

A new metric for assessing IMRT modulation complexity and plan deliverability

Andrea L. McNiven, Michael B. Sharpe, and Thomas G. Purdie

Citation: *Medical Physics* **37**, 505 (2010); doi: 10.1118/1.3276775

View online: <http://dx.doi.org/10.1118/1.3276775>

View Table of Contents: <http://scitation.aip.org/content/aapm/journal/medphys/37/2?ver=pdfcov>

Published by the American Association of Physicists in Medicine

Articles you may be interested in

The impact of leaf width and plan complexity on DMLC tracking of prostate intensity modulated arc therapy
Med. Phys. **40**, 111717 (2013); 10.1118/1.4824434


Impact of curved surface for clinical plan verification in intensity modulated radiation therapy using 2d array I'mRT MatriXX
AIP Conf. Proc. **1454**, 61 (2012); 10.1063/1.4730688

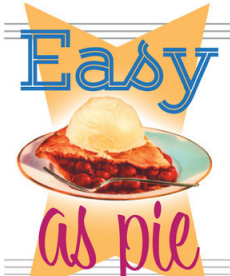
Fractal analysis for assessing the level of modulation of IMRT fields
Med. Phys. **38**, 5385 (2011); 10.1118/1.3633912

Assessing software upgrades, plan properties and patient geometry using intensity modulated radiation therapy (IMRT) complexity metrics
Med. Phys. **38**, 2027 (2011); 10.1118/1.3562897

A new smoothing procedure to reduce delivery segments for static MLC-based IMRT planning
Med. Phys. **31**, 1158 (2004); 10.1118/1.1713279

RadImage is a registered trademark of RadImage, Inc.

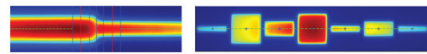




Easy
as pie

RITG148⁺


Custom Designed
TG-148 Tests
For Tomotherapy QA



RIT is your only source for the tests specified for helical tomotherapy in the TG-148 report. These automated QA tests include:

• Automated QA testing	• MLC alignment test
• Y-jaw divergence/beam centering	• Couch translation/gantry rotation
• Y-jaw/gantry rotation plane alignment	• Laser localization
• Gantry angle consistency	• Image quality tests (Cheese Phantom)
• Treatment field centering	• Built in trending and reporting with RITrend

These tests are included in both our RITComplete, and RITG148+ products.



Call 719.590.1077,
option 4, or email
mac@radimage.com
today to set up your
personal demo.

A new metric for assessing IMRT modulation complexity and plan deliverability

Andrea L. McNiven,^{a)} Michael B. Sharpe, and Thomas G. Purdie

Radiation Medicine Program, Princess Margaret Hospital, University Health Network, Toronto, Ontario M5G 2M9, Canada and The Department of Radiation Oncology, University of Toronto, Toronto, Ontario M5G 2M9, Canada

(Received 19 June 2009; revised 19 October 2009; accepted for publication 23 November 2009; published 12 January 2010)

Purpose: To evaluate the utility of a new complexity metric, the modulation complexity score (MCS), in the treatment planning and quality assurance processes and to evaluate the relationship of the metric with deliverability.

Methods: A multisite (breast, rectum, prostate, prostate bed, lung, and head and neck) and site-specific (lung) dosimetric evaluation has been completed. The MCS was calculated for each beam and the overall treatment plan. A 2D diode array (MapCHECKTM, Sun Nuclear, Melbourne, FL) was used to acquire measurements for each beam. The measured and planned dose (PINNACLE³, Phillips, Madison, WI) was evaluated using different percent differences and distance to agreement (DTA) criteria (3%/3 mm and 2%/1 mm) and the relationship between the dosimetric results and complexity (as measured by the MCS or simple beam parameters) assessed.

Results: For the multisite analysis (243 plans total), the mean MCS scores for each treatment site were breast (0.92), rectum (0.858), prostate (0.837), prostate bed (0.652), lung (0.631), and head and neck (0.356). The MCS allowed for compilation of treatment site-specific statistics, which is useful for comparing different techniques, as well as for comparison of individual treatment plans with the typical complexity levels. For the six plans selected for dosimetry, the average diode percent pass rate was 98.7% (minimum of 96%) for 3%/3 mm evaluation criteria. The average difference in absolute dose measurement between the planned and measured dose was 1.7 cGy. The detailed lung analysis also showed excellent agreement between the measured and planned dose, as all beams had a diode percentage pass rate for 3%/3 mm criteria of greater than 95.9%, with an average pass rate of 99.0%. The average absolute maximum dose difference for the lung plans was 0.7 cGy. There was no direct correlation between the MCS and simple beam parameters which could be used as a surrogate for complexity level (i.e., number of segments or MU). An evaluation criterion of 2%/1 mm reliably allowed for the identification of beams that are dosimetrically robust. In this study we defined a robust beam or plan as one that maintained a diode percentage pass rate greater than 90% at 2%/1 mm, indicating delivery that was deemed accurate when compared to the planned dose, even under stricter evaluation criterion. MCS and MU threshold criteria were determined by defining a required specificity of 1.0. A MCS threshold of 0.8 allowed for identification of robust deliverability with a sensitivity of 0.36. In contrast, MU had a lower sensitivity of 0.23 for a threshold of 50 MU.

Conclusions: The MCS allows for a quantitative assessment of plan complexity, on a fixed scale, that can be applied to all treatment sites and can provide more information related to dose delivery than simple beam parameters. This could prove useful throughout the entire treatment planning and QA process. © 2010 American Association of Physicists in Medicine. [DOI: [10.1118/1.3276775](https://doi.org/10.1118/1.3276775)]

Key words: IMRT, quality assurance, complexity, deliverability, treatment planning

I. INTRODUCTION

The evolution of radiation therapy from conventional 3D-conformal treatment plans to intensity modulated radiation therapy (IMRT) has impacted patient treatment and the clinical process in many ways. IMRT can provide several advantages compared to conformal treatments, including an improvement in outcome,^{1,2} the potential for escalating the dose to the target,¹ and a decrease in toxicities.^{3,4} These improvements, however, are also often associated with an increase in treatment complexity, as compared to conformal treatment plans. The increase in complexity affects many steps of the

treatment process,^{5–8} including treatment planning, the quality assurance workload, and the treatment delivery itself. A comprehensive QA program includes consideration of all aspects of the treatment process, including QA of the linear accelerator, any image guidance tools utilized, the treatment planning system, data transfer and record and verify systems, as well as patient-specific IMRT measurements. The implementation of the recommendations, particularly with respect to patient-specific measurements, varies greatly from one institution to another.

In terms of a IMRT treatment plan, complexity has been

described previously as “the frequency and amplitude of fluctuations in the intensity distribution of a beam,”⁹ while others have defined increasing complexity simply by an increased number of monitor units.¹⁰ In general, a high degree of complexity is typically associated with a large number of monitor units (MUs), a large number of segments (for step and shoot IMRT), small segment size, and complex segment shape. Intuitively, all of these aspects could contribute to a more complex fluence map for a single beam. A high degree of complexity is not necessarily a negative feature of a treatment plan, as it may be required due to the geometry and location of the target and organs at risk, as it has been established that there is a tradeoff between complexity and treatment quality,¹⁰ with respect to the achievement of planning objectives.

Webb¹¹ introduced the modulation index (MI) with the intent of defining the degree of complexity of fluence maps. The MI quantifies the variation between adjacent bixel intensities, with the main purpose of using the MI to describe the varying degrees of conformality that can be achieved in planning. In this vein, most work with respect to characterizing IMRT complexity has focused on the complexity of the fluence map. Several different approaches to optimization have been introduced to decrease the complexity of the fluence map.^{10,12–18} This can include limiting the number of MU,^{10,12} number of segments,^{14,17} or imposing limits^{15,18,19} on the extent of modulation allowed during the treatment optimization process. The analysis of these techniques often involves an assessment of plan quality, in terms of the ability to meet specific goals for target coverage or organ at risk dose constraints, by evaluating the dose volume histogram (DVH).

In addition to the tradeoff between IMRT complexity and plan quality (i.e., target coverage, critical structure sparing), another tradeoff that should be considered during the treatment planning process is that between IMRT complexity and the dosimetric accuracy of the treatment plan. The deliverability of the IMRT plan refers to the technical feasibility of delivery (no undeliverable segments). In this work we extend the definition of deliverability to describe the ability to deliver a dose that accurately represents the dose calculated by the treatment planning system (TPS). Dosimetric accuracy can be affected by MLC position accuracy and linear accelerator performance. The dosimetric impact of errors in leaf position increases as the separation between opposing leaves decreases,²⁰ a situation that can be associated with smaller and more irregularly shaped segments. An important step in IMRT QA is the characterization of the linac performance (MU, flatness, symmetry) at small MU.⁵ In certain situations dosimetric accuracy (intended as compared to delivered dose) can be affected due to a phenomenon described as the overshoot effect.^{21,22} This is related to small MU segments, which can be associated with more complex treatment plans. Previous work has also been completed to investigate the effect of increasing the degree of plan complexity, as defined based on the fluence map, on the delivered dose. Mohan *et al.*⁹ showed that filtered fluence maps (lower total number of MU) resulted in a better agreement between planned and measured doses for a sliding window technique. Giorgia *et*

*al.*²³ found that the complexity of the fluence pattern (as assessed using the MI) was related to delivery accuracy, as they found that the agreement between planned and delivered doses decreased as complexity increased. The quality of the dosimetric results (agreement between planned and measured dose) depended on the dose calculation algorithm. This potential tradeoff between increased complexity and dosimetric accuracy can be complicated further if the increased treatment times due to a larger number of MU and the interplay between intrafraction motion and leaf motion are considered.

A single metric that could be integrated into the treatment planning process to assist in both the treatment planning process as well as the post-planning QA process would, therefore, be of interest. The ability to define a level of complexity, that is, typical for a treatment site could assist in both the treatment planning process, by providing a numerical target for assessing the tradeoff between achieving dosimetric objectives and limiting the delivery time. A well-defined or limited range of scores could also allow for easy comparison between treatment sites, which is not possible using existing complexity measures. Much work has been completed to try to improve the QA process by improving efficiency or reducing the extensive QA workload that is recommended for IMRT delivery,^{5–8} specifically in the area of patient-specific QA. This includes the investigation of 2D arrays,^{22,24–29} which can help to improve the efficiency of pretreatment patient-specific dose measurements. New *in vivo* dosimetry techniques³⁰ have also been proposed to eliminate the need for measurements prior to the commencement of treatment. Different pretreatment filtration methods^{23,31,32} could allow for the identification of gross errors prior to measurement, eliminate the need for certain measurements by establishing thresholds for which good results are statistically expected, or by improving the workflow by more quickly identifying potential errors. A complexity metric that was related to deliverability could potentially allow for a reduction in the quantity of patient-specific QA measurements, or when used as a filtering tool, a score could identify plans that require specific attention due to a score outside of a certain threshold.

The modulation complexity score (MCS), a single metric for assessing beam and overall plan complexity and deliverability, has been developed with this intention. The goal of this work is to evaluate the utility of this tool, as compared to typical parameters used to infer degree of complexity (e.g., number of MU). First, the evaluation of numerical data and the potential utility in the treatment planning process will be investigated. Second, the ability of the MCS to provide information about the accuracy of dose delivery will also be assessed in a site-specific context.

II. METHODS AND MATERIALS

II.A. Calculation of the modulation complexity score

The MCS incorporates information from the treatment planning system, such as variability in leaf positions, degree of irregularity in field shape, segment weight, and area into a

TABLE I. Site summary of clinical treatment plans for which the MCS and associated scores were calculated using the QA tool.

Site	Number of Plans	MCS	
		Average	Standard deviation
Breast	48	0.920	0.015
Rectum	28	0.858	0.016
Prostate	49	0.837	0.030
Prostate bed	43	0.652	0.057
Lung	42	0.631	0.133
Head and neck	33	0.356	0.104

single score, ranging from 0 to 1.0, that can be applied to any IMRT plan. An open rectangular field has “zero” complexity and would have a MCS of 1.0. Two parameters, the leaf sequence variability (LSV) and the aperture area variability (AAV), are calculated and combined to calculate the MCS (Appendix). In basic terms, the score decreases from a value of 1.0 (defined at a single rectangular open field) through the addition of segments that are smaller and more irregularly shaped compared to the maximum beam aperture.

II.B. MCS evaluation: Multiple treatment sites

Different clinical treatment sites have an inherent difference in the level of complexity that would be required to create a clinically acceptable plan, based on the differences in typical target shape, size, and location with respect to critical structures. For the purpose of an initial evaluation of the MCS, several clinical treatment sites have been chosen that represent a range of complexity: breast, rectum, prostate, postprostatectomy radiation, lung, and head and neck. All of the IMRT plans investigated in this study were delivered in the clinic, having passed oncology and physics QA processes, and are step and shoot IMRT. The MCS scores (per beam and per treatment plan) were calculated for 243 treatment plans for these six sites (Table I). Treatment planning was completed using PINNACLE³ treatment planning system, Version 8.0h (Phillips, Madison, WI), and dose was calculated using the adaptive convolve algorithm. Plans included in the statistics encompassed the three types of linear accelerators used for delivery at our institution and the corresponding beam models (Varian with 120-leaf MLC, Varian Medical Systems, Palo Alto, CA and Synergy and Synergy S, Elekta Oncology Systems, Crawley, UK). One plan from

each site was chosen for dosimetric evaluation (Table II). All chosen plans were planned for and delivered using either 6 or 18 MV photon beam (Synergy, Elekta Oncology Systems, Crawley, UK). One linear accelerator and beam model was chosen to help limit the variables in the analysis. The plans were chosen to be as close to the site average complexity as possible, given the choice to choose treatment plans corresponding to a single accelerator type.

II.C. MCS evaluation: Lung

Radical lung IMRT plans were chosen to evaluate the MCS metric for a single site analysis, as lung treatments vary greatly in complexity. Additionally, lung is a treatment site for which less complex plans may be desirable in order to maximize the robustness of the treatment with respect to variations in breathing motion, to minimize the effect of interplay between motion and leaf motion for the delivery of the IMRT segments. Forty two IMRT lung plans, that were generated for and subsequently used for clinical purposes, have been evaluated using the MCS tool. Thirteen of these plans, that covered the range of MCS, were chosen for this study. These plans also cover a large range of the typical and subjective measures of complexity such as the number of beams, MU, and segments. All of the selected plans were planned to deliver a prescription of 66 Gy in 33 fractions using a 6 MV photon beam (Synergy, Elekta Oncology Systems, Crawley, UK).

II.D. Dose measurements

2D dose measurements were completed using a MapCHECKTM 2D diode array^{26,27} (Sun Nuclear, Melbourne, FL), which has a diode resolution of 1 cm (horizontal and vertical) and 0.7 cm (diagonal) in the central 10 × 10 cm² area and 2 and 1.4 cm resolutions in the outer area. The diode array was placed and leveled on the treatment couch (no backscatter was used). Dose calibration of the array was completed the same day as the measurements, following the MapCHECKTM procedure, using a 10 × 10 cm² field and a known dose. The array was centered (using the crosshairs and confirming alignment with a 10 × 10 cm² field). For the IMRT measurements, the beams were all delivered, one beam at a time, at 0° gantry angle, with the array setup at 92 cm SSD and 5 cm water-equivalent depth (3 cm solid water plus inherent array buildup).

Measurements were compared to planar dose maps ex-

TABLE II. Summary of plan characteristics for plans included in dosimetric evaluation.

Site	Prescription dose	Number of beams	Number of segments	MU per beam (range)	Total MU	MU per segment	MCS per beam (range)	Plan MCS
Breast	42.4 Gy/16 fractions	2	7	175–189	364	52.0	0.896–0.922	0.909
Rectum	50 Gy/25 fractions	3	11	98–120	322	29.3	0.808–0.840	0.823
Prostate	78 Gy/39 fractions	7	68	52–64	396	5.8	0.692–0.762	0.739
Prostate Bed	66 Gy/33 fractions	7	59	68–105	601	10.2	0.455–0.691	0.580
Lung	66 Gy/33 fractions	5	32	77–120	463	14.5	0.524–0.742	0.645
Head and Neck	70 Gy/35 fractions	9	94	80–159	1051	11.2	0.019–0.319	0.165

ported from PINNACLE³ based on different dose percentage agreement and distance to agreement (DTA) criteria (3%/3 mm and 2%/1 mm). The planar dose maps were calculated with 0.1 cm resolution in the planar dose tool in PINNACLE³ and exported for import into the MAPCHECKTM software for analysis. The ranges of percentage agreement and DTA criteria were chosen to include values documented in the literature for clinical use,^{24,29} as well as more stringent criteria to more strictly evaluate the MCS score. Analysis was completed on a per beam basis using absolute dose, using the treatment planning system calculated planar dose map as the reference, and a 5% threshold was used in the analysis (points with dose less than the 5% isodose line are not included in analysis). A diode passes the evaluation criteria either if the percentage dose difference, as compared to the planned dose, meets the criteria or if the point meets the DTA criterion. To analyze the plan as a whole, a weighted average of all of the beams was used. This was calculated using the diode percentage pass rate for each beam, weighted by the number of MU per beam relative to the total number of MU for the treatment plan. This weighting method is the same as that used to calculate the plan MCS by the QA tool (Appendix).

Additional measurements were completed that helped to assess the sensitivity of the results to the resolution of the diode measurements. The effect of increased dosimetric sampling (achieved through combining multiple MAPCHECKTM measurements) on the diode percentage pass rates was analyzed for two of the lung plans. Two plans were chosen, one of higher and one of lower complexity, and measurements were acquired five separate times: once with the diode array centered, and four more with the array shifted 5 mm in the left, right, superior, and inferior direction. Results could then be combined to achieve a higher density of sampling for the measurements. In the central part of the MAPCHECKTM array, this effectively increases the resolution from 1 cm in the horizontal and vertical directions to 0.5 cm. In the outer array, nonuniform sampling is achieved, with a maximum distance between data points in the horizontal and vertical directions of 1 cm, instead of the original 2 cm diode spacing.

III. RESULTS

III.A. MCS evaluation: Multiple sites

The analysis of many clinical treatment plans allows for the accumulation of treatment site and technique-specific statistics (Table I). The MCS scores increase as the inherent complexity of the treatment (shape and location of the target and its proximity to critical structures) decreases. The standard deviation of the scores also provides insight to the standardization of a technique, as the breast tangent treatment has a very tight standard deviation with respect to the MCS. The MCS tool then allows for the comparison of the MCS of the current plan (green vertical line) with all plans of that type evaluated previously (Fig. 1).

Six different IMRT plans were evaluated in the multiple site analysis. The mean absolute dose difference (between the maximum measured dose and the corresponding point in

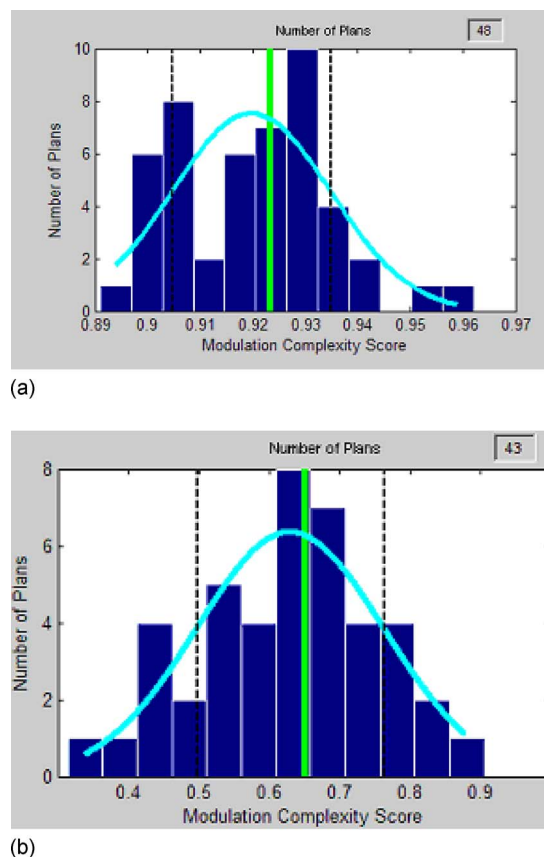


FIG. 1. Illustration of the comparison of a single plan's MCS score (thinner vertical line) compared to all plans analyzed by the MCS QA tool for (a) two-field breast tangent treatments and (b) radical lung treatments. The differences in scale for the MCS highlight the variability in standardization of the plans between the two treatment sites.

the planned dose map) was 1.7 cGy, with a standard deviation of 1.3 cGy. Using 3%/3 mm evaluation criteria every beam had a diode percentage pass rate greater than 96.0%, with an average of 98.7% and 1.0% standard deviation. Figure 2 illustrates the relationship between the percentage pass rate for the 3%/3 mm criteria with MCS and the number of MU. Figures 2(a) and 2(b) include the results for each individual beam. MCS allows for the separation of the different plan types [Fig. 2(a)]. Head and neck, which is a site often identified as high in complexity, has low scores. In contrast, breast and rectum, which are much simpler in terms of complexity (consisting of a large aperture, delivering most of the dose with a few smaller, low MU segments), had the highest scores. No such separation between plan types is evident using MU as a measure of complexity [Fig. 2(b)]. It would be possible to compare total plans using MCS on the same scale. There is no correlation, across multiple treatment sites, between dosimetric results (for either 3%/3 mm or 2%/1 mm criteria) and the MCS or number of MU.

III.B. MCS evaluation: Radical lung IMRT plans

Evaluation of the radical lung IMRT plans using the 2D diode array resulted in a mean maximum absolute dose difference of 0.8 cGy with a standard deviation of 0.7 cGy. The

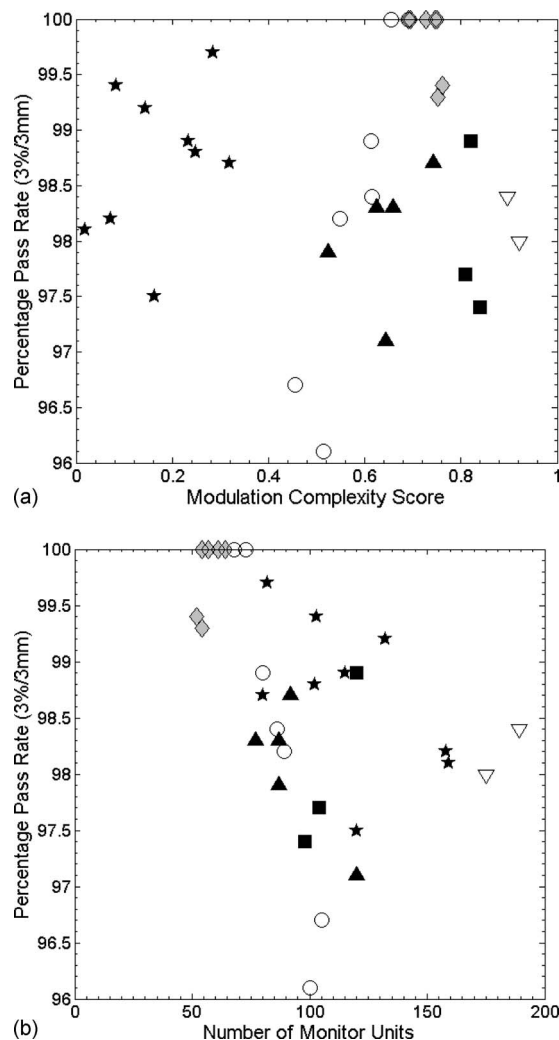


FIG. 2. Evaluation of the percentage of diodes passing the 3%/3 mm evaluation criterion for multiple treatment sites for each beam as a function of (a) modulation complexity score and (b) the number of MU for breast (▼), rectum (■), prostate (◆), prostate bed (▲), lung (●), and head and neck (star).

average percentage pass rate was 99.0% (per beam) for the 3%/3 mm evaluation criterion, with a standard deviation of 1.0%.

Figure 3 illustrates the relationship between the MCS and several beam parameters that have been used previously to characterize complexity (number of MU and the number of beam segments) for the lung plans evaluated. There is limited or no correlation between the complexity score and the number of MU and the number of segments, or the average number of MU per segment, as indicated by correlation coefficient values of 0.41 and 0.18, respectively.

The percentage of diodes meeting the evaluation criteria of 3%/3 mm and 2%/1 mm as a function of MCS and MU is shown in Fig. 4. Figures 4(a) and 4(b) illustrate the excellent agreement between the measured and planned doses for all investigated fields. The use of a more stringent evaluation criteria (2%/1 mm) allows for a greater range of percentage pass rates for further evaluation and for the identification of dosimetrically robust beams or plans, for which the percent-

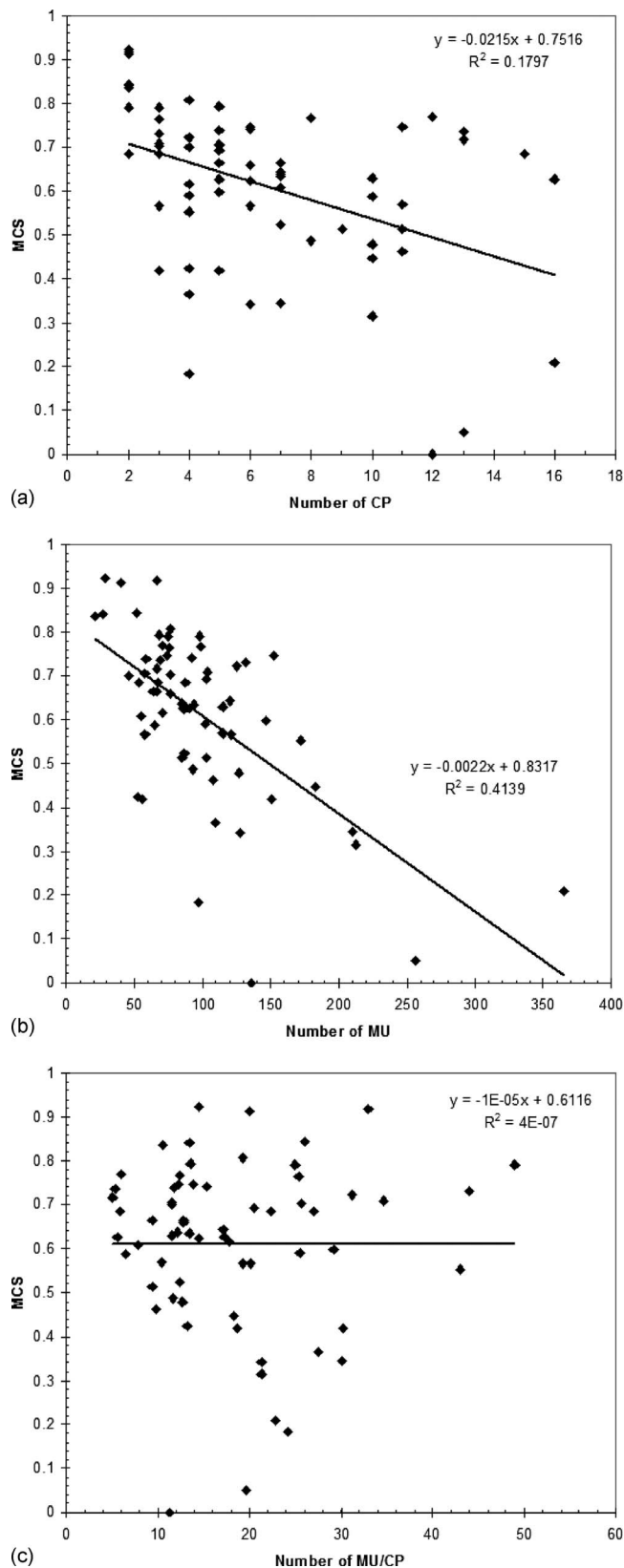


FIG. 3. Modulation complexity scores for each beam for the 13 lung IMRT plans evaluated in this study, as a function of the number of (a) beam segments, (b) MU, and (c) MU per segment. The lack of a strong correlation between MCS and simple beam parameters indicates that the incorporation of multiple factors into a single score can provide different information for the treatment planning and QA processes than the individual beam parameters.

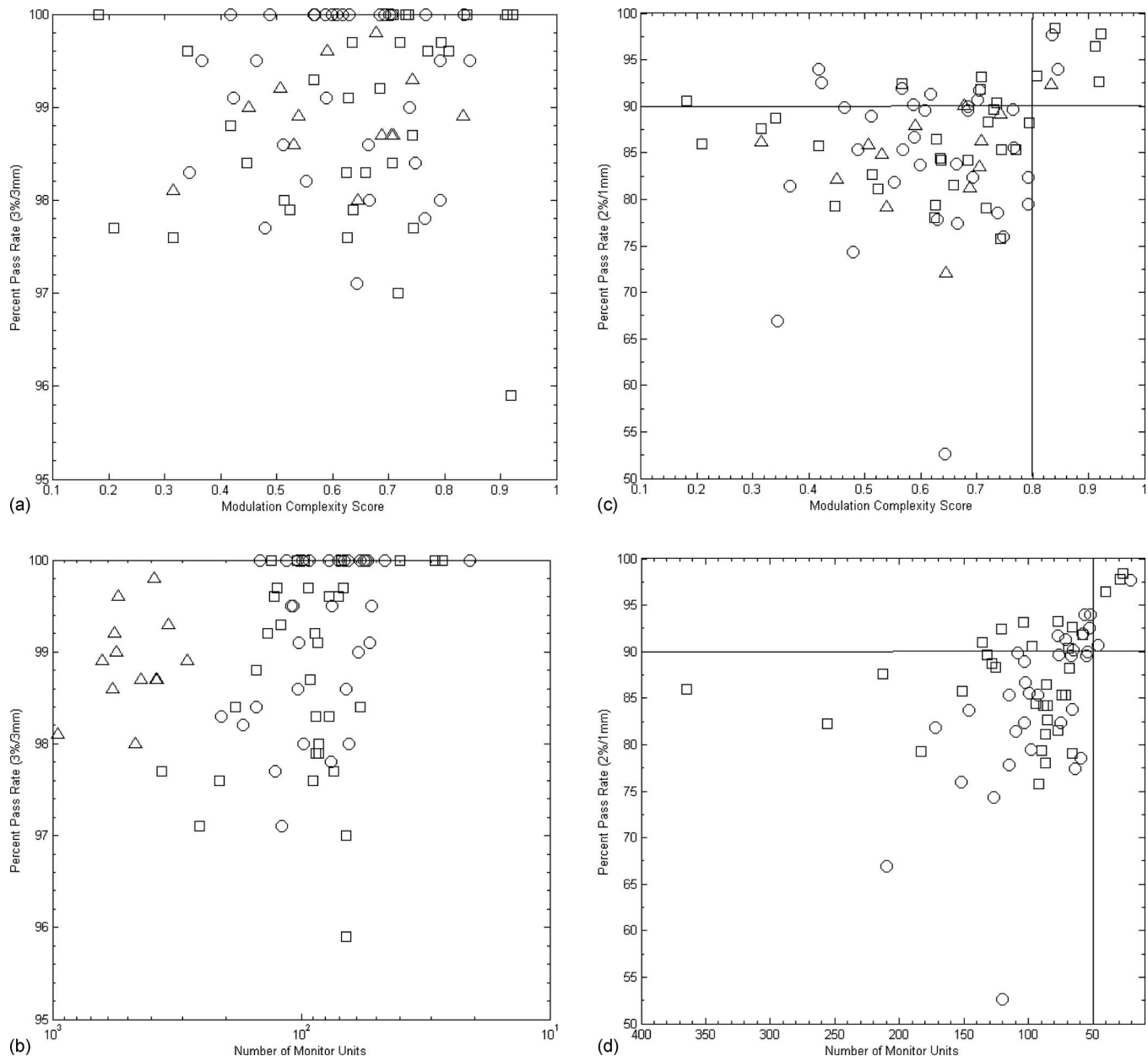


FIG. 4. The percentage of diodes that met the evaluation criteria for absolute dose comparison between planned and measured doses for percentage dose difference and distance to agreement criteria of 3%/3 mm [(a) and (b)] and 2%/1 mm [(c) and (d)]. Results are shown for total plans (triangles), and individual beams with field size, as evaluated by the number of diodes used in the dosimetric analysis, in the top half of beams (squares) and in the lower half (circles). Total plans are not included in (d), but the treatment plan identified as robust in (c), was not identified by number of MU. Each beam and overall treatment plan was assessed for complexity using the MCS and the number of MU. Robust delivery requirements of >90% pass rate were used in conjunction with a 1.0 specificity requirement to determine MCS (>0.8) and MU (<50) thresholds for receiver-operator curve analysis for sensitivity [quadrants illustrated in (c) and (d)].

age pass rate decreases less than other beams [Figs. 4(c) and 4(d)]. For both MCS and MU there is a separation of some data points [upper right quadrants of Figs. 4(c) and 4(d)] representing beam with sustained high pass rates. This indicates the potential for using the complexity metric to identify dosimetrically robust beams, or in other words, a beam for which good dosimetric results could be expected.

Thresholds were then identified for further analysis of these results. For this work, a dosimetrically robust delivery was considered to be one for which the pass rate remained

larger than 90% for a 2% dose difference and 1% DTA criterion. Beams or plans with a MCS greater than 0.8 or less than 50 MU were identified as being able to accurately identify dosimetrically robust plans all of the time (100% specificity). In other words, all beams with a metric meeting the threshold criteria were dosimetrically robust (pass rate larger than 90%, using 2%/1 mm criteria). These thresholds are depicted as the lines separating Figs. 4(c) and 4(d) into quadrants. These quadrants were then used in receiver-operator curve (ROC) analysis to identify the ability of a single metric

TABLE III. Comparison of the ability of the MCS or the number of MU to identify dosimetrically robust beams for the 2%/1 mm dose difference and DTA evaluation. The percentage pass rate criteria of 90% was chosen as a threshold to indicate robust delivery. The 86% pass rate was the median pass rate for the 2%/1 mm evaluation.

% difference and DTA criterion	Percentage pass rate criteria	Complexity evaluator	Threshold value	Sensitivity	Specificity	Positive predictive value	Negative predictive value
2%/1 mm	90%	MU	<50	0.23	1.00	1.00	0.77
		MCS	>0.8	0.36	1.00	1.00	0.80
	86%	MU	<55	0.23	1.00	1.00	0.57
		MCS	>0.8	0.21	1.00	1.00	0.56
	86%	MU	<97	0.44	0.43	0.43	0.44
		MCS	>0.643	0.56	0.58	0.56	0.58
	90%	MU	<71	0.64	0.84	0.61	0.86
		MCS	>0.721	0.41	0.77	0.41	0.77

(the MCS or the number of MU) to accurately identify IMRT treatment beams that will be accurately delivered at the treatment unit. Sensitivity was improved using the MCS metric (0.36) as compared to the sensitivity achieved using the number of MU as the complexity metric (0.23). The negative predictive values, which describes the ability to accurately identify the proportion of beams that are not dosimetrically robust according to our definition, were 0.80 for MCS and 0.77 for MU, showing very little difference between MU and MCS. Table III contains the results for the evaluation of MCS and MU for various evaluation criteria and thresholds. The threshold percentage pass rate was decreased to 86% (the median pass rate for 2%/1 mm), in conjunction with median MCS or MU values or 100% specificity.

Increased sampling was completed for two lung plans by shifting the diode array and combining results. For the plan as a whole, there was a negligible decrease in the diode percentage pass rate for evaluation using the 3%/3 mm evaluation criteria of 0.1% for each plan. The difference between the pass rates for high and low density sampling increased for the stricter evaluation criterion. For the overall plan results, the percentage pass rate decreased by 1.6% for both the complex and simpler plan. The maximum difference for a single beam was a decrease in pass rate of 8.9%. The simple plan had an overall MCS of 0.834, and four of the five beams had an MCS > 0.8 which would identify them as robust according to our threshold criteria. Therefore, although the increased sampling did decrease the pass rate

(probably due to higher sensitivity to gradients), the higher density sampling did not affect the identification of any of the robust beams, as percentage pass rates for those beams in question remained greater than 80%.

Analysis on the basis of field size was completed to investigate the potential dependence of pass rate on field size due to differences in diode spacing throughout the diode array. The data points (n=79) in Fig. 4 are separated into three groups: (1) total plan results (n=13), (2) the top half of beams in terms of field size (assessed simply by the number of diodes included in the analysis) (n=33), and (3) the bottom half of beams with respect to field size (n=33). Table IV shows the variation in MCS scores and the diode percentage pass rate based on the field size. There is no difference in the MCS score or pass rate based on field size, using the number of diodes involved in the analysis as representative of the maximum aperture. For certain cases, however, there may be a field size dependence associated with the location of the gradients. If the majority of the modulation and the resulting dose gradients occur outside of the central portion of the array, the sensitivity to these gradients, and any potential dosimetric errors, may be understated. One example is illustrated in Fig. 5. This beam has a MCS of 0.182, much lower than the 0.80 threshold used to identify robustness, and is identified as a large field in Fig. 4 as 233 diodes were used in the analysis. The percentage pass rate for a 2%/1 mm criteria was 90.1%.

TABLE IV. Analysis of IMRT lung beams based on field size, as determined by the number of diodes involved in the dose difference and DTA analysis, with respect to the MCS and percentage pass rate.

		Total plans	All lung beams	Beams with smaller field size (# diodes < median)	Beams with larger field size (# diodes < median)
Number of diodes	Mean (range)	n=13 N/A	n=66 242.1 (148–366)	n=33 200.9 (148–230)	n=33 283.4 (231–366)
	Median		230.5	207	281
MCS	Average	0.610	0.611	0.622	0.600
	Standard deviation	0.140	0.192	0.133	0.239
Average pass rate (%)	3%/3 mm	98.9	99.0	99.3	98.8
	2%/1 mm	84.7	85.8	84.6	87.0

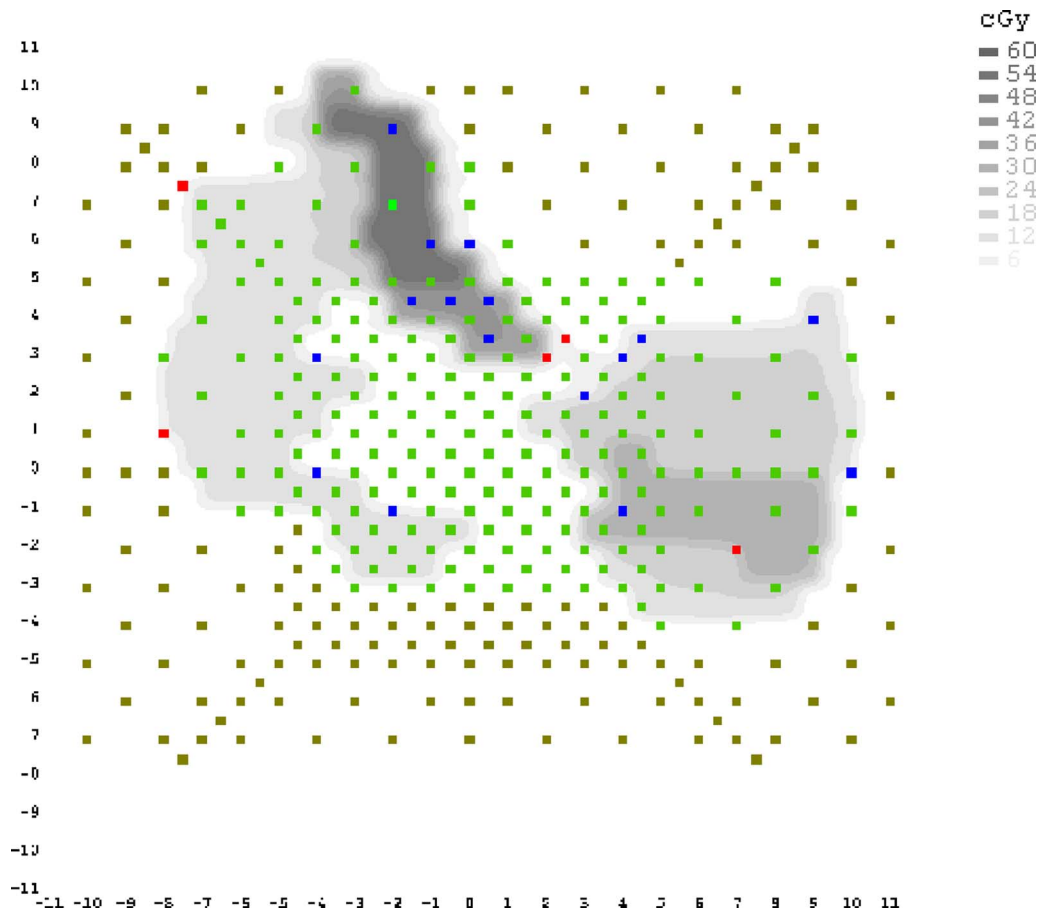


FIG. 5. MAPCHECK™ results for a single lung beam, with a MCS of 0.182, evaluated using a 2%/1 mm criterion. The beam contained 4 segments and 97 MU in total. 233 diodes were included in the analysis with the diodes depicted by lighter gray squares meeting the criteria. Black squares represent that diodes that failed to pass the criteria, measuring either a lower or higher dose than that which was calculated by the TPS. Majority of the sharp gradients and small segments that would contribute to the low MCS are outside of the inner 10×10 cm² region of the array.

IV. DISCUSSION

A good quality assurance tool is one that is easy to understand, widely applicable, easy to implement in the clinical process, and one that can provide useful information to the user that could be incorporated into the clinical process. This describes the MCS in terms of its simple calculation, applicability to analysis of individual beams as well as complete IMRT treatments and utility at various points within the clinical radiotherapy process.

The modulation complexity score has a fixed range and maximum score of 1.0, allowing for the comparison of different treatment sites, as well as single beam or total treatment plan analysis, on the same scale. This is not possible using a simple beam parameter such as the number of monitor units alone (Fig. 1). Additionally, the tool is easy to understand in this respect, as the average MCS score for a treatment site (Table I) decreases with increased inherent complexity. This allows for the separation of treatment plans along the MCS scale as illustrated in Fig. 2(a). The ability to compile plan or beam statistics and to compare current treatment plans to “typical” complexity scores could prove useful during training, planning, plan selection, and physics QA. Intuitively, the complexity required to achieve a clinically

acceptable head and neck plan would be much higher than that needed for a breast plan [increased target shape complexity, more organs at risk, potential for requiring concavities in the dose distribution (e.g., around the cord or brainstem)]. In addition to the wide applicability in terms of treatment type, the fixed range of MCS also allows for the assessment of complexity on an individual beam or a treatment plan in its entirety, more easily than the number of MU [Fig. 2(b)]. The MCS may provide more information to the quality assurance process than traditional surrogates for complexity, such as the number of MU, number of segments, or number of MU per segment, by incorporating multiple factors into the complexity assessment. This is illustrated by the lack of correlation between MCS and segment number and weak correlation with the number of MU (Fig. 3).

With respect to training for treatment planning, the statistics can help guide a planner in ways to potentially improve the treatment plan throughout the optimization process and to understand the type of complexity that is typically achievable. The ability to characterize plan complexity with a simple measure could also aid in the decision making process. In the same manner, during the treatment planning process, the ability to compare a plan’s MCS with typical values

could provide clues to when the plan complexity could potentially be decreased. For instance, it could be noted that the plan has a higher complexity than usual, and perhaps a smaller number of segments could be incorporated into the optimization process. The subcomponents of the MCS outlined in Appendix, the LSV and AAV could also be used to identify features of the plan that are atypical.

MCS could also play a role in the clinic for the purposes of plan selection. If two potential plans were similar in terms of target coverage and organ at risk sparing, a simpler plan would be desirable as that could be related to higher deliverability. This would be of particular interest in areas that are highly susceptible to intrafraction motion, such as the lung or liver, for which the interplay between motion and IMRT segments could affect the integrity of the treatment delivery.

Finally, we have shown that the MCS can potentially provide a more sensitive link to the deliverability of a IMRT beam or treatment plan than the number of MU alone. Compared to MU, the MCS was able to identify more beams, as well as a total plan, that were robustly delivered. The relationship between complexity and dosimetric outcomes, although limited, has the potential to impact the clinical quality assurance workflow. The amount of IMRT patient-specific quality assurance could theoretically be reduced if robust delivery was identified or the current workflow could be adjusted or augmented with this new information. The lung data investigated in this work for the initial investigation show that approximately 10% of the data points were identified as robust. This would not necessarily have a large impact on the QA workload, but depending on the current workflow for the IMRT QA process, it may be of assistance, by at least highlighting beams or plans which may have a complexity that should be investigated more carefully.

For lung IMRT, as implemented in our clinic, plans or beams with a MCS greater than 0.8 were dosimetrically robust. The threshold was chosen to have 100% specificity, as high specificity is essential if such a threshold was to be incorporated into the clinical quality assurance process. Ideally, if robust accurate delivery could be ensured (100% specificity), then potentially QA measurements could be eliminated for those beams. It should also be noted that for a typical dose per fraction prescription of 200 cGy, the plans had an average of approximately 5.5 beams and 95 MU per beam (MU range 27–365). Beams with less than 50 MU are not common.

The threshold values in this study are expected to be institution specific, depending on the IMRT QA process (type of dosimeter), the treatment planning system and linear accelerator (depending on accuracy of beam model and linac type), and the treatment site being considered. The comparison between planned dose and measured dose has been shown to be dependent on the beam model and the measurement data upon which the model has been based.³³ Additionally, the diode spacing of the MAPCHECKTM device and the effect on IMRT QA has been investigated.²⁵ Buonamici *et al.* found that the sampling did not play a significant role in IMRT QA results, however, they do caution that special consideration may be required for asymmetric fields. It would

seem reasonable to expect that if the majority of high dose gradients are outside of the higher density sampling region ($10 \times 10 \text{ cm}^2$), that errors in these regions may be underrepresented, leading to artificially high percentage pass rates. Initial investigation of this affect was completed by artificially increasing the diode resolution by obtaining multiple measurements for two plans. Our results indicated that generally the sampling did not have a large effect on the percentage pass rate, although the magnitude of the affect did increase with tighter evaluation criteria. Although the maximum field aperture does not appear to be a factor in the dosimetric analysis, further investigation is required to fully investigate the impact of asymmetric fields. The increased diode spacing, however, may have a larger effect on beams for which the majority of the gradients are outside of the central region of the diode array. An example of this is shown in Fig. 5, as the fluence is fairly uniform throughout the central region of the diode array, and the gradients and smaller segment sizes, contributing to a poor MCS, are primarily located where the diode spacing is greater. This field was counted as a false negative in the ROC analysis. Some of the false negatives, which decrease the sensitivity of the MCS to identify robust deliverability, could potentially be attributed to larger fields whose complexity is not accurately being captured due to the location of the gradients.

In addition to the relation between field size, and the location of the gradients with respect to the density of the diode measurements, future work will also investigate the use of test fields to identify beam model or site-specific criteria. Further studies would also be required to identify MCS thresholds that could be applied to treatment sites other than lung. For example, the breast plan evaluated was a two-tangent IMRT plan with half-beam blocks. Without shifting the diode array, the majority of the field, including most of the dose gradients created by the segments, is outside of the high-density area of the array. In such cases, the already weak dosimetric link may be hard to identify, even using stricter evaluation criterion.

It should be noted as well that all of our measurements were completed with clinically acceptable plans, and all of the dosimetric results for a 3%/3 mm criteria were excellent. The incorporation of poor plans into the analysis in the future, allowing for a broader range of data points, may provide more insight into the relationship between complexity and dosimetric deliverability. As stated previously, all measurements were also completed with an absolute dose calibration performed the day of the measurement, and a verification of the alignment of the diode array with respect to a $10 \times 10 \text{ cm}^2$ field. This process reduces the impact of linac output variation and setup error of the array itself or due to any differences between the lasers or crosshair position and the mechanical isocenter. Alignment with the square field also accounts for any offset in jaw position. All of these factors, in addition to continual refinement of the beam model due to a mature IMRT program and quality control procedures,³⁴ potentially contribute to the stringent evaluation criteria required to identify a link between delivery and complexity.

V. CONCLUSIONS

The MCS incorporates multiple factors contributing to the complexity of a radiotherapy treatment beam into a single numerical score. Complexity is often intuitively associated with accuracy in dose delivery, however, concepts such as the number of MU or beam segments are technique dependent, can represent a large of values and do not necessarily reflect any impact on the end result, which is the accuracy of the delivered dose as compared to the planned dose. The MCS allows for a quantitative assessment of plan complexity, on a finite scale that can be applied to all treatment sites, and can provide more information related to dose delivery than simple beam parameters. This could prove useful throughout the entire treatment planning and QA process.

ACKNOWLEDGMENTS

The authors would like to gratefully acknowledge Dr. Tim Craig for his input and discussions throughout the manuscript preparation process, and Dr. Daniel L  tourneau for discussions and advice with respect to the measurement and calibration procedure and analysis.

APPENDIX: MODULATION COMPLEXITY SCORE CALCULATION

The score for each beam is calculated based on three parameters that are extracted from the treatment planning system: segment shape, area, and weight.

The leaf sequence variability (LSV) parameter has been defined to characterize the variability in segment shape for a specific plan. The shape of each segment is considered, based on the change in leaf position between adjacent MLC leaves. This is calculated for leaves on each bank that define a specific segment. The LSV is defined using N , the number of open leaves constituting the beam and the coordinates of the leaf positions (pos). Leaves are not considered if they are positioned under the jaws. The position of each leaf is incorporated by defining pos_{max} . The maximum distance between positions for a leaf bank is defined as

$$\text{pos}_{\text{max}} = \langle \max(\text{pos}_{N \in n}) - \min(\text{pos}_{N \in n}) \rangle_{\text{leaf bank}}. \quad (\text{A1})$$

The LSV is then calculated as follows:

$$\text{LSV}_{\text{segment}} = \left\langle \frac{\sum_{n=1}^N (\text{pos}_{\text{max}} - (\text{pos}_n - \text{pos}_{n+1}))}{N \times \text{pos}_{\text{max}}} \right\rangle_{\text{left bank}} \times \left\langle \frac{\sum_{n=1}^N (\text{pos}_{\text{max}} - (\text{pos}_n - \text{pos}_{n+1}))}{N \times \text{pos}_{\text{max}}} \right\rangle_{\text{right bank}}. \quad (\text{A2})$$

The second IMRT segment characteristic that is considered for the overall determination of complexity is the area of the beam aperture. The aperture area variability (AAV) is used to characterize the variation in segment area relative to the maximum aperture defined by all of the segments. Segments that are more similar in area to the maximum beam aperture contribute to a larger score. The AAV is calculated using the leaf position information as follows:

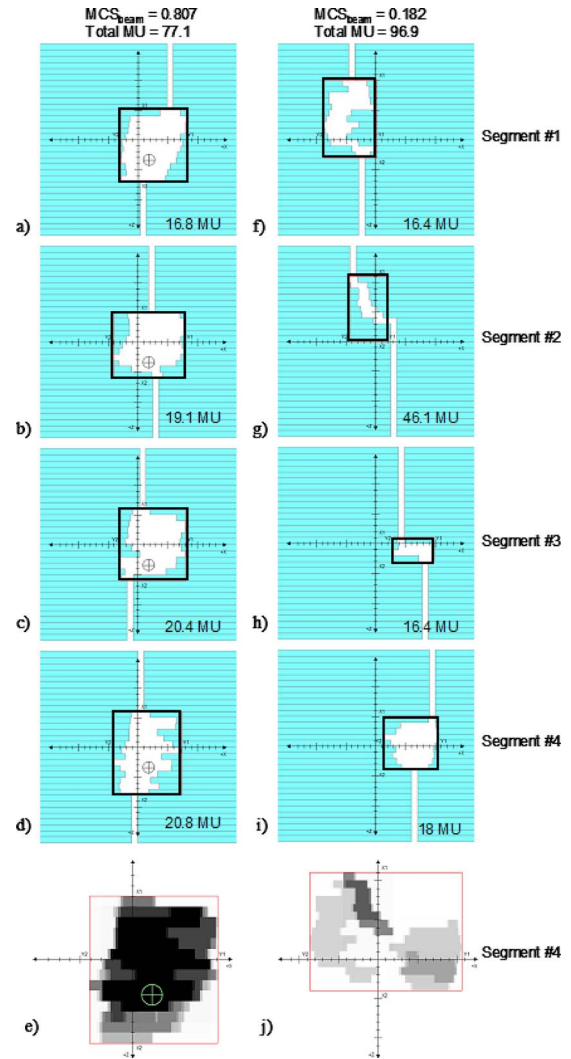


FIG. 6. Illustrations of the MLC patterns for each of the four IMRT segments, and the resulting fluence maps for [(a)–(e)] a beam with a high MCS (0.807) and [(f)–(j)] a beam with low MCS (0.182). Leaf positions under the jaws (outside of the black rectangle) are not considered in the calculation.

$\text{AAV}_{\text{segment}}$

$$= \frac{\sum_{a=1}^A \langle \text{pos}_a \rangle_{\text{left bank}} - \langle \text{pos}_a \rangle_{\text{right bank}}}{\sum_{a=1}^A \langle \max(\text{pos}_a) \rangle_{\text{left bank} \in \text{beam}} - \langle \max(\text{pos}_a) \rangle_{\text{right bank} \in \text{beam}}} \quad (\text{A3})$$

where A is the number of leaves in the leaf bank.

Finally, the relative segment weight is also incorporated into the final complexity score, with segments with a larger number of MU contributing more to the MCS score. The weight is incorporated along with AAV and LSV into the MCS calculation. The MCS_{beam} is the product of the $\text{LSV}_{\text{segment}}$ and $\text{AAV}_{\text{segment}}$ weighted by the relative MU of each segment in the beam,

$$\text{MCS}_{\text{beam}} = \sum_{i=1}^I \text{AAV}_{\text{segment } i} \times \text{LSV}_{\text{segment } i} \times \frac{\text{MU}_{\text{segment } i}}{\text{MU}_{\text{beam}}}, \quad (\text{A4})$$

where I is the number of segments in the beam.

The total plan complexity is defined by the MCS_{plan} metric. The MCS_{plan} is the MCS_{beam} weighted by the relative MU of each beam in the plan,

$$MCS_{\text{plan}} = \sum_{j=1}^J MCS_{\text{beam } j} \times \frac{MU_{\text{beam } j}}{MU_{\text{plan}}}, \quad (\text{A5})$$

where J is the number of beams in the total plan.

Figure 6 illustrates the impact of variations in leaf position or area for two different beams, both with four segments, but a large difference in MCS_{beam} scores. The first beam [Figs. 6(a)–6(e)] has a much more uniform beam aperture, and little variability between leaf positions from segment to segment. Thus the values of LSV and AAV would be higher than that of the second beam, resulting in a higher MCS.

^{a)}Electronic mail: andrea.mcniven@rmp.uhn.on.ca

- ¹D. A. Kuban, S. L. Tucker, L. Dong, G. Starkschall, E. H. Huang, M. R. Cheung, A. K. Lee, and A. Pollack, "Long-term results of the M. D. Anderson randomized dose-escalation trial for prostate cancer," *Int. J. Radiat. Oncol., Biol., Phys.* **70**, 67–74 (2008).
- ²N. Lee, P. Xia, J. M. Quivey, K. Sultanem, I. Poon, C. Akazawa, P. Akazawa, V. Weinberg, and K. K. Fu, "Intensity-modulated radiotherapy in the treatment of nasopharyngeal carcinoma: an update of the UCSF experience," *Int. J. Radiat. Oncol., Biol., Phys.* **53**, 12–22 (2002).
- ³M. J. Zelefsky, Z. Fuks, L. Happersett, H. J. Lee, C. C. Ling, C. M. Burman, M. Hunt, T. Wolfe, E. S. Venkatraman, A. Jackson, M. Skwar-chuk, and S. A. Leibel, "Clinical experience with intensity modulated radiation therapy (IMRT) in prostate cancer," *Radiother. Oncol.* **55**, 241–249 (2000).
- ⁴J. P. Pignol, I. Olivetto, E. Rakovitch, S. Gardner, K. Sixel, W. Beckham, T. T. Vu, P. Truong, I. Ackerman, and L. Paszat, "A multicenter randomized trial of breast intensity-modulated radiation therapy to reduce acute radiation dermatitis," *J. Clin. Oncol.* **26**, 2085–2092 (2008).
- ⁵G. A. Ezzell, J. M. Galvin, D. Low, J. R. Palta, I. Rosen, M. B. Sharpe, P. Xia, Y. Xiao, L. Xing, and C. X. Yu, "Guidance document on delivery, treatment planning, and clinical implementation of IMRT: report of the IMRT Subcommittee of the AAPM Radiation Therapy Committee," *Med. Phys.* **30**, 2089–2115 (2003).
- ⁶J. M. Galvin, G. Ezzell, A. Eisbrauch, C. Yu, B. Butler, Y. Xiao, I. Rosen, J. Rosenman, M. Sharpe, L. Xing, P. Xia, T. Lomax, D. A. Low, and J. Palta, "Implementing IMRT in clinical practice: a joint document of the American Society for Therapeutic Radiology and Oncology and the American Association of Physicists in Medicine," *Int. J. Radiat. Oncol., Biol., Phys.* **58**, 1616–1634 (2004).
- ⁷E. E. Klein, J. Hanley, J. E. Bayouth, 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**, 4197–4212 (2009).
- ⁸*Guidelines for the Verification of IMR*, edited by B. J. Mijnheer and D. Georg (ESTRO, Brussels, 2008).
- ⁹R. Mohan, M. Arnfield, S. Tong, Q. Wu, and J. Siebers, "The impact of fluctuations in intensity patterns on the number of monitor units and the quality and accuracy of intensity modulated radiotherapy," *Med. Phys.* **27**, 1226–1237 (2000).
- ¹⁰D. Craft, P. Suss, and T. Bortfeld, "The tradeoff between treatment plan quality and required number of monitor units in intensity-modulated radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* **67**, 1596–1605 (2007).
- ¹¹S. Webb, "Use of a quantitative index of beam modulation to characterize dose conformality: illustration by a comparison of full beamlet IMRT, few-segment IMRT (fsIMRT) and conformal unmodulated radiotherapy," *Phys. Med. Biol.* **48**, 2051–2062 (2003).
- ¹²M. M. Matuszak, E. W. Larsen, and B. A. Fraass, "Reduction of IMRT beam complexity through the use of beam modulation penalties in the objective function," *Med. Phys.* **34**, 507–520 (2007).
- ¹³M. M. Coselmon, J. M. Moran, J. D. Radawski, and B. A. Fraass, "Im-proving IMRT delivery efficiency using intensity limits during inverse planning," *Med. Phys.* **32**, 1234–1245 (2005).
- ¹⁴J. Dai and Y. Zhu, "Minimizing the number of segments in a delivery sequence for intensity-modulated radiation therapy with a multileaf collimator," *Med. Phys.* **28**, 2113–2120 (2001).
- ¹⁵M. M. Matuszak, E. W. Larsen, K. W. Jee, D. L. McShan, and B. A. Fraass, "Adaptive diffusion smoothing: a diffusion-based method to reduce IMRT field complexity," *Med. Phys.* **35**, 1532–1546 (2008).
- ¹⁶P. Xia and L. J. Verhey, "Multileaf collimator leaf sequencing algorithm for intensity modulated beams with multiple static segments," *Med. Phys.* **25**, 1424–1434 (1998).
- ¹⁷D. M. Shepard, M. A. Earl, X. A. Li, S. Naqvi, and C. Yu, "Direct aperture optimization: a turnkey solution for step-and-shoot IMRT," *Med. Phys.* **29**, 1007–1018 (2002).
- ¹⁸S. Webb, D. J. Convery, and P. M. Evans, "Inverse planning with constraints to generate smoothed intensity-modulated beams," *Phys. Med. Biol.* **43**, 2785–2794 (1998).
- ¹⁹J. D. Fenwick, W. A. Tome, H. A. Jaradat, S. K. Hui, J. A. James, J. P. Balog, C. N. DeSouza, D. B. Lucas, G. H. Olivera, T. R. Mackie, and B. R. Paliwal, "Quality assurance of a helical tomotherapy machine," *Phys. Med. Biol.* **49**, 2933–2953 (2004).
- ²⁰T. 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**, 1919–1927 (1998).
- ²¹G. A. Ezzell and S. Chungbin, "The overshoot phenomenon in step-and-shoot IMRT delivery," *J. Appl. Clin. Med. Phys.* **2**, 138–148 (2001).
- ²²J. G. Li, J. F. Dempsey, L. Ding, C. Liu, and J. R. Palta, "Validation of dynamic MLC-controller log files using a two-dimensional diode array," *Med. Phys.* **30**, 799–805 (2003).
- ²³N. Giorgia, F. Antonella, V. Eugenio, C. Alessandro, A. Filippo, and C. Luca, "What is an acceptably smoothed fluence? Dosimetric and delivery considerations for dynamic sliding window IMRT," *Radiat. Oncol.* **2**, 42 (2007).
- ²⁴S. Both, I. M. Alecu, A. R. Stan, M. Alecu, A. Ciura, J. M. Hansen, and R. Alecu, "A study to establish reasonable action limits for patient-specific quality assurance in intensity-modulated radiation therapy," *J. Appl. Clin. Med. Phys.* **8**, 1–8 (2007).
- ²⁵F. B. Buonamici, A. Compagnucci, L. Marrazzo, S. Russo, and M. Buc-cioli, "An intercomparison between film dosimetry and diode matrix for IMRT quality assurance," *Med. Phys.* **34**, 1372–1379 (2007).
- ²⁶P. A. Jursinic and B. E. Nelms, "A 2-D diode array and analysis software for verification of intensity modulated radiation therapy delivery," *Med. Phys.* **30**, 870–879 (2003).
- ²⁷D. 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).
- ²⁸J. G. Li, G. Yan, and C. Liu, "Comparison of two commercial detector arrays for IMRT quality assurance," *J. Appl. Clin. Med. Phys.* **10**, 62–74 (2009).
- ²⁹B. E. Nelms and J. A. Simon, "A survey on planar IMRT QA analysis," *J. Appl. Clin. Med. Phys.* **8**, 76–90 (2007).
- ³⁰L. N. McDermott, M. Wendling, J. J. Sonke, M. van Herk, and B. J. Mijnheer, "Replacing pretreatment verification with in vivo EPID dosimetry for prostate IMRT," *Int. J. Radiat. Oncol., Biol., Phys.* **67**, 1568–1577 (2007).
- ³¹R. A. Siochi, E. C. Pennington, T. J. Waldron, and J. E. Bayouth, "Radiation therapy plan checks in a paperless clinic," *J. Appl. Clin. Med. Phys.* **10**(1), 43–62 (2009).
- ³²M. De Brabandere, A. Van Esch, G. J. Kutcher, and D. Huyskens, "Quality assurance in intensity modulated radiotherapy by identifying standards and patterns in treatment preparation: a feasibility study on prostate treatments," *Radiother. Oncol.* **62**, 283–291 (2002).
- ³³G. Yan, C. Fox, C. Liu, and J. G. Li, "The extraction of true profiles for TPS commissioning and its impact on IMRT patient-specific QA," *Med. Phys.* **35**, 3661–3670 (2008).
- ³⁴S. L. Breen, D. J. Moseley, B. Zhang, and M. B. Sharpe, "Statistical process control for IMRT dosimetric verification," *Med. Phys.* **35**, 4417–4425 (2008).