Medical Physics
The International Journal of Medical Physics Research and Practice

On the insensitivity of single field planar dosimetry to IMRT inaccuracies

Jon J. Kruse

Citation: Medical Physics **37**, 2516 (2010); doi: 10.1118/1.3425781
View online: http://dx.doi.org/10.1118/1.3425781
View Table of Contents: http://scitation.aip.org/content/aapm/journal/medphys/37/6?ver=pdfcov
Published by the American Association of Physicists in Medicine

On the insensitivity of single field planar dosimetry to IMRT inaccuracies

Jon J. Kruse[a)]
Department of Radiation Oncology, Mayo Clinic, Rochester, Minnesota 55905

(Received 11 November 2009; revised 6 April 2010; accepted for publication 7 April 2010; published 12 May 2010)

Purpose: To report on the sensitivity of single field planar measurements in identifying IMRT plans with poor calculational accuracy.
Methods: Three IMRT plans for head and neck cancer were subjected to extensive quality assurance. The plans were recalculated on a cylindrical phantom and between eight and 18 low gradient points were measured in each plan with an ion chamber. Every point measured in these plans agreed to within 4% of the dose predicted by the planning system and the plans were judged acceptable for clinical use. Each plan was then reoptimized with aggressive dose constraints so that the new treatment fields were more highly modulated than the ones from the original plans. Very complex fields can be calculated less accurately and ion chamber measurements of these plans in the cylindrical phantom confirmed significant dosimetric errors—Several of the measured points in each plan differed from the calculated dose by more than 4%, with a maximum single deviation of 10.6%. These three plans were judged unacceptable for clinical use. All six plans (three acceptable, three unacceptable) were then analyzed with two means of individual field planar dosimetry: Portal imaging with an electronic portal imaging device (EPID) and an ion chamber array. Gamma analysis was performed on each set of planar measurements with 2%/2 mm distance to agreement (DTA) and 3%/3 mm DTA criteria to try to determine a gamma analysis threshold which would differentiate the flawed plans from the acceptable ones.
Results: With the EPID and 2%/2 mm DTA criteria, between 88.2% and 92.8% of pixels from the acceptable IMRT plans passed the gamma analysis, and between 87.5% and 91.9% passed for the unacceptable IMRT plans. With the ion chamber array and 2%/2 mm DTA criteria, between 92.4% and 94.9% of points in the acceptable plans passed the gamma analysis, while 86.8% to 98.3% of the points in the unacceptable plans passed the gamma analysis. The difference between acceptable and unacceptable plans was diminished further when gamma criteria were expanded to 3%/3 mm DTA. A fraction of pixels passing the gamma analysis was found to be a poor predictor of dosimetric accuracy with both planar dosimeters, as well as both sets of gamma criteria.
Conclusions: Deconstruction of an IMRT plan for field-by-field QA requires complex analysis methods such as the gamma function. Distance to agreement, a component of the gamma function, has clinical relevance in a composite plan but when applied to individual, highly modulated fields, it can mask important dosimetric errors. While single field planar dosimetry may comprise one facet of an effective QA protocol, gamma analysis of single field measurements is insensitive to important dosimetric inaccuracies of the overall plan. [DOI: 10.1118/1.3425781]

Key words: IMRT, dosimetry, quality assurance

I. INTRODUCTION

Despite many years of widespread clinical experience, a consensus on the proper approach and thresholds for patient specific IMRT quality assurance (QA) has yet to emerge. For conventional radiation therapy, the accepted standard is that the dose delivered to the patient should be within 5% of the planned value,[1] recognizing that imperfect dose calculation from the planning system represents only one potential source of deviation between planned and delivered dose. AAPM Task Group 53, in a report that predated the widespread adoption of IMRT, suggested that planning systems should be able to model high dose, low gradient areas to within 2%.[2] Subsequent work has shown that IMRT planning systems are often unable to achieve this level of accuracy,[3] compelling the implementation of enhanced pretreatment QA of IMRT plans. A recent report by AAPM Task Group 119 (Ref. 4) provides guidance on the type of commissioning and pretreatment QA measurements that may be performed, as well as confidence limits for the results of these tests.

The literature suggests that while the number of IMRT plans with poor dosimetric accuracy may be low, it is not zero and QA protocols should be sensitive to aberrant calculations. Dong *et al.*[5] reported on the results of 1591 ion chamber measurements of 751 IMRT plans and found deviations ranging from −12.7% to 11.7%. They note that three of the five cases with large deviations had complicated intensity patterns, suggesting that complex plans should garner a higher level of attention from the physicist responsible for QA. Breen *et al.*[6] published a similar summary of their experience in verifying IMRT plans. In that series, doses mea-

sured within the planning target volumes (PTVs) demonstrated a sizable range (−4.3% to 6.6%). Measurements within organs at risk (OR) had a mean discrepancy of 4.5%, with a range of −4.5% to 16.3%. When the treatment planning system was upgraded with a more sophisticated MLC model, PTV dose deviations were between −3.0% and 5.0% and OR doses ranged between −5.8% and 4.4%. Since 2001, the Radiological Physics Center has been auditing institutions' abilities to plan and deliver IMRT treatments to an anthropomorphic head and neck phantom which is loaded with film and TLDs. The criteria for passing the audit require that the dose recorded by the TLDs agree to within 7% of the planned dose. By June 2008, 416 institutions had irradiated the phantom 572 times and the overall pass rate was only 74%; an improvement over the 57% pass rate through mid 2003.[7]

Measurement of a cumulative IMRT plan with an ion chamber is a straightforward approach to IMRT QA. The primary advantage of measuring dose in the total plan is that when all fields are combined in a single plan, as in an actual patient treatment, clinically relevant regions of the plan such as those covering a target volume generally exhibit low dose gradients. Even some low dose volumes, such as the region that spares the spinal cord in a head and neck treatment, are often free of dose gradients in the total plan. Therefore, ion chamber measurements can be interpreted unambiguously; the agreement between measurement and the calculation is readily compared to clinically meaningful accuracy thresholds.

A complimentary, and occasionally competing, technique for IMRT QA is single field planar dosimetry. Each beam is delivered individually to a phantom, and the dose is measured in a plane orthogonal to the beam direction using film, electronic portal imaging devices (EPIDs), or arrays of diodes or ion chambers.[8–13] While with individual field planar dosimetry, the dose in the treatment field can be measured after a single delivery, the difficulty lies in the interpretation of the measurements. A high dose, low gradient volume in the total plan often becomes an assortment of very complex dose patterns when the plan is deconstructed into individual fields. A method for analysis of complex planar dose distributions, the gamma function,[14] has become a prevalent tool for comparison of calculated and measured dose distributions from single IMRT fields. The challenges in implementation of a gamma index for analysis of IMRT dosimetry lie in selecting gamma analysis parameters and translating a map of gamma values to clinically relevant action thresholds. If gamma is applied to a dose plane within the composite plan, regions that fall out of tolerance can be evaluated in terms of the impact on the tissue being irradiated by that area of the plan as long as the geometry of the phantom is not dramatically different from the shape of the patient. However, if a portion of an IMRT *field* fails the gamma criteria, it may be one of only nine fields to irradiate any part of the patient. Areas of failure within single fields are difficult to correlate from field to field, and the cumulative effect on the plan's accuracy cannot be judged.

A common strategy for analyzing gamma maps from IMRT fields is to calculate the fraction of pixels in each dose plane for which the gamma function returns a value greater than one. In this work the author has developed two collections of IMRT plans: One set which is calculated accurately enough to be used for patient treatment in the author's center and one set which is not. Fields of all treatment plans are measured with two different planar dosimeters and gamma analysis is performed with two tolerances for each dosimeter system. The aim of this study is to assess the sensitivity of single field gamma analysis in detecting realistic, known dosimetric errors.

II. METHOD AND MATERIALS

II.A. Treatment plans and delivery

This study was initiated with three clinical IMRT plans for head and neck cancer, selected at random from clinical practice at the author's institution. Each case was planned with nine fields, delivering a prescription dose of 60 Gy to a large PTV with a concomitant boost carrying a smaller target volume to 70 Gy. The spinal cord was expanded 5 mm radially in all three cases, and the resulting spinal cord planning organ at risk volume was kept below 45 Gy. A number of other normal structures such as parotids, esophagus, brainstem, larynx, and oral cavity were also considered in the optimization process. Extensive QA, to be described in Sec. II B, was performed on each plan at the initiation of this study to confirm suitable dosimetric accuracy over the entire irradiated volume. These three plans, henceforth referred to as "acceptable" plans, represent the control data set and gamma analysis of planar dosimetry for these plans were compared to three plans with known dosimetric deficiencies.

In order to develop IMRT plans with dosimetric inaccuracies, each of the clinical plans was copied and reoptimized with more aggressive dose volume constraints and cost function priorities. In some cases, for example, the maximum cord dose was reduced from 45 to 40 Gy, along with similar reductions to other normal tissue constraints, while the allowed maximum dose to target volumes was also reduced. None of the dosimetric constraints were far outside what may be used clinically. The new plans were more highly modulated than their acceptable counterparts, and doses in highly modulated plans tend to be calculated less accurately by the IMRT planning systems,[5] although *a priori* prediction of dosimetric accuracy is not possible. Examples of moderately and highly modulated fields from one of the plans may be seen in Figs. 3 and 5. Each of these plans was subjected to the same extensive set of ion chamber measurements to verify that the dosimetric accuracy of these plans would deem them unsuitable for clinical application. If the dosimetric accuracy of the new plans was not outside the acceptable range for clinical use in the author's center, the optimization was repeated again with increasingly aggressive dosimetric constraints. Eventually, for each acceptable plan, an unacceptable counterpart was developed which was not suitable for clinical use in the author's center.

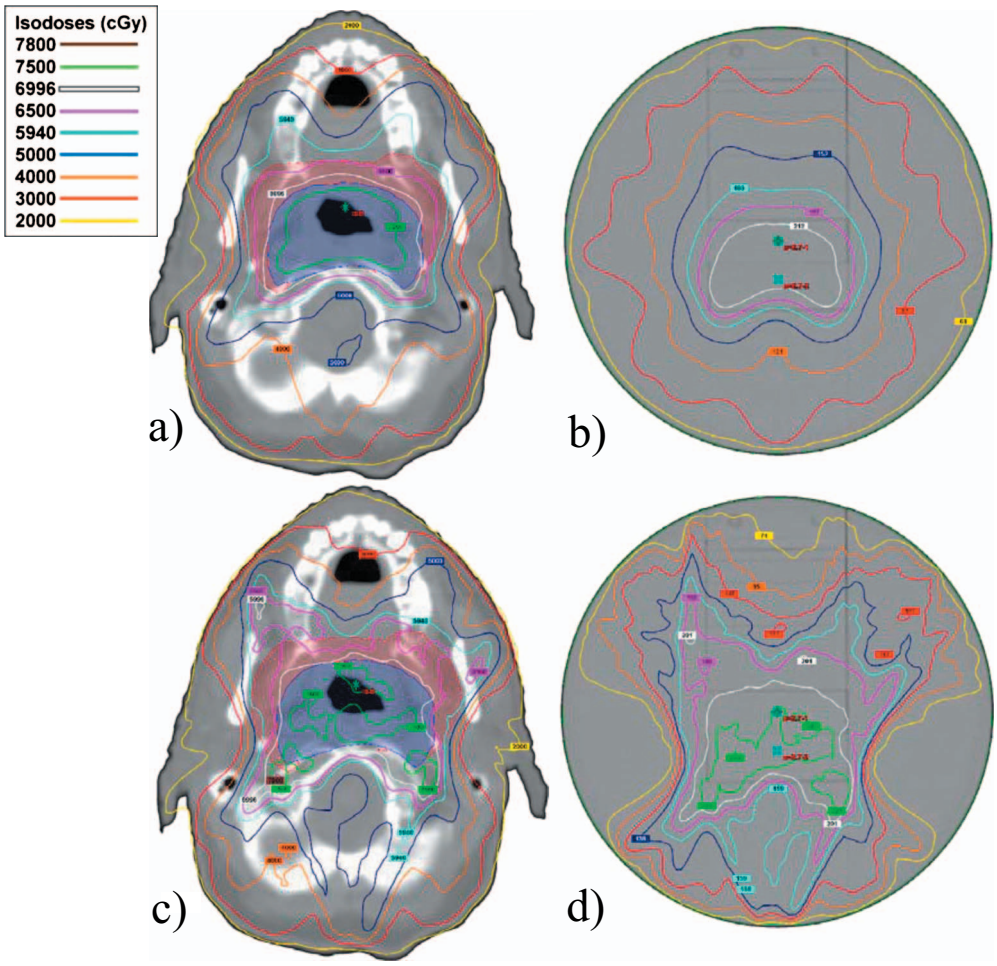

FIG. 1. (a) Axial plane of an IMRT plan that was determined by phantom measurements to be clinically acceptable. (b) Same plane of the IMRT plan, recast on the phantom used for ion chamber measurements to verify clinical acceptability of the plan. (c) Axial plane of an IMRT plan that was determined to be unacceptably inaccurate. (d) Same plane of the unacceptable IMRT plan, recast on the phantom used for ion chamber QA.

Figure 1(a) shows an axial slice of one of the acceptable IMRT plans. Figure 1(c) is the same slice in the unacceptable treatment plan for the same patient. The dose levels in the shaded target volumes are similar for both plans, but the fields in the unacceptable plan are more highly modulated than the accurate one. At the time of this study, the IMRT plans were computed with Varian's pencil beam convolution algorithm. The dose distributions have since been recalculated using the newer analytic anisotropic algorithm and no appreciable difference between the two algorithms was noted for calculations within homogeneous phantoms.

All plans were delivered on a Varian 21EX linac, fitted with a Millennium 120 leaf multileaf collimator.

II.B. Volumetric ion chamber QA

All six plans, three acceptable plans and three unacceptable plans, were cast on a cylindrical phantom for dosimetric verification. The superior end of the phantom has a hemispherical cap and an angular scale is printed on the inferior end to indicate rotation about the superior-inferior axis. A 2 mm diameter, 5 mm long ion chamber (PTW Pinpoint Model T31006, Radiation Products Design, Albertville, MN) is placed in the phantom, and its position can be adjusted radially in millimeter increments. Through a combination of phantom rotation, radial positioning of the ion chamber holder, and longitudinal translation of the phantom, any point in an IMRT plan can be measured with this system.[15] Figure 1(b) shows an axial slice of an acceptable plan cast on the cylindrical phantom and Fig. 1(d) is the same slice of the unacceptable plan for that same patient. Dosimetric accuracy has not been found to be a global property of many IMRT plans and so multiple high dose and low dose points in the cumulative plans were chosen for measurement with the ion chamber. Six, 15, and ten points were measured in high dose regions of plans one, two, and three, respectively. Two to three low dose points were also measured in the spinal cord region for each of the plan pairs. The same points were measured in each pair of acceptable and unacceptable plans, and the points were chosen to be free of dose gradients in the neighborhood of the small volume ion chamber. For an IMRT plan to be considered for clinical application in the author's center, multiple ion chamber measurements are made in low gradient regions of the plan and all must agree to within 4% of the calculation. The extensive ion chamber

measurements performed here, with between eight and 18 points measured in high and low dose, low gradient locations, are meant to establish the suitability of the acceptable plans and to document the inaccuracies of the unacceptable plans.

II.C. Planar EPID dosimetry

One of the techniques for single field IMRT dosimetry studied in this work used the linac-mounted EPID (PortalVision aS500, Varian Medical Systems, Palo Alto, CA) as a planar detector. The EPID detector consists of 512×384 pixels with a resolution of 0.78×0.78 mm^2. An algorithm in the IMRT planning system converted each of the IMRT fields to a portal dose image (PDI) which predicted the response of the imager when exposed to the field. The imager was positioned 5 cm below isocenter and operated in a mode which integrated the response of the device to each field.[12] EPID response was reported in terms of "calibrated units" (CU), such that a 10×10 cm^2 open field of 100 monitor units on the surface of the detector returned a value of 0.907 CU in the center of the integrated image. No further calibration of the imager was required at the time of measurement, but all IMRT irradiations in this study were accompanied by 10×10 cm^2 irradiation of 100 monitor units to validate the constancy of the EPID calibration to within 1% of the reference value. The position of the EPID relative to central axis of the beam was recorded by the imager support arm so that measured PDIs were automatically registered to predicted images. Positional calibration of the imager was verified to within 1 mm at the time of each measurement by analysis of the 10×10 cm^2 image used to validate CU calibration.

In addition to generating PDI prediction images, the IMRT planning system includes software for review of measured PDIs. Gamma comparisons were done between predicted and measured portal dose images with two separate criteria: 2% dose/2 mm distance to agreement (DTA) and 3% dose/3 mm DTA. The EPID analysis software performed the gamma analysis by evaluating the percentage dose difference with respect to the maximum dose in the plane, with no option to derive the relative dose difference with respect to the dose in a specific point, i.e., locally. The analysis was performed with a low dose threshold such that any pixels that received less than 10% of the maximum dose in the image were excluded from the comparison. The 10% threshold excluded points outside the irradiated area of each field, but not low dose regions within each field. For each of the two criteria sets, the percentage of pixels passing the gamma criteria was recorded for each field. Although each of the IMRT plans was designed with nine gantry angles, the limited range of MLC leaves within a carriage position required most of the fields to be split into two subfields for each gantry angle. The EPID system does not sum fields, so each subfield was measured and analyzed separately.

II.D. Ion chamber array dosimetry

The second planar detector used in this study was the MATRIXX array of ion chambers (IBA Dosimetry, Schwarzenbruck, Germany). The active area of the array is 24.4×24.4 cm^2, and it consists of 1020 vented ion chambers. The effective depth of the ion chamber points of measurement is 3 mm, and an additional 5 cm of Solid Water (Gammex, Middleton, WI) was added to the top of the array for buildup. 10 cm of Solid Water below the array contributed backscatter. Fields were delivered individually to the surface of the array with the gantry fixed at zero degrees. Each measurement session was initiated by irradiating the array with a 10×10 cm^2 field of 100 monitor units to verify the calibration and positioning of the MATRIXX to within 1% and 1 mm, respectively.

Reference dose planes were computed in the IMRT planning system by replacing the patient with a block of Solid Water and computing each beam individually on a 2.5 mm dose grid at zero gantry angle. Dose planes were exported from the planning system in DICOM format to the analysis software accompanying the ion chamber array (OmniPro-I'mRT, IBA Dosimetry, Schwarzenbruck, Germany). The dose points measured by the MATRIXX array were linearly interpolated onto a 2.5 mm grid for comparison with the calculated dose planes. As in the EPID procedure, each field measurement was analyzed with 2%/2 mm and 3%/3 mm gamma criteria, with a threshold that excluded from the comparison any chamber receiving less than 10% of the maximum dose in the plane. As opposed to the EPID analysis, the MATRIXX software calculates relative dose deviations locally, with no option to calculate the deviation relative to the maximum dose in the plane. The fraction of pixels passing the gamma analysis for each criterion set was recorded for each beam. The MATRIXX array allows users to sum multiple irradiations into a single dose plane, so in this analysis, individual subfields resulting from multiple carriage positions at a single gantry angle were recombined for comparison to calculated dose planes.

III. RESULTS

III.A. Volumetric ion chamber QA

Figure 2 summarizes the results of ion chamber measurements in low gradient regions of the IMRT plans calculated in a cylindrical phantom. The x-axis in Fig. 2 corresponds to the dose predicted by the planning system for each point, while the y-axis is the deviation between measurement and calculation. Deviation of greater than 4% between a measurement and calculation for any single point is categorized as a failed result and would qualify a plan as unacceptable for clinical application in the author's center. None of the 39 points measured in the acceptable plans failed this criterion; 24 of the 39 points measured in the unacceptable plans differed from the calculation by more than 4%, with a maximum difference of 10.6% between calculated and measured dose. The general trend of inaccuracy in all the IMRT plans was consistent in that the measured doses in the high dose

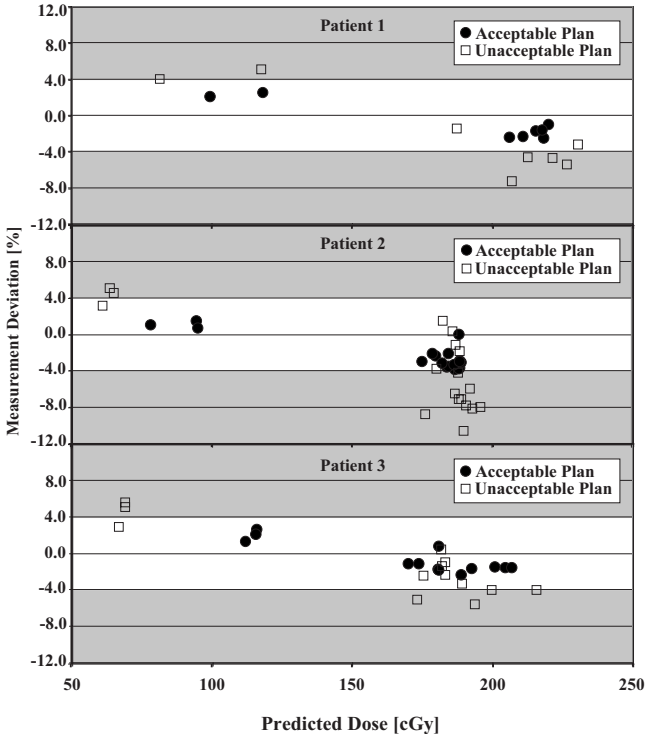

FIG. 2. Results of ion chamber measurements in the cumulative dose distributions for all IMRT plans. The deviation between measured and calculated dose is plotted against the dose calculated by the planning system. Deviations of greater than 4% between measurement and calculation (shaded regions) are qualified as failed results in the author's center.

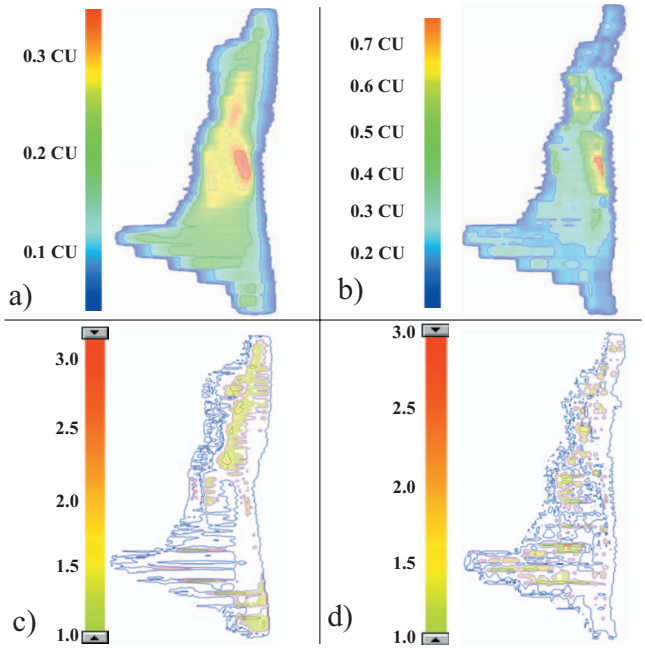

FIG. 3. (a) Single PDI measured by the EPID from plan that was determined by phantom measurements to be clinically acceptable. (b) The same field from an unacceptable plan for the same patient. (c) and (d) are gamma plots from comparison of measurements (a) and (b) to the predicted PDIs. Colored pixels are points that failed a 2%/2 mm gamma analysis.

regions of the plan were lower than predicted, while the measurements in the low dose regions were higher than predicted. While the observed inaccuracies for most clinical plans are small, this represents an inherent trend for IMRT plans generated at the author's center. This trend cannot be eliminated through adjustment of the modeling parameters in the planning system, as they monotonically influence all calculated doses. Accuracy of high dose regions could be improved, for example, but only at the expense of low dose regions in the plan.

III.B. Planar EPID dosimetry

Figure 3 shows a pair of IMRT fields for one of the patient's acceptable and unacceptable plans, as measured with the EPID. Panel A is the measured dose map for one field of the acceptable plan, and below it, in panel C, is the gamma function map comparing it to the portal dose prediction as computed by the software within the treatment planning package. Colored pixels are the ones which failed the 2%/2 mm DTA analysis. On the right, in panels B and D are the analogous displays from the unacceptable plan for that patient. For this pair of fields, 88.8% of the analyzed pixels passed the gamma criteria in the unacceptable plan, while only 85.2% passed in the acceptable plan. Field-by-field gamma analysis scores from the EPID measurements are shown in Fig. 4. The top row of plots shows the percentage of pixels passing the gamma analysis with 2%/2 mm DTA criteria and the bottom row shows the analysis with 3%/3 mm DTA. The plots show two data points per field number, as most of the fields were split into two MLC carriage groups and must be analyzed separately. The exceptions are fields 3 and 7 for patient 3.

Table I summarizes the results of single field gamma analysis for all six plans, performed with the EPID and portal dose prediction algorithm. With 2% dose and 2 mm DTA criteria, the acceptable plans' gamma scores, averaged over all fields, were 91.2%, 88.2%, and 92.8%, while the three unacceptable plans scored 91.9%, 87.5%, and 87.6%. There is no clear distinction in overall gamma score between the acceptable and unacceptable IMRT plans. When the gamma criteria are expanded to 3% and 3 mm DTA, the scores from the acceptable plans range between 96.5% and 98.2%, while the unacceptable plans score between 96.4% and 98.1%.

III.C. Ion chamber array dosimetry

Figure 5 shows a pair of IMRT fields for one of the patient's plans, as measured with the MATRIXX array. Panel A is the measured dose map for one field of the acceptable plan, and below it, in panel C, is the gamma function binary display. Red pixels in panel C are pixels that failed the 2%/2 mm DTA gamma analysis with the MATRIXX array. On the right, in panels B and D, are the analogous displays from the unacceptable plan for that patient. For this pair of fields, 88% of the analyzed pixels pass the gamma criteria in both the acceptable and unacceptable plans. Field-by-field gamma analysis scores from the MATRIXX array measurements are shown in Fig. 6. The top row of plots shows the percentage of pixels passing the gamma analysis with 2%/2 mm DTA

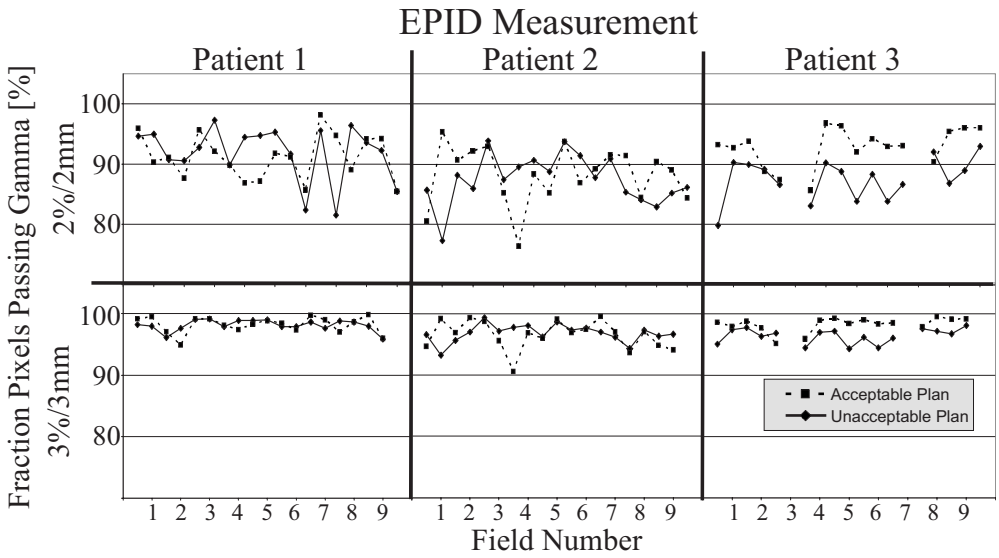

FIG. 4. Field-by-field display of the gamma analysis results for EPID-based QA of six IMRT plans. The fraction of pixels passing the gamma analysis is plotted for each of the fields when using 2%/2 mm DTA criteria (top row) and 3%/3 mm DTA (bottom row). The solid lines connect the measurements of the unacceptable IMRT plans, while the dashed lines connect measurements of the acceptable plans. Missing points in the third column of plots correspond to fields which were delivered with a single MLC carriage position.

criteria and the bottom row shows the analysis with 3%/3 mm DTA.

Table II summarizes the gamma analysis for all six plans, as measured with the MATRIXX ion chamber array. Using gamma criteria of 2% and 2 mm DTA, two of the three acceptable plans have better average scores, but the deviation from field to field is large compared to that difference.

IV. DISCUSSION

Analyzing the fraction of pixels passing gamma analysis in individual beams delivered at gantry, collimator, and couch angles set to 0° is not a sensitive test for detecting dosimetric inaccuracies in IMRT plans. There are cases in the EPID measurements, such as patient 3, in which the average gamma score for the unacceptable plan (87.6%, with 2%/2 mm criteria) is somewhat lower than the score for the corresponding acceptable plan (92.8%). Unfortunately, hope of choosing a global threshold indicating dosimetric inaccuracy in a plan is diminished with the realization that the average gamma score for the unacceptable plan for patient 3 is very close to the score for the acceptable plan for patient 2 (88.2%). Worse yet is the result for patient 1, in which the unacceptable plan has a slightly better average gamma score (91.9% for 2%/2 mm criteria) than the acceptable plan (91.2%).

As in the EPID analysis, the MATRIXX analysis of the unacceptable plan for patient 1 actually returns a somewhat higher average gamma score than its acceptable counterpart. When the gamma criteria are expanded to 3% and 3 mm DTA, the differences between the acceptable and unacceptable plans diminishes, as it did with the EPID measurements. Only one of the three cases, patient 2, has a lower average gamma score for the unacceptable plan (96.9%) than for the acceptable plan (97.6%). For the other two patients, all four plans return average gamma scores over 99% with the 3%/3 mm DTA analysis.

The challenges in analyzing single field planar maps are choosing a quantity to measure and determining a clinically relevant threshold for that quantity. Percentage of points passing a gamma comparison between calculated and measured fields has emerged as a prevalent QA strategy, although there are limited data to support any particular action threshold. Nelms *et al.*[16] conducted a survey of 139 institutions regarding their patient specific IMRT QA thresholds. They

TABLE I. Summary of gamma analysis of IMRT plans measured field-by-field with an EPID. Scores for each plan are the percentage of pixels passing the gamma criteria, averaged over all fields in a plan. The standard deviations of the scores, calculated for all of fields in each plan, are reported in parentheses.

	Average gamma passing percentage: EPID analysis			
	2% dose/2 mm DTA		3% dose/3 mm DTA	
	Acceptable plan	Unacceptable plan	Acceptable plan	Unacceptable plan
Patient 1	91.2 (3.8)	91.9 (4.6)	98.2 (1.4)	98.1 (0.9)
Patient 2	88.2 (4.8)	87.5 (4.0)	96.5 (2.3)	96.8 (1.4)
Patient 3	92.8 (3.3)	87.6 (3.5)	98.2 (1.2)	96.4 (1.2)

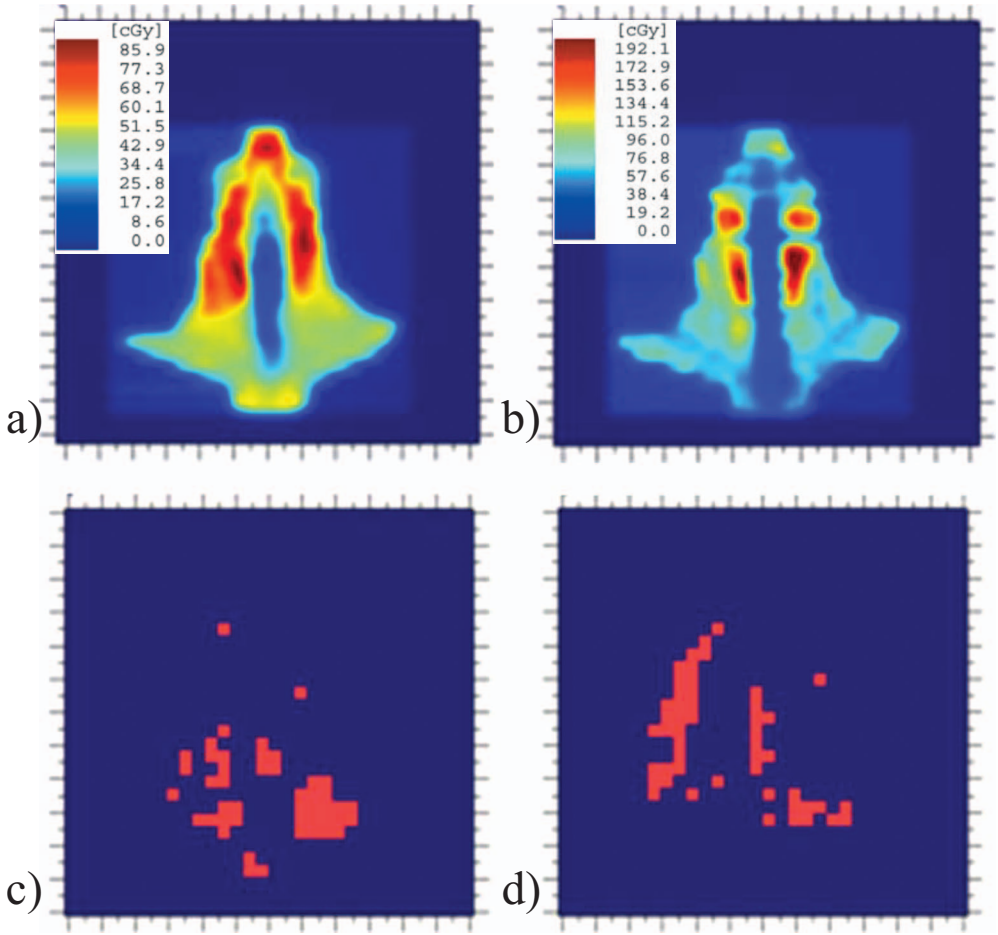

FIG. 5. (a) Single dose plane measured by the MATRIXX ion chamber array for a plan that was determined by phantom measurements to be clinically acceptable. (b) The same field from an unacceptable plan for the same patient. (c) and (d) are binary gamma plots from comparison of measurements (a) and (b) to the calculated dose planes. Red pixels are points that failed a 2%/2 mm gamma analysis.

found that the majority of institutions performed their gamma analyses with 3% dose difference and 3 mm DTA criteria. Among institutions who had established acceptance criteria for IMRT plans, most required 90%–95% of points to pass the gamma analysis. A weakness of the study, acknowledged by the authors, is that planar IMRT QA does not indicate how total patient dose may be affected by any deviations noted in the planar measurements. It should be noted that all six plans examined in the current study exceeded 95% scores for 3%/3 mm gamma analysis with both dosimeters.

Howell *et al.*[17] reported on their efforts to establish action levels for EPID-based QA of IMRT plans. Three parameters were included in their analysis: Maximum gamma, average gamma, and percentage of points with gamma >1. A total of 1152 fields from 152 IMRT plans were evaluated, and statistical distributions of these parameters were presented. The percentage of pixels with gamma >1 varied with plan complexity, but the results for their head and neck plans were similar to those in the present work, with 7.6% of the points failing a 3%/3 mm DTA analysis, averaged over 397 fields. Their study advocates for increased scrutiny of plans in which results are more than 1 or 2 standard deviations from their historical experience, but unfortunately there is no mention of any independent effort to determine whether the plans which comprise that history were calculated accurately by the planning system. Further, the sensitivity of the gamma parameters to plan inaccuracies is not addressed. In a similar study, van Zijtveld *et al.*[18] analyzed EPID measurements of IMRT fields from 75 patients. They examined the sensitivity of gamma as a function of dose and distance to agreement parameters and introduced an automated decision tree for physicist intervention based on a fraction of pixels failing gamma analysis and the area of those regions of failure. Examples of fields failing the gamma analysis due to malfunctioning leaves or delivery of the wrong plan were presented, but the underlying assumption of their work was that plans which pass the gamma analysis are free of dosimetric errors.

The report of AAPM Task Group 119 shares the gamma scores for a number of IMRT plans which were also subject to ion chamber measurement of the cumulative plan.[4] The gamma scores in this manuscript are comparable to that report's average score of 97.9% with 3%/3 mm analysis. The sensitivity of gamma analysis to dosimetric errors is not addressed by Task Group 119, although the report warns against relying solely on per field gamma analysis for detection of dosimetric inaccuracies.

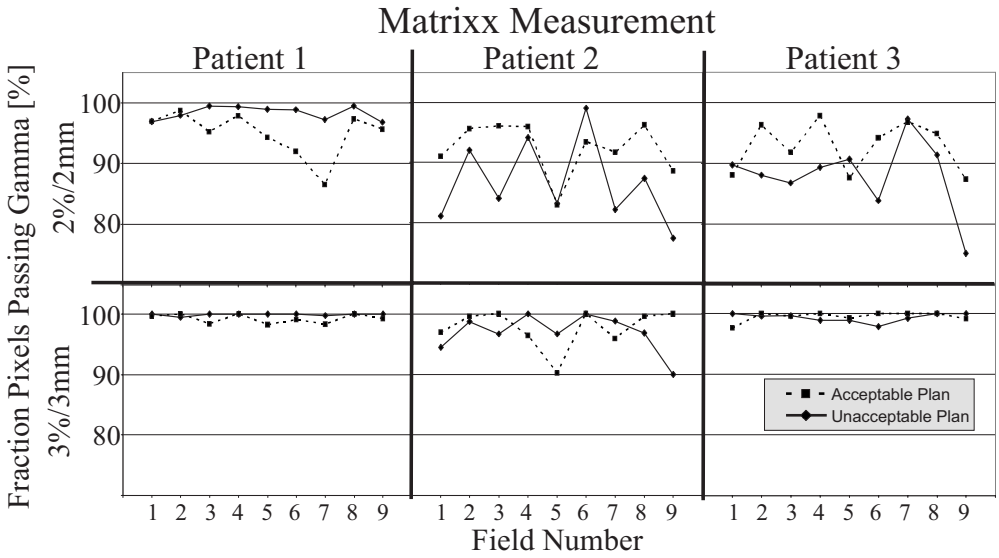

FIG. 6. Field-by-field display of the gamma analysis results for QA of six IMRT plans performed with the MATRIXX ion chamber array. The fraction of pixels passing the gamma analysis is plotted for each of the fields when using 2%/2 mm DTA criteria (top row) and 3%/3 mm DTA (bottom row). The solid lines connect the measurements of the unacceptable IMRT plans, while the dashed lines connect measurements of the acceptable plans.

By comparison to other works in the literature, the current study suffers from a weakness of limited statistics. The unique strength of this study, however, is that it is performed with controls of plans with known, realistic, dosimetric inaccuracies. These inaccuracies, demonstrated by extensive ion chamber measurements of the composite plan, were not manufactured by artificially modifying machine parameters. Instead, the unacceptably inaccurate plans were developed by asking the inverse planning algorithm to create a plan with highly modulated yet realistic fluences in exactly the way a dosimetrist may do for a clinical case. With an ideal QA protocol, the physicist would measure a parameter that readily differentiates acceptable IMRT plans from ones with substandard dosimetric accuracy. Even the small set of plans studied here, three acceptable and three unacceptable, clearly indicates that planar analysis of individual IMRT fields is not sensitive to these errors.

The problem with field-by-field measurement of IMRT plans is that clinically vital regions of uniform dose in the composite plan are often a collection of extremely steep dose gradients when deconstructed into individual fields. Analysis of single fields with extreme gradients necessitates use of distance to agreement, but this blunts the sensitivity of the analysis to true dosimetric errors as well. One potential variable unaddressed by this study is whether the dose deviations within the gamma analysis should be calculated locally, or relative to the maximum dose in a plane. Unfortunately, this could not be studied directly here because each planar system allowed for only one method of gamma calculations: global deviations for the EPID and local for the ion chamber array. Each method of analysis is represented by one of the systems here and there is no apparent difference in sensitivity between the two, although the systems also feature very different spatial resolutions.

Clearly, the gold standard method of plan verification employed here, measurement of the composite plan with an ion chamber can also be somewhat insensitive to dosimetric errors. The data in Fig. 2, results of ion chamber measurements for these plans, demonstrate that dosimetric accuracy is a regional quality of the plan, and not a global phenomenon. Sampling just a single point in patient 2's unacceptable plan, for example, may indicate an accuracy of better than 2%, but not indicate the presence of another region in which the deviation between calculation and measurement is greater than

TABLE II. Summary of gamma analysis of IMRT plans measured field-by-field with a MATRIXX ion chamber array. Scores for each plan are the percentage of pixels passing the gamma criteria, averaged over all fields in a plan. The standard deviations of the scores, calculated for all of fields in each plan, are reported in parentheses.

	Average gamma passing percentage: MATRIXX analysis			
	2% dose/2 mm DTA		3% dose/3 mm DTA	
	Acceptable plan	Unacceptable plan	Acceptable plan	Unacceptable plan
Patient 1	94.9 (3.8)	98.3 (1.1)	99.2 (0.8)	99.9 (0.2)
Patient 2	92.4 (4.5)	86.8 (7.0)	97.6 (3.2)	96.9 (3.2)
Patient 3	92.8 (4.2)	88.0 (6.1)	99.5 (0.8)	99.3 (0.7)

10%. In the author's clinic, ion chamber measurements are complemented by EPID measurements of each field. Isodose and profile overlays from the EPID measurements qualitatively demonstrate the accuracy and integrity of the treatment fields over the entire treatment volume. A qualitative comparison of calculated and measured isodose plots does not guarantee dosimetric accuracy of the entire plan volume, but this work shows that gamma analysis does not provide this information either. The qualitative analysis is meant only to show that there are no gross errors in the delivered fluences far from the points sampled by an ion chamber.

It would appear that for the time being, the accuracy of IMRT planning systems dictates a continued need for patient specific IMRT QA which is sensitive to dosimetric inaccuracies. Promising developments which may deliver viable sensitivity to dosimetric errors over a large volume include the use of measured EPID images to back-calculate a volumetric dose distribution in the patient,[19,20] or placement of a planar detector array in a phantom for measurement of a composite plan.

V. CONCLUSION

By comparing gamma analysis of three acceptable IMRT plans to the same analysis for three unacceptable plans, the goal of this study was to determine a threshold score which could differentiate between good and poor IMRT cases. Instead, this work has shown that while planar dosimetry may comprise one facet of an effective IMRT QA protocol, gamma scores could not reliably identify a plan with poor dosimetric accuracy. Effective patient specific IMRT QA should still include a measurement of composite dose in the complete plan.

[a]Electronic mail: kruse.jon@mayo.edu

[1]G. J. Kutcher, L. Coia, M. Gillin, W. F. Hanson, S. Leibel, R. J. Morton, J. R. Palta, J. A. Purdy, L. E. Reinstein, G. K. Svensson, M. Weller, and L. Wingfield, "Comprehensive QA for radiation oncology: Report of AAPM Radiation Therapy Committee Task Group 40," Med. Phys. **21**, 581–618 (1994).
[2]B. Fraass, K. Doppke, M. Hunt, G. Kutcher, G. Starkschall, R. Stern, and J. Van Dyke, "American Association of Physicists in Medicine Radiation Therapy Committee Task Group 53: Quality assurance for clinical radiotherapy treatment planning," Med. Phys. **25**, 1773–1829 (1998).
[3]J. R. Palta, C. Liu, and J. G. Li, "Quality assurance of intensity-modulated radiation therapy," Int. J. Radiat. Oncol., Biol., Phys. **71**, S108–S112 (2008).
[4]G. A. Ezzell, J. W. Burmeister, N. Dogan, T. J. LoSasso, J. G. Mechalakos, D. Mihailidis, A. Molineu, J. R. Palta, C. R. Ramsey, B. J. Salter, J. Shi, P. Xia, N. J. Yue, and Y. Xiao, "IMRT commissioning: Multiple institution planning and dosimetry comparisons, a report from AAPM Task Group 119," Med. Phys. **36**, 5359–5373 (2009).
[5]L. Dong, J. Antolak, M. Salehpour, K. Forster, L. O'Neill, R. Kendall, and I. Rosen, "Patient-specific point dose measurement for IMRT monitor unit verification," Int. J. Radiat. Oncol., Biol., Phys. **56**, 867–877 (2003).
[6]S. L. Breen, D. J. Moseley, B. Zhang, and M. B. Sharpe, "Statistical process control for IMRT dosimetric verification," Med. Phys. **35**, 4417–4425 (2008).
[7]A. Molineu, J. Lowenstein, N. Hernandez, P. Alvarez, D. Followill, and G. Ibbott, "Has IMRT delivery improved in the last 5 years?," Med. Phys. **35**, 2762–2762 (2008).
[8]C. Martens, I. Claeys, C. De Wagter, and W. De Neve, "The value of radiographic film for the characterization of intensity-modulated beams," Phys. Med. Biol. **47**, 2221–2234 (2002).
[9]B. Warkentin, S. Steciw, S. Rathee, and B. G. Fallone, "Dosimetric IMRT verification with a flat-panel EPID," Med. Phys. **30**, 3143–3155 (2003).
[10]M. Bucciolini, F. B. Buonamici, and M. Casati, "Verification of IMRT fields by film dosimetry," Med. Phys. **31**, 161–168 (2004).
[11]D. Letourneau, M. Gulam, D. Yan, M. Oldham, and J. W. Wong, "Evaluation of a 2D diode array for IMRT quality assurance," Radiother. Oncol. **70**, 199–206 (2004).
[12]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).
[13]G. J. Budgell, Q. Zhang, R. J. Trouncer, and R. I. Mackay, "Improving IMRT quality control efficiency using an amorphous silicon electronic portal imager," Med. Phys. **32**, 3267–3278 (2005).
[14]D. A. Low, W. B. Harms, S. Mutic, and J. A. Purdy, "A technique for the quantitative evaluation of dose distributions," Med. Phys. **25**, 656–661 (1998).
[15]C. M. Ma, S. B. Jiang, T. Pawlicki, Y. Chen, J. S. Li, J. Deng, and A. L. Boyer, "A quality assurance phantom for IMRT dose verification," Phys. Med. Biol. **48**, 561–572 (2003).
[16]B. E. Nelms and J. A. Simon, "A survey on planar IMRT QA analysis," J. Appl. Clin. Med. Phys. **8**, 2448–2462 (2007).
[17]R. M. Howell, I. P. Smith, and C. S. Jarrio, "Establishing action levels for EPID-based QA for IMRT," J. Appl. Clin. Med. Phys. **9**, 2721–2730 (2008).
[18]M. van Zijtveld, M. L. P. Dirkx, H. C. J. de Boer, and B. J. M. Heijmen, "Dosimetric pre-treatment verification of IMRT using an EPID; clinical experience," Radiother. Oncol. **81**, 168–175 (2006).
[19]M. van Zijtveld, M. L. P. Dirkx, H. C. J. de Boer, and B. J. M. Heijmen, "3D dose reconstruction for clinical evaluation of IMRT pretreatment verification with an EPID," Radiother. Oncol. **82**, 201–207 (2007).
[20]S. Steciw, B. Warkentin, S. Rathee, and B. G. Fallone, "Three-dimensional IMRT verification with a flat-panel EPID," Med. Phys. **32**, 600–612 (2005).