluded that the PF, DRGS, and DRMLC tests offer a convenient and accurate machine QC tool for linear accelerators used for VMAT. It is still important to bear in mind that these tests do not replace the patient-specific QC for which additional measurements are needed.

[a]Our department has a research collaboration agreement with Varian Medical Systems yet not related to the research activities presented in this paper.

[b]Author to whom correspondence should be addressed. Electronic mail: maibjoer@rm.dk; Telephone: (+45)89494465 [from 2010-11-29: (+45)78464465]; Fax: (+45)89494522.

[1]P. Xia and L. J. Verhey, "Multileaf collimator leaf sequencing algorithm for intensity modulated beams with multiple static segments," Med. Phys. **25**(8), 1424–1434 (1998).

[2]C. C. Ling, C. Burman, C. S. Chui, G. J. Kutcher, S. A. Leibel, T. LoSasso, R. Mohan, T. Bortfeld, L. Reinstein, S. Spirou, X. H. Wang, Q. Wu, M. Zelefsky, and Z. Fuks, "Conformal radiation treatment of prostate cancer using inversely-planned intensity-modulated photon beams produced with dynamic multileaf collimation," Int. J. Radiat. Oncol., Biol., Phys. **35**, 721–730 (1996).

[3]P. R. Poulsen, B. Cho, K. Langen, P. Kupelian, and P. J. Keall, "Three-dimensional prostate position estimation with a single x-ray imager utilizing the spatial probability density," Phys. Med. Biol. **53**, 4331–4353 (2008).

[4]C. X. Yu, "Intensity-modulated arc therapy with dynamic multileaf collimation: An alternative to tomotherapy," Phys. Med. Biol. **40**(9), 1435–1449 (1995).

[5]M. A. Earl, D. M. Shepard, S. Naqvi, X. A. Li, and C. X. Yu, "Inverse planning for intensity-modulated arc therapy using direct aperture optimization," Phys. Med. Biol. **48**(8), 1075–1089 (2003).

[6]D. M. Shepard, D. Cao, M. K. N. Afghan, and M. A. Earl, "An arc-sequencing algorithm for intensity modulated arc therapy," Med. Phys. **34**(2), 464–470 (2007).

[7]K. Otto, "Volumetric modulated arc therapy: IMRT in a single gantry arc," Med. Phys. **35**(1), 310–317 (2008).

[8]S. Korreman, J. Medin, and F. Kjær-Kristoffersen, "Dosimetric verification of RapidArc treatment delivery," Acta Oncol. (Madr) **48**(2), 185–191 (2009).

[9]E. Schreibmann, A. Dhabaan, E. Elder, and T. Fox, "Patient-specific quality assurance method for VMAT treatment delivery," Med. Phys. **36**(10), 4530–4535 (2009).

[10]C. X. Yu, X. A. Li, L. Ma, D. Chen, S. Naqvi, D. Shepard, M. Sarfaraz, T. W. Holmes, M. Suntharalingam, and C. M. Mansfield, "Clinical implementation of intensity-modulated arc therapy," Int. J. Radiat. Oncol., Biol., Phys. **53**(2), 453–463 (2002).

[11]T. Teke, A. M. Bergman, W. Kwa, B. Gill, C. Duzenli, and I. A. Popescu, "Monte Carlo based, patient-specific RapidArc QA using linac log files," Med. Phys. **37**(1), 116–123 (2010).

[12]P. A. Jursinic, R. Sharma, and J. Reuter, "MapCHECK used for rotational IMRT measurements: Step and shoot, Tomotherapy, RapidArc," Med. Phys. **37**(6), 2837–2846 (2010).

[13]D. Létourneau, J. Publicover, J. Kozelka, D. J. Moseley, and D. A. Jaffray, "Novel dosimetric phantom for quality assurance of volumetric modulated arc therapy," Med. Phys. **36**(5), 1813–1821 (2009).

[14]M. Iori, E. Cagni, M. Paiusco, P. Munro, and A. E. Nahum, "Dosimetric verification of IMAT delivery with a conventional EPID system and a commercial portal dose image prediction tool," Med. Phys. **37**(1), 377–390 (2010).

[15]C.-S. Chui, S. Spirou, and T. LoSasso, "Testing of dynamic multileaf collimation," Med. Phys. **23**(5), 635–641 (1996).

[16]T. LoSasso, C.-S. Chui, and C. C. Ling, "Physical and dosimetric aspects of a multileaf collimation system used in the dynamic mode for implementing intensity modulated radiotherapy," Med. Phys. **25**(10), 1919–1927 (1998).

[17]J. Chang, C. H. Obcemea, J. Sillanpaa, J. Mechalakos, and C. Burman, "Use of EPID for leaf position accuracy QA of dynamic multi-leaf collimator (DMLC) treatment," Med. Phys. **31**(7), 2091–2096 (2004).

[18]G. J. Budgell, J. H. L. Mott, P. C. Williams, and K. J. Brown, "Requirements for leaf position accuracy for dynamic multileaf collimation," Phys. Med. Biol. **45**(5), 1211–1227 (2000).

[19]M. Partridge, P. M. Evans, M. van Herk, L. S. Ploeger, G. J. Budgell, and H. V. James, "Leaf position verification during dynamic beam delivery: A comparison of three applications using electronic portal imaging," Med. Phys. **27**(7), 1601–1609 (2000).

[20]A. Van Esch, J. Bohsung, P. Sovari, M. Tenhunen, M. Paiusco, M. Iori, P. Engström, H. Nyström, and D. P. Huyskens, "Acceptance tests and quality control (QC) procedures for the clinical implementation of intensity modulated radiotherapy (IMRT) using inverse planning and the sliding window technique: Experience from five radiotherapy departments," Radiother. Oncol. **65**, 53–70 (2002).

[21]G. Nicolini, E. Vanetti, A. Clivio, A. Fogliatta, S. Korreman, J. Bocanek, and L. Cozzi, "The GLAaS algorithm for portal dosimetry and quality assurance of RapidArc, an intensity modulated rotational therapy," Radiat. Oncol. 3(24) (2008).

[22]C. C. Ling, P. Zhang, Y. Archambault, J. Bocanek, G. Tang, and T. LoSasso, "Commissioning and quality assurance of RapidArc radiotherapy delivery system," Int. J. Radiat. Oncol., Biol., Phys. **72**(2), 575–581 (2008).

[23]A. Fredh, S. Korreman, and P. Munck af Rosenschöld, "Automated analysis of images acquired with electronic portal imaging device during delivery of quality assurance plans for inversely optimized arc therapy," Radiother. Oncol. **94**, 195–198 (2010).

[24]J. L. Bedford and A. P. Warrington, "Commissioning of volumetric modulated arc therapy (VMAT)," Int. J. Radiat. Oncol., Biol., Phys. **73**(2), 537–545 (2009).

[25]A. Van Esch, T. Depuydt, and D. P. Huyskens, "The use of an aSi-based EPID for routine absolute dosimetric pre-treatment verification of dynamic IMRT fields," Radiother. Oncol. **71**, 223–234 (2004).

[26]W. S. Rasband, IMAGEJ, U.S. National Institutes of Health, Bethesda, MD (http://rsb.info.nih.gov/ij/), 2009.

[27]S. R. Sternberg, "Biomedical image processing," Computer **16**(1), 22–34 (1983).
[28]E. E. Klein, J. Hanley, J. Bayouth, F.-F. Yin, W. Simon, S. Dresser, C. Serago, F. Aguirre, L. Ma, B. Arjomandy, and C. Liu, "Task Group 142 report: Quality assurance of medical accelerators," Med. Phys. **36**(9), 4197–4212 (2009).
[29]M. Alber *et al.*, ESTRO Booklet No. 9: Guidelines for verification of IMRT (ESTRO, 2008).
[30]G. Nicolini, E. Vanetti, A. Clivio, A. Fogliata, G. Boka, and L. Cozzi, "Testing the portal imager GLAaS algorithm for machine quality assurance," Radiat. Oncol. 3(14) (2008).
[31]P. S. Basran and M. K. Woo, "An analysis of tolerance levels in IMRT quality assurance procedures," Med. Phys. **35**(6), 2300–2307 (2008).