

An analysis of tolerance levels in IMRT quality assurance procedures

Parminder S. Basran^{a)}

Department of Radiation Oncology, University of Toronto, Toronto, Ontario M4N 3M5, Canada and Department of Medical Physics, Odette Cancer Centre, Sunnybrook Health Sciences Centre, 2075 Bayview Avenue, Toronto, Ontario M4N 3M5, Canada

Milton K. Woo

Department of Radiation Oncology, University of Toronto, Toronto, Ontario M4N 3M5, Canada, Department of Medical Physics, Odette Cancer Centre, Sunnybrook Health Sciences Centre, 2075 Bayview Avenue, Toronto, Ontario M4N 3M5, Canada, and Department of Medical Biophysics, University of Toronto, Toronto, Ontario M4N 3M5, Canada

(Received 3 January 2008; revised 1 April 2008; accepted for publication 2 April 2008; published 14 May 2008)

Increased use of intensity modulated radiation therapy (IMRT) has resulted in increased efforts in patient quality assurance (QA). Software and detector systems intended to streamline the IMRT quality assurance process often report metrics, such as percent discrepancies between measured and computed doses, which can be compared to benchmark or threshold values. The purpose of this work is to examine the relationships between two different types of IMRT QA processes in order to define, or refine, appropriate tolerance values. For 115 IMRT plans delivered in a 3 month period, we examine the discrepancies between (a) the treatment planning system (TPS) and results from a commercial independent monitor unit (MU) calculation program; (b) TPS and results from a commercial diode-array measurement system; and (c) the independent MU calculation and the diode-array measurements. Statistical tests were performed to assess significance in the IMRT QA results for different disease site and machine models. There is no evidence that the average total dose discrepancy in the monitor unit calculation depends on the disease site. Second, the discrepancies in the two IMRT QA methods are independent: there is no evidence that a better—or worse—monitor unit validation result is related to a better—or worse—diode-array measurement result. Third, there is marginal benefit in repeating the independent MU calculation with a more suitable dose point, if the initial IMRT QA failed a certain tolerance. Based on these findings, the authors conclude at some acceptable tolerances based on disease site and IMRT QA method. Specifically, monitor unit validations are expected to have a total dose discrepancy of 3% overall, and 5% per beam, independent of disease site. Diode array measurements are expected to have a total absolute dose discrepancy of 3% overall, and 3% per beam, independent of disease site. The percent of pixels exceeding a 3% and 3 mm threshold in a gamma analysis should be greater than or equal to 95% for non-head and neck IMRT cases, and 88% for head and neck IMRT cases. The IMRT QA methodology described here is neither unique nor ubiquitous, and the ability to deliver a safe IMRT does not simply require IMRT QA tests to pass a given tolerance; however, the selection of a tolerance should be meaningful when assessing a complex plan. The methodology in defining appropriate tolerances, described in this article, is based on an interpretation of IMRT QA results from IMRT plans deemed safe to deliver. © 2008 American Association of Physicists in Medicine. [DOI: [10.1118/1.2919075](https://doi.org/10.1118/1.2919075)]

Key words: IMRT, quality assurance, radiation treatment planning, dose verification, dose calculation

I. INTRODUCTION

In intensity modulated radiation therapy (IMRT) techniques, patient delivery quality assurance has become an integral part of the treatment process. Much of the justification, philosophy and requirements for the IMRT quality assurance (QA) are given in the AAPM Guidance document on IMRT.¹ In addition, there are task groups working on new standards for IMRT, such as the AAPM Task Groups 119, 120, and the Ontario (Canada) organizational standards for IMRT, but a comprehensive report is still forthcoming. Generally, IMRT is a complex technique involving many steps in simulation,

planning, file transfers, setup, on-line verification; thus, stringent QA procedures that ensure the process is carried out as prescribed and planned are highly essential.

Historically, ion chamber and film have been used from the beginning. In fact the first commercial IMRT treatment planning system, Corvus,² developed a standard built-in function to generate a phantom plan from which a measurement using ion chamber and film could be compared to the plan calculations.³ The ion chamber measurement is intended to serve as a verification of the monitor unit calculation by the treatment planning system, as per the recommendation of ICRU Report 24,⁴ and the film measurement is intended to

verify the overall integrity of the treatment plan as transmitted from the treatment planning system to the delivery linac. At that time, the availability of independent monitor unit calculation software was limited. Because of the many segments involved, it was impractical to manually check the monitor units for each segment. The measurement provided an acceptable way of assuring the delivery of the plan. As IMRT techniques matured over the years, film and ion chamber measurements in the IMRT QA process have remained and have become an accepted standard method of IMRT patient QA.

Clearly, the use of IMRT is increasing rapidly, and with it the necessity of patient-specific QA. The IMRT QA itself is a heavy workload on the physics department, often occurring after clinical operations hours. At the same time, many new developments have also occurred in the area of IMRT QA. These include the development of new phantoms such as the Quasar IGRT phantom,⁵ new devices such as diode or ion chamber matrix detectors,^{6,7} and electronic portal imaging devices (EPID) devices that could replace film measurement. Software programs have been developed to serve as a means of validating the monitor unit calculations by the treatment planning systems. These new developments are intended to streamline the IMRT QA process and minimize the IMRT QA workload. Some IMRT QA methods reduce this workload by reducing the effort to verify the plan, while others reduce the effort by minimizing the need for measurement.

However, these new developments have a relatively short history, and little experience has been accumulated. In fact, most centers have developed their own particular IMRT QA processes, employing different hardware devices and software programs. These processes usually evolve into a decision tree structure, where the outcome of one test process might trigger a second level of different tests. One example is a process where a monitor unit calculation program is used first, and an ion chamber measurement is done if the monitor unit calculation results do not agree with the treatment planning system within a predetermined tolerance. This may be followed by beam segments verification using an EPID.

One aspect of the process that has not been examined objectively and analytically is the tolerance levels used in driving the decision tree above. Traditionally, from ICRU Report 24,⁴ an accuracy level of 5% was the accuracy level required for treatment plan dose calculations. When IMRT first became clinically practical, this tolerance level was directly adopted. As IMRT becomes more prevalent and the precision in planning and delivery increases, the natural tendency is to demand more accurate dose delivery and validation. With the advent of more sophisticated algorithms and more complex fluence delivery patterns, values of 2%–3% in high dose and low gradient regions and 4% in high dose and high gradient regions are achievable.⁸ As a result, tighter tolerance has often been adopted, although no universal standard has been accepted. However, the implication of this tight tolerance level on the patient IMRT QA process has not been addressed satisfactorily. One of the objectives of this work is to examine this aspect in more detail.

The objectives this article are (1) to examine the results of two different types of IMRT QA processes (independent monitor unit calculation and a two-dimensional diode-array measurement) in the hopes of selecting appropriate values for IMRT QA tolerances; (2) to examine whether there are relationships between the different metrics provided by the IMRT QA; and (3) within the context of the described IMRT QA decision tree, to examine whether more efficient IMRT QA processes might be possible. These objectives are achieved first with a statistical evaluation of discrepancies in the metrics reported by the IMRT planning system and an independent monitor unit (MU) checking program and diode array measurement. Second, we examine the possible relationships between the independent dose calculation and the corresponding IMRT QA measurements. Finally, we arrive at some recommendations that could further streamline the IMRT QA process to meet the future workload demands of IMRT. The purpose of this article is not to discuss what tolerance values others should use, but to demonstrate how an analysis of IMRT QA results may be used to provide some guidance on selecting appropriate tolerances in their practice.

II. METHODS AND MATERIALS

Over the course of 3 months, a retrospective analysis of IMRT QA results was performed in our center. All IMRT plans were categorized according to anatomical site, defined here as GU (prostate), HN (head and neck), lung (LUNG) and “other”. Because of the low number of other cases, they are ignored in this analysis. The total number of monitor units, photon beam used for treatment, number of target doses in the IMRT plan and planned dose per fraction were recorded for all IMRT plans. It should be noted that all IMRT QA plans had been checked and deemed clinically acceptable in terms of dosimetric coverage, dose sparing, and were physically deliverable by the treatment unit, but still required some type of dose validation.

The decision tree for the IMRT QA process adopted in our center is shown in Fig. 1. All plans have an independent monitor unit computation. If the disease site is the prostate gland, where the target is small and fluence patterns are not highly irregular, no additional IMRT QA is required if the independent MU computation calculates the individual beams within 5% of the treatment planning system, and the total dose from all beams is within 3% of the treatment planning system (TPS). (Note also that ablative doses to the prostate, or doses greater than 300 cGy/fraction, require a measurement.) If either tolerance is exceeded, then a measurement is performed. If the disease site is not prostate, then a two-dimensional dose measurement of all beams is also performed. It is worthwhile to point out that, unlike in the United States, there is no requirement for an IMRT measurement in order to adhere to billing practices in countries such as Canada where this study was undertaken; therefore, there is no obligation to perform an IMRT measurement on

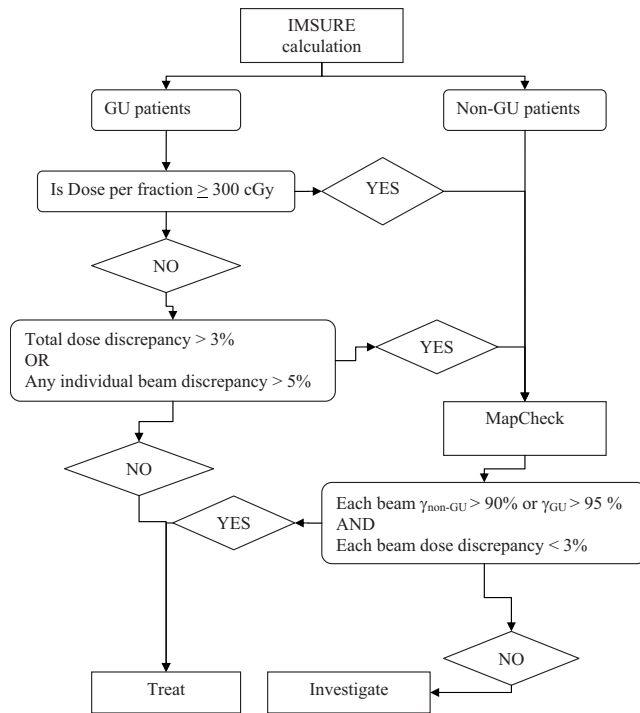


FIG. 1. Current decision tree for IMRT quality assurance along with tolerance levels for MU calculation and planar dose measurements prior to this study. Note that for GU patients, an IMRT measurement with MapCheck may not be required.

all IMRT patients. A brief description of the planning process is given, details of the IMRT QA processes, data collected, and methods of data analysis are described.

II.A. IMRT planning process

An isocentric IMRT plan is generated within the Pinnacle (Version 8.0d/h, Philips Medical Systems, Cleveland, OH) TPS where the prescription dose is delivered to a target structure. Note that in our IMRT planning process, the use of asymmetric jaws arrangements is preferred over isocenter shifts. The IMRT is delivered on both Siemens Primus (Siemens Medical Systems, Malvern, PA) and Elekta Synergy (Elekta Oncology Systems, Sunnyvale, CA) machines using 6 MV photons with segmented delivery of static MLC fields. During the planning, a “dose point” is created within the target receiving the highest dose where a uniform dose is required. Each beam uses this dose point as the reference point for the monitor unit calculation and isodose distribution normalization. The sum of all beams’ reference point doses will roughly equal the dose per fraction of the target structure. Once the TPS plan is approved, documents are generated and an electronic DICOM RT file, which contains all aspects of the IMRT delivery, is exported to the record and verify system.

The DICOM RT file is then converted to a format accessible to the record and verify system, MOSAIQ (Version 1.2, IMPAC Medical Systems, Sunnyvale, CA) in a file called hereafter the “RTP” file. After the RTP file is imported into the record and verify system, the RTP file is moved to a

location containing only those files that have been imported into the record and verify system. The imported RTP files are used for the independent dose calculation.

II.B. Independent monitor unit calculation

The first method of the IMRT QA is the monitor unit validation, facilitated by the commercial software IMSURE (Version 3.0, Standard Imaging, Middleton, WI). This software calculates the dose from each IMRT beam using a three-source analytical model of head-scatter, along with look-up tables for tissue phantom ratios and phantom scatter factors.⁹ The RTP file contains some, but not all, structures within the DICOM RT file. Some important structures in the RTP file are each beam’s source to surface distance (SSD), physical distance from the dose point to the patient surface (r), monitor units in each segment, and the dose to the dose point as calculated by the treatment planning system. Note that r generally does not equal the distance from the SSD to the dose point since the dose point is generally not on the central axis. By specifying source to surface distance of the dose point, or point SSD (PSSD), and the coordinates of the dose point with respect to the isocenter, IMSURE can compute the physical depth of the dose point. Along with the dose point coordinates, and PSSDs, the radiological effective depths of the dose points are also entered into the IMSURE software for the dose calculation. IMSURE then computes the dose from each beam and reports the dose discrepancy between the calculated and TPS dose from each beam, and the dose discrepancy between the total dose calculated by IMSURE and the TPS.

For the IMSURE to TPS analysis, individual beam discrepancies and total beam discrepancy between IMSURE and the planning system were recorded. Also, the “worst beam” discrepancy value, defined as the discrepancy with the largest minimum or maximum value, was recorded. The average and standard deviation of all individual beam discrepancies in an IMRT plan was also calculated.

II.C. Diode-array measurement

The second IMRT QA method is the MapCheck (Version 3.03, Sun Nuclear Corporation, Melbourne, FL) diode array system using a process similar to that described by Létourneau *et al.*¹⁰ To generate the MapCheck QA geometry, the intended IMRT plan is replanned in “QA mode” within the TPS, where the beam geometry of the patient is replaced with a computed tomography scan of a solid water phantom. The dimensions of the solid water are approximately $30 \times 30 \times 30$ cm. The gantry angles and collimator angles from each beam are zeroed such that all beams irradiate the center of the solid water at a plane perpendicular to the solid water phantom. The prescription in the QA plan is modified such that monitor units from each beam deliver the prescribed monitor units in the treatment plan. The segment shapes and monitor units from each beam are checked against the treatment planning segment shapes and monitor units prior to calculation. The SSDs for all beams are fixed at 90 cm, such that the isocenter is 10 cm deep in the solid water phantom.

The dose to a plane 10 cm deep in the solid water phantom is computed at a dose resolution of 3×3 mm, using the “planar dose” routine in the TPS. The computed dose to the plane perpendicular to the beam at 10 cm depth is then exported in an ASCII format and stored on a network drive. The ASCII file contains the dose per monitor unit from the beam and so the total monitor units from each beam must be recorded for the MapCheck to TPS analysis.

Prior to each measurement, the MapCheck device is calibrated for absolute dose with a flood field of radiation. This is done by delivering a large beam (22×22 cm) to the diode array with a monitor unit setting roughly equal to the average MUs delivered in each beam. The large field irradiates all diode detectors and the diode signals are compared against a benchmark set of diode readings. Then, after applying a dose per MU factor at a 10 cm depth, correction factors are applied to provide a measure of the absolute dose delivered to each diode. With this approach, the diode readings in absolute dose mode are accurate to within 1%, similar to those reported by Létourneau *et al.*¹⁰

Because there is an inherent 2 cm of buildup on the diode array, 8 cm of solid water material are placed on the phantom and aligned with the beam. If the fluence patterns are delivered far from the central axis such that the MapCheck diodes are only partially irradiated, a “shift” is invoked in order to exploit the higher density of diodes and hence, more dose samples in the center of the diode array. In the record and verify system, all IMRT QA plans are delivered in QA mode where the intended fields are delivered but not recorded as dose delivered to the patient. Because the gantry angles and collimator angles are deliberately set to zero, they must be “overridden” in the record and verify system during the QA measurement.

The measured and calculated planar dose grids are then compared using MapCheck software. Beams are analyzed in “absolute dose” mode, using the calibration factors obtained immediately prior to the IMRT QA measurement. Corrections for orientation and alignment of the measured to computed planar doses are applied if required. Diode readings lower than 10% of the highest diode signal are ignored in the analysis since they are normally under the jaws in the low dose and low gradient regions where the diode response is less reliable and the signal to noise ratio to the diodes becomes a concern. Measured and calculated distributions are compared via gamma analysis,¹¹ using a 3%/3 mm tolerance, and isodose display. For each beam, a diode “point dose” is chosen in a high dose, low gradient region and the dose to the diode is compared to the dose to the corresponding coordinates in the calculated dose grid from the TPS. When comparing doses from all beams, point dose measurements from all beams are summed and compared to the corresponding sum of point doses from the TPS.

For the MapCheck to TPS IMRT QA plan analysis, individual beam dose discrepancies, total dose discrepancies, and gamma values were recorded for each beam. Also, the worst beam dose discrepancy value, defined by the discrepancy with the largest minimum or maximum beam discrepancies,

TABLE I. Averages, μ , of the total dose discrepancies in the independent monitor unit calculations (IMSURE) and TPS. Also shown are the standard deviations, σ , number of data points, N , and the maximum and minimum observed errors.

	N	μ (%)	σ (%)	Max (%)	Min (%)
Total	115	2.1	2.0	7.4	-3.7
GU	55	3.2	1.5	7.2	-0.5
HN	52	1.2	1.9	5.5	-2.2
LUNG	8	0.2	1.8	2.7	-2.8

and the worst beam gamma value, defined by the beam with the lowest gamma, were recorded and identified.

II.D. Sample size and statistical analysis

A total of 115 IMRT plans were evaluated. None were excluded in the analysis. Of the 115 IMRT plans, 84 also had MapCheck measurements. In total, 55 GU, 52 HN, and 8 LUNG IMRT cases were reviewed. All plans described here were deemed safe to deliver and were consequently used for treatment (i.e., there were no modifications to any of the existing plans). Microsoft Excel (Microsoft Corporation, Redmond, WA) was used to compile data and perform basic statistical tests for analyzing the data. Average discrepancies between TPS/IMSURE and TPS/MapCheck were tested for significance using one or two-sided t tests. Fisher’s F test was used to test for differences in standard deviations, and correlation coefficients were computed to examine possible correlations in the data.

III. RESULTS

III.A. Comparison of IMSURE and TPS

Table I displays the total dose discrepancy between IMSURE and the TPS (positive number indicates that IMSURE calculation is higher than TPS). Also shown are the discrepancies based on treatment site.

Table II displays a summary of individual beam discrepancies between IMSURE and the TPS. A Fisher’s F test reveals that the standard deviations from GU and HN sites are the same ($P=0.999$), suggesting that the range in individual beam discrepancies is similar for both HN and GU; however,

TABLE II. Summary of IMSURE and TPS and individual dose discrepancies, where N is the number of data points, and μ and σ are the mean and standard deviations.

	N	μ (%)	σ (%)
Total	788	1.8	4.2
GU	380	3.2	1.7
HN	355	1.4	2.2
LUNG	53	-0.6	2.8

TABLE III. Summary of MapCheck and TPS and total dose discrepancies, where N is the number of data points, and μ and σ are the mean and standard deviations.

	N	μ (%)	σ (%)
Total	84	-0.2	0.8
GU	21	-0.5	0.7
HN	53	-0.1	0.9
LUNG	10	0.5	1.5

the range in individual beam discrepancies for LUNG plans is significantly different from GU and HN ($P < 0.001$).

III.B. Comparison of MapCheck and TPS

Table III displays a summary of total dose discrepancy between MapCheck and the TPS IMRT QA plan calculation (negative number indicates that MapCheck measurement is lower than TPS). Clearly, the average dose discrepancy from all beams is within zero, or well within error of the measurement, and total dose discrepancy is zero independent of site or machine type. There were no total dose discrepancies greater than $\pm 2.5\%$.

Table IV displays a summary of individual dose point discrepancies between MapCheck and the TPS. Again, all discrepancies are within measurement error and the individual beam dose discrepancy is zero, regardless of site or machine type. Note again the small standard deviation, suggesting robustness in the measurement process. There were no absolute dose beam discrepancies greater than $\pm 6.0\%$.

Table V displays a summary of gamma analyses comparing the MapCheck measurements to the TPS. Using a one-sided t test, the average gamma values of GU and HN sites were significantly different ($P < 0.001$), as were the average values of GU and LUNG ($P < 0.001$), and HN and LUNG ($P < 0.001$). This may be explained by the increased degree of complexity often presented in HN IMRT.

III.C. Comparison of IMSURE and MapCheck

While IMSURE and MapCheck are two different types of IMRT QA, for the sake of completeness we tested for possible relationships between IMSURE calculations and MapCheck measurements.

TABLE IV. Summary of MapCheck and TPS and individual beam dose discrepancies, where N is the number of data analyzed, and μ and σ are the mean and standard deviations.

	N	μ (%)	σ (%)
Total	556	0.3	1.5
GU	150	0.5	0.7
HN	353	0.2	0.6
LUNG	53	-0.1	0.4

TABLE V. Summary of MapCheck and TPS and gamma analysis, where N is the number of data points, μ and σ are the mean and standard deviations, respectively, and "Minimum" refers to the minimum value of gamma within that data set.

	N	μ	σ	Minimum
Total	556	96.5	4.0	77.6
GU	150	98.8	2.0	86.7
HN	353	95.5	3.5	77.6
LUNG	53	97.3	1.6	90.2

Figure 2 displays a scatter plot of total dose discrepancies from MapCheck versus that from IMSURE. There is no correlation between a poor total dose MapCheck result and a poor IMSURE result (correlation coefficient, or $R < 0.001$), regardless of machine type or site.

Figure 3 displays a scatter plot of individual beam dose discrepancies from MapCheck versus discrepancies from IMSURE. Again, there is no correlation between MapCheck and IMSURE beam discrepancies ($R < 0.003$).

Similar relationships were observed between a beam's gamma value and individual beam discrepancies from IMSURE ($R = 0.03$) and a beam's gamma value and Mapcheck beam dose discrepancies ($R = 0.07$).

III.D. Reassessment of point-dose calculations

A striking result from the independent monitor unit validation is the spread of individual beam dose discrepancies (see x axis on Fig. 3). Individual beam dose discrepancies of up to $\pm 25\%$ are observed. Some of these large errors are due to the fact that some beams deliver a small dose (less than 5% of the total dose per fraction) and so the effects of a small absolute dose error are disproportionate to the relative

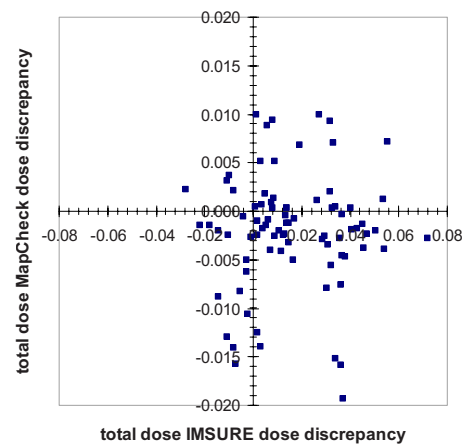


FIG. 2. A scatter plot of the total dose discrepancies from MapCheck measurements compared to those of the IMSURE results, with respect to the TPS plan calculation. The dose to a point from each beam is summed in the MapCheck measurements and compared to the expected beam doses in the TPS. The discrepancy between measured total dose and TPS total doses is plotted on the Y axis. The dose to a point from each beam is summed in the IMSURE calculation and compared to the expected beam doses in the TPS. The discrepancy between the IMSURE total dose and TPS total doses is plotted on the X axis.

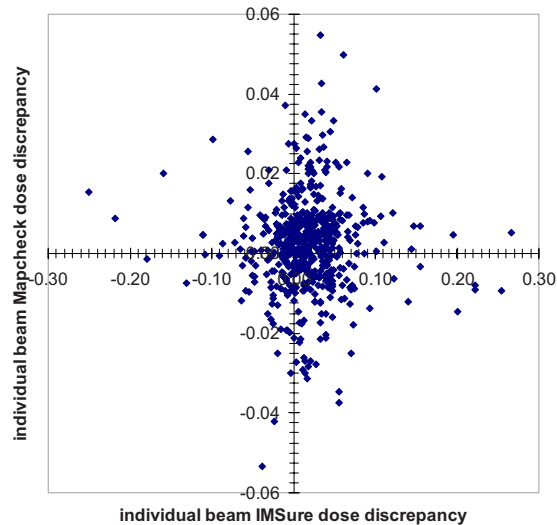


FIG. 3. A scatter plot of the individual beam dose discrepancies from MapCheck measurements compared to those of the IMSURE results, with respect to the TPS plan calculation. The dose to a point from each beam in the MapCheck measurements is compared to the expected beam dose in the TPS. The discrepancy between measured beam dose and TPS beam doses is plotted on the Y axis. The dose to a point from each beam in the IMSURE calculation is compared to the expected beam dose in the TPS. The discrepancy between the IMSURE beam dose and TPS beam doses is plotted on the X axis.

error. However, even when such beams are ignored in the data analysis, large individual beam discrepancies ($>10\%$) are still observed. One possible explanation is that the selected dose point is in a dose gradient or is close to MLC leaf ends in one or few of the segments.

We investigated whether poor individual beam results were due to a poor dose point selection. Only HN cases were chosen in this analysis since HN plans often have irregular segment shapes. A tolerance level of 3% overall and 4% per beam was defined. If either tolerance was not met in the IMSURE calculation, a recalculation of IMSURE was performed using a new dose point. The new dose point was carefully selected within the IMRT plan in a high and uniform dose region over a $2 \times 2 \times 2 \text{ cm}^3$ volume. A total of 39 HN IMRT plans were reevaluated. Of the 39 plans, 22 (56%) passed both the individual beam dose and total dose discrepancy threshold. After selecting the new dose point, the average of the total dose discrepancy remained unchanged (1.8%), with only a slight reduction in the standard deviation ($\sigma=2.1\%$, as opposed to 2.5%). Table VI displays the change in individual beam dose discrepancies. While it appears that there is a slight reduction in the average dose error, there is an increase in the standard deviation, suggesting that the spread in the individual beam doses increases when choosing a new dose point. Further, if the direction of the error is ignored, there is a slight reduction in the average error and a slight reduction in the maximum error observed in the plan (maximum error of 7.1% as opposed to 9.4%) but, again the standard deviation increases. As a result of the reselection of the dose point, 25 of the 39 (64%) plans meet the individual beam threshold of 5% and total dose threshold of 4%.

TABLE VI. Individual beam dose discrepancies between IMSURE and TPS. Directional discrepancies take into account the sign (positive or negative) of the individual dose discrepancies, whereas the nondirectional discrepancies are results when the absolute value of the discrepancy is used.

Beam dose discrepancies (directional)	
	$\mu(\sigma)$
Before	2.1% (2.6%)
After	1.3% (3.0%)
Beam dose discrepancies (nondirectional)	
	$\mu(\sigma)$
Before	2.8% (1.8%)
After	2.5% (2.3%)

IV. DISCUSSION

From this analysis, we have come to a number of conclusions. First, there is no evidence that the total dose discrepancies from the monitor unit validation are dependent on the site. Therefore, there is no reason to not use a single value for the total dose discrepancy threshold regardless of site treated. We have chosen to use a total dose discrepancy tolerance of 3% for all sites when using the monitor unit validation. The tolerances for individual beam dose discrepancies will be discussed later in this section.

Second, we have demonstrated that the two QA methods are completely independent of each other. There is no conclusive evidence that a better—or worse—monitor unit validation result is related to a better—or worse—MapCheck result. One of the major concerns in our current process is the fact that all QA results are included as an official record in the patient chart. When beams show a large discrepancy ($>5\%$) in the monitor unit calculation, the subsequent MapCheck measurement is used to “justify” acceptance of the plan. Although there might be a visual inspection of the location of the point of dose calculation in IMSURE, there is no significant effort put into explaining each poor IMSURE result. By showing that there is no correlation between the IMSURE calculation and the MapCheck measurement, we essentially remove the need to explain the poor IMSURE result; the two tests are independent and the MapCheck result does not add value in interpreting the IMSURE result. Moreover, the MapCheck result is deemed a more robust QA procedure since analytical MU checking programs can have difficulty in computing dose close to MLC leaf edges and modeling head-scatter with sufficient accuracy in highly modulated beams. This is likely the primary reason why IMSURE fails to meet our IMRT QA tolerances.

If both IMRT QA processes are viewed as processes that validate the delivered dose, then either the IMSURE or MapCheck process will suffice. IMSURE is the first choice because of the lesser resource requirement. Neither test should be considered a more accurate means of validating the delivered dose. The MapCheck does, however, have added value by verifying that the machine can deliver the desired segments, and also provides a means of validating

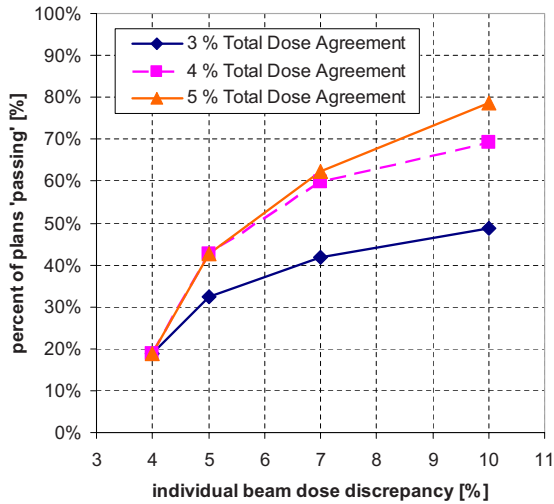


FIG. 4. The percentage of plans passing the IMRT QA as a function of individual and total dose discrepancy thresholds. For example, if a total dose discrepancy of 4% is selected, then the number of plans passing the IMRT QA is shown in the dashed line for different individual beam dose discrepancies.

the two-dimensional dose distribution delivered from the machine. Note that this aspect of IMRT delivery could, in principle, be performed through some other means, such as portal dosimetry or dose reconstruction methods.

Third, we have demonstrated that there is only a modest benefit from reselecting a new dose point within IMSURE. For individual beams the location of the point could critically determine the individual beam dose discrepancy even though the total dose agreement is acceptable. While there is a chance that a “better” result will be obtained in the individual beam dose discrepancy, the increased time commitment has limited gains. In our process it is not efficient for the dose calculation point to be selected with detailed examination of the field size and shape, and it is not practical to reselect the point if the IMSURE calculation fails.

Further to this, it may be justifiable to adjust the tolerance level for individual beams while maintaining the overall discrepancy level from all beams. Figure 4 displays the percentage of all (HN, GU, and LUNG) plans that pass total dose tolerances for different individual beam discrepancies. For example, if we chose a total dose discrepancy tolerance of 3% and an individual beam discrepancy of 4%, then only 20% of plans pass the monitor unit calculation test. If the individual beam discrepancy is changed from 4% to 5%, while total dose discrepancy is kept at 3%, then 32% of plans now pass, or about 12% more. Alternatively, repeating the monitor unit validation with a new dose point would result in 8% more plans passing the 3% total dose, 4% individual beam dose tolerance, as discussed in Sec. III. The implication of this is the reduced level of resources required for the IMRT QA, which in our case is a MapCheck measurement. Either a change in the individual beam tolerance or introducing a recalculation of a new dose point in IMSURE would result in fewer plans requiring a MapCheck measurement, which has a significant impact on resources in

TABLE VII. Revised tolerances, from Fig. 1, for independent monitor unit validations (IMSURE) and diode array measurements (MapCheck), based on the analysis of IMRT QA data in this article.

		MapCheck	Non-HN (%)	HN (%)
IMSURE	All sites	Total	2	2
		Beam	3	3
		Gamma	95	88

a busy clinical environment. We have chosen to use a total dose discrepancy threshold of 3% and an individual beam dose discrepancy of 5% in order to gain some efficiencies in our IMRT QA practice, while also abiding by internationally accepted guidelines in ICRU-24.⁴

Finally, we have chosen to use gamma thresholds based on our statistical findings. Specifically, tolerances for acceptable gamma values are based on the (one-sided) 95% confidence interval of gamma values that have been accepted in the past. In this case, we chose gamma values of 88, 95, and 95 for HN, GU, and LUNG cases, respectively (see Table VII). MapCheck dose thresholds given in Table VII are also based on measurement uncertainties.

One suggestion for manufacturers of TPSs and independent MU checking programs is to provide multiple points for dose calculation, or even isodose distributions. By sampling more points, there is more opportunity to select points which may not be subjected to high gradient regions that may be present in the patient, while also providing a means of validating doses to other target structures. Some manufacturers of monitor unit validation programs already provide this mechanism.

V. CONCLUSIONS

We conclude this work with the following statements:

- (a) There is no evidence that the total dose discrepancy from the monitor unit validation is dependent on the site.
- (b) While individual beam dose discrepancies may be large when performing independent monitor unit calculations, there is only a modest benefit from reselecting a new dose point.
- (c) Rather than reselecting dose points to achieve a better result from the monitor unit check, one can equally loosen individual beam dose tolerances.
- (d) The selection of tolerance values are not trivial, but a statistical analysis of QA results from IMRT plans deemed safe to deliver can be used in establishing or refining tolerances in IMRT QA procedures.

It should be noted that the IMRT QA methodology described here is neither unique nor ubiquitous, and each clinic should develop its own QA processes and tolerances based on its equipment and needs. Furthermore, passing tolerances from a QA test does not imply that an IMRT plan is safe to deliver: the experience and knowledge of a qualified medical

physicist is essential in determining whether an IMRT plan is safe for treatment. Our need was to develop meaningful tolerances in our IMRT QA processes for plans deemed safe to deliver. The development of such tolerances plays a small, but important, role in the effort to maintain the safe delivery of complex treatment plans. With the extended use of IMRT to other sites it is hoped that a similar analysis described in this article will be performed in order to provide meaningful thresholds and tolerances in IMRT QA.

ACKNOWLEDGMENTS

The authors would like to acknowledge Anula Nico, Robert Tkaczyk, Andrew Fung, Philip Au, and Sharon VanDoo-dewaard for performing the MapCheck measurements. Standard Imaging is gratefully acknowledged for their support in providing IMSURE licenses for testing and research purposes. Matt Skinner is gratefully acknowledged for setting up the MapCheck QA protocol, discussions, and contributing to the IMRT QA data. Kathy Mah is kindly acknowledged for her contributions to the IMRT QA data.

^{a)}Telephone: 416 480 6100 Ext: 1087; Fax: 416 480 6801. Electronic mail: parminder.basran@sunnybrook.ca

¹G. 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).

²Corvus Treatment Planning System, NOMOS Radiation Oncology, A Division of North American Scientific, Cranberry Township, Pennsylvania.

³J. S. Tsai, D. E. Wazer, M. N. Ling, J. K. Wu, M. Fagundes, T. DiPetrillo, B. Kramer, M. Koistinen, and M. J. Engler, "Dosimetric verification of the dynamic intensity-modulated radiation therapy of 92 patients," *Int. J. Radiat. Oncol., Biol., Phys.* **40**, 1213–1230 (1998).

⁴International Commission on Radiation Units and Measurements, *ICRU Report 24, Determination of Absorbed Dose in a Patient Irradiated by Beams of X or Gamma Rays in Radiotherapy Procedures* (ICRU, Washington, DC, 1976).

⁵QUASAR™ Penta-Guide Modus Medical Devices Inc., London, Ontario, Canada.

⁶MapCheck, Sun Nuclear Corp., Florida.

⁷MultiCube, IBA Dosimetry, Schwarzenbruck, Germany.

⁸F. M. Khan, *Treatment Planning in Radiation Oncology*, 2nd ed. (Lippincott Williams & Wilkins, New York, 2007).

⁹Y. Yang, L. Xing, J. G. Li, J. Palta, Y. Chen, G. Luxton, and A. Boyer, "Independent dosimetric calculation with inclusion of head scatter and MLC transmission for IMRT," *Med. Phys.* **30**, 2937–2947 (2003).

¹⁰D. Létourneau, 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).

¹¹D. 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).