Article Type: Research Article

Evaluation of Machine Log-Files/MC based Treatment Planning and Delivery QA as Compared to ArcCHECK QA

Carl W Stanhope[1,2], Douglas G Drake[1], Jian Liang[1], Markus Alber[3,4], Matthias Söhn[3], Charbel Habib[1], Virgil Willcut[5], Di Yan[1]

[1]Department of Radiation Oncology, William Beaumont Hospital, Royal Oak, MI, 48073, USA

[2]Department of Medical Physics, Wayne State University, Detroit, MI, 48202, USA

[3]ScientificRT GmbH, Munich, 81373, Germany

[4]Department of Radiation Oncology, Heidelberg University Hospital, Heidelberg, 61920, Germany

[5]Elekta AB, Stockholm, 113 57, Sweden

Corresponding Author

Di Yan, PhD

Department of Radiation Oncology

William Beaumont Hospital

3601 W 13 Mile Rd

Royal Oak, MI 48073

Di.Yan@beaumont.edu

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/mp.12926

Abstract:

Purpose: A treatment planning/delivery QA tool using linac log files (LF) and Monte Carlo (MC) dose calculation is investigated as a standalone alternative to phantom-based patient-specific QA (ArcCHECK (AC)).

Methods: Delivering a variety of fields onto MapCHECK2 and ArcCHECK, diode sensitivity dependence on dose rate (in-field) and energy (primarily out-of-field) was quantified. AC and LF QAs were analyzed w.r.t. delivery complexity by delivering 12x12cm static fields/arcs comprised of varying numbers of abutting sub-fields onto ArcCHECK. 11 clinical dual-arc VMAT patients planned using Pinnacle's convolution-superposition (CS) were delivered on ArcCHECK and log file dose (LF-CS and LF-MC) calculated. To minimize calculation time, reduced LF-CS sampling (1/2/3/4° control point spacing) was investigated. Planned ('Plan') and LF-reconstructed CS and MC doses were compared with each other and AC measurement via statistical (mean ± StdDev(σ)) and gamma analyses to isolate dosimetric uncertainties and quantify the relative accuracies of AC QA and MC-based LF QA.

Results: Calculation and ArcCHECK measurement differed by up to 1.5% in-field due to variation in dose rate and up to 5% out-of-field. For the experimental segment-varying plans, despite CS calculation deviating by as much as 13% from measurement, Plan-MC and LF-MC doses generally matched AC measurement within 3%. Utilizing 1º control point spacing, 2%/2mm LF-CS vs. AC pass rates (97%) were slightly lower than Plan-CS vs. AC pass rates (97.5%). Utilizing all log file samples, 2%/2mm LF-MC vs. AC pass rates (97.3%) were higher than Plan-MC vs. AC (96.5%). Phantom-dependent, calculation algorithm-dependent (MC vs. CS), and delivery error-dependent dose uncertainties were 0.8±1.2%, 0.2±1.1%, and 0.1±0.9% respectively.

Conclusion: Reconstructing every log file sample with no increase in computational cost, MC-based LF QA is faster and more accurate than CS-based LF QA. Offering similar dosimetric accuracy compared to AC measurement, MC-based log files can be used for treatment planning QA.

Introduction

Patient-specific IMRT and VMAT plan quality assurance (QA) is conventionally carried out using phantom-based systems such as Delta4, ArcCHECK, COMPASS, or Dosimetry Check. However, use of these systems is fairly time consuming, incapable of tracking delivered per-fraction dose, and incapable of determining the root-cause of failures. Phantom-based QA also ignores patient-specific anatomical variation. Thus, there exists the need for a system that is automatable, capable of tracking heterogeneous anatomical dose, and capable of root-cause analysis while maintaining or improving upon the effectiveness of phantom-based QA. Fulfilling each of these requirements, Monte Carlo (MC) based log file (LF) QA is one possible alternative. Kumar et al. shows that log files can carry out IMRT QA with the same accuracy as film.[1] Log files are also a viable method of VMAT QA for both phantom and patient geometries.[2,3] However, despite its advantages, it remains debatable whether LF QA can replace conventional techniques[4].

The primary concern with the use of log files is that they rely upon a robust machine QA protocol. For example, in a Varian machine, MLC leaf motor counts are converted into leaf position and outputted to the log file. As a result, a malfunctioning MLC motor, faulty t-nut, or loss of encoder counts may result in incorrect log file reported MLC positions. Agnew et al. illustrates this exact problem - a loose t-nut resulted in incorrect log file MLC leaf positions[5]. However, for Elekta machines, MLC positional information is determined based upon the optically measured position of reflectors on each leaf, thus eliminating the possibility of these errors assuming that the optical

reflectors have been calibrated correctly. A second concern is that LF QA ignores dose calculation error, thus placing a greater emphasis on having a robust dose calculation algorithm. In general, dose calculation algorithms have been thoroughly vetted and are accurate within 2-5% for homogeneous phantom geometry[6-7]. However, given the increased importance of dose calculation accuracy on patient anatomy, two dose calculation algorithms were assessed in this study and included in log file QA: Pinnacle's adaptive convolution-superposition (CS) algorithm and a research version of ScientificRT's SciMoCa Monte Carlo (MC) algorithm. By comparing clinical CS dose and LF-MC-based QA dose, some semblance of dose calculation accuracy verification is determined.

LF QA is susceptible to fewer error sources than AC QA. ArcCHECK error sources that log file QA is immune to include setup error and various detector inaccuracies. Several previous studies have analyzed diode sensitivity as a function of instantaneous dose rate, dose rate, energy, and field size. Saini et al. shows that diodes typically under-respond on the order of 1-4% for lower instantaneous dose rates (~25 cGy/min) as opposed to typical clinical dose rate (> 200 cGy/min).[8] Letourneau et al. shows MapCHECK diodes under-respond by up to 2% at 50 MU/min compared to 600 MU/min.[9] Diode energy dependence has been explored by several studies, albeit to a lesser extent.[10-12] Yin et al. shows that unshielded Scanditronix diodes at 15cm depth can over-respond by as much as 15% to lower energy patient-scattered radiation. However, for out-of-field diodes where the radiation spectrum can be significantly softer, one should expect an even greater relative over-response by the diodes. For example, Rickner et al. theorizes that diodes will over-respond to 100keV and 500keV photons by 70% and 30% respectively.[13] Czarnecki et al. showed field size calibration factors to vary by up to 2.0% between 1x1cm and 10x10cm field sizes.[14] However, after applying field size correction factors, Chaswal et al. shows diode dose error as a result of field size dependence to be within 0.4% for 5x5cm to 20x20cm fields.[15] To double check these sensitivity dependencies, as well as to fully understand their exact effects on our system, several of these error

sources were re-evaluated. Furthermore, their effect on measurement, calculation, and overall impact on QA results were evaluated.

Another caution of AC QA is on the significance of gamma metrics. Research has identified gamma passing rates to be relatively insensitive to clinically relevant delivery errors.[16] Given this concern, there has been a general movement towards DVH-based QA metrics.[17] However, when it comes to DVH-based metrics, phantom-based systems tend to fall short. By comparing 3DVH and log-file reconstructed doses to the original TPS dose, Tyagi et al. shows that log files are better able to reproduce delivered dose on the patient anatomy, and thus more suitable for DVH-based QA.[3] 3DVH is SunNuclear's ArcCHECK-specific method of reconstructing 3D dose on the patient geometry from measurement. Subsequently, log files will also be more suitable in adaptive radiotherapy programs. Log file QA on patient-anatomy is left as a future study.

This is largely a comparative study, seeking to determine whether log file QA can provide ArcCHECK equivalent accuracy. First, accuracy of the log file is investigated and the Monte Carlo beam model is validated. Next, the log file dose reconstruction process along with strategies for reducing the computational cost of said process was detailed. Calculated vs. AC-measured dose differences were quantified as functions of diode dose rate and energy dependence. LF and AC QAs were evaluated for experimental plans of varying delivery complexity. Lastly, clinical LF QA and AC QA results were compared.

2. Methods and Materials

For this study, an Elekta Infinity linac equipped with the Agility beam-limiting device was utilized. Plans were selected from previously treated patients that were planned and optimized using the Pinnacle3 v9.8 treatment planning system. Eleven dual-arc VMAT patients were selected for this study (9 H&N, and 2 low dose rate brain). The head and neck cases were selected for their high degree of complexity and somewhat lower gamma pass rates. The two brain cases were selected to compare how log file and ArcCHECK QAs respond to low dose rate deliveries. All 22 arcs were delivered onto Sun Nuclear's ArcCHECK phantom and log files recorded. All beams were flattened, 6 MV.

To compare AC QA to LF QA, five dose distributions were determined for each plan: Plan-CS, Plan-MC, LF-CS, LF-MC, and AC. 'Plan' denotes the dose distribution was calculated from the original clinical treatment plan file optimized in Pinnacle, whereas 'LF' doses were calculated from log files. CS and MC denote the two dose calculation algorithms. 'AC' denotes ArcCHECK measurement. To begin, Plan-CS is recalculated using Monte Carlo to get Plan-MC. Next, each plan was delivered on the ArcCHECK phantom, measurement taken, and log files recorded. For LF-CS dose calculation, in-house code was used to reconstruct the log file beams in Pinnacle with a reduced number of log file samples (see Section 2.2). For MC dose calculation, the DICOM toolkit (DCMTK) library by OFFIS was utilized to convert log files into DICOM RTPLANs containing every log file sample with non-zero MU. DCMTK is a collection of open-source C/C++/ANSI libraries and applications implementing the DICOM standard. Whereas LF-MC is calculated using every useful log file sample, LF-CS is calculated using a reduced number of log file samples.

2.1 Preliminary Validations

Elekta's Log File Convertor for Integrity R3.2 records linac delivery parameters (dose rate, gantry/collimator angle, leaf/collimator positions, MU) every 40ms. MU readings are determined via redundant MU chambers. Jaw positions are determined electronically via two redundant potentiometer voltages. MLC positions are determined from the video image positions of optical reflectors on each MLC leaf. The accuracy of jaw and leaf positioning, as represented in the log file, hinges on machine calibration and QA. Jaw and MLC calibration is carried out annually as well as after any relevant field service. Calibration is performed using the internal Calibration Workflows contained in the Integrity R3.2 TCS, specifically the optical, diaphragm, and leaf workflows. For monthly QA, picket fence patterns are acquired with the iViewGT EPID and a BB array phantom aligned to the crosshairs, and compared to baseline images acquired just after Calibration Workflows, to verify that leaves and jaws are within 1mm of their baseline positions. If any of the collimation components appear to be out of tolerance, the appropriate Calibration Workflow is re-run and the appropriate picket fence test repeated.

The ArcCHECK phantom was supported by the iBeam evo H&N extension. To account for this added attenuation, a couch ROI was added to the phantom geometry. To validate our couch ROI, a stack of solid water was placed on top of the attached H&N extension and irradiated from beneath. Ion chamber measurements were taken at a depth of 10cm for seventeen 10x10cm static fields with 10º spacing between each field. Gantry angles ranged from 100 to 260 degrees. Field sizes are given in centimeters throughout this study.

The Monte Carlo algorithm utilized by this study was ScientificRT´s SciMoCa algorithm. SciMoCa shares its fundamental concept with the voxel Monte Carlo (VMC) family of codes, e.g. VMC++ or XVMC.[18-21] SciMoCa is also based off of a more recent series of papers by Sikora et al.[22-24] The beam model utilized is not clinically commissioned and is currently used for research purposes only. For all calculations, unless otherwise stated, particle histories were simulated to achieve 1% dosimetric variance at 70% max dose. Beam model data for the 6MV Elekta Agility linac was acquired using IBA's Blue Phantom^2. Inline/crossline profiles, PDDs, and output factors were acquired for various field sizes as shown in Table 3. Output factors less than 5x5, PDDs less than 5x5, and profiles less than 20x20 were acquired using a Sun Nuclear EDGE diode. All other measurements were acquired using an IBA CC13 ion chamber (0.13cm^3 volume, 5.8mm length, 3.0mm radius). Crossplane profiles were offset half a leaf width (2.5mm) in the inplane direction to minimize the effect of interleaf leakage and to scan close to the center of the Agility leaf tips which appear tapered in EPID images. All profiles were measured at 1.5/5/10/20 cm depths. Ideally, large field profiles should be acquired using a microdiamond or microchamber. For the ≥20x20cm profiles taken with a CC13 chamber, substantial penumbral blurring and thus poorer agreement between measurement and calculation is expected. The resulting beam model was validated by comparing calculated doses to each of the aforementioned measurements. All calculations utilized 1mm dose grid resolution in the direction of measurement. In regards to commissioning and QA of treatment planning dose calculation, Smilowitz et al. recommends low-gradient in-field regions match within 1.5%, penumbra agree within 3mm DTA, and out-of-field measurements agree within 3% of max dose.[25] Therefore, the fraction of calculations that meet each of these criteria was determined. Out-of-field doses were also compared using more strict 2% and 1% criteria. Low-gradient was defined as less than 1cGy/mm. Out-of-field was defined as less than 10% of CAX dose. Additional per-field calculations included percent difference in output factor, mean percent difference in PDD from 0-35cm, mean percent difference in beam width (50%-50%) over all eight profiles, and mean percent difference in

penumbral width (80%-20%) over all eight profiles. Percent PDD differences were calculated relative to maximum dose. Due to insufficient scan length for the 40x40 field, out-of-field dose and penumbral width was left uncalculated and unanalyzed.

2.2 Strategies for Reducing the Computational Cost of Log File QA

Log files are reconstructed by converting n number of log file samples into n VMAT control points, where a control point is a planning parameter that defines the linac state (collimation positions, delivered MU, gantry angle) at some point in the planned treatment delivery. However, Pinnacle places an upper limit of 359 control points on each field, thus limiting the number of log file samples that can be reconstructed to 359, despite thousands of samples per log file. In addition to this software limit, utilizing every one of thousands of log file samples would result in extremely long CS calculation times. This is because CS dose calculation time scales linearly with the number of control points for the CS algorithm. There are at least a few ways to reduce this cost: (1) utilize CPU and/or GPU parallelization, (2) utilize a dose calculation algorithm whose computational cost is independent of the number of control points (e.g. Monte Carlo), and/or (3) reduce the number of log file samples reconstructed. For this study, methods 2 and 3 were investigated. CS dose calculation utilized Oracle's Sun Server X4-2. MC calculation utilized an Intel i7-3770 and 8GB of RAM.

The effect of reducing the number of log file samples reconstructed was investigated by utilizing 1/2/3/4º control point spacing. For example, a VMAT arc starting at -179º, ending at 179º, and reconstructed with a 3º control point spacing would have the following control points: -179.xº, -176º, -173º, … 175º, 178º, 179.xº. Variable x stands for the exact gantry angle at which radiation started and stopped delivering radiation according to the log file. In addition to control point

spacing, the effect of dose grid resolution (2/3/4 mm) on log file reconstructed dose accuracy was analyzed. Reconstructive effectiveness was quantified by comparing log file gamma pass rates (LF-CS vs. AC) with pass rates of the original plan (Plan-CS vs. AC). This metric measures dosimetric improvement due to log files accounting for machine delivery errors. Because LF and AC both account for delivery errors, this metric should be greater than zero.

2.3 Discrepancy between Calculated and Measured Diode Doses

2.3.1 In-Field Discrepancy: Dose Rate Effect

Calculated (Plan-CS) vs. measured (MapCHECK2 and ArcCHECK) doses were compared for a variety of dose rates (35, 70, 140, 280, and 570 MU/min). 150MU was delivered onto the MapCHECK2 diode array for 3x3cm and 20x20cm field sizes and each dose rate. Constant machine output was verified by placing 2cm solid water on top of the MapCHECK and inserting an ion chamber at CAX. In-field and out-of-field dose rate effects were investigated; penumbral diodes directly on the field edge were excluded from the analysis. MapCHECK2 was initially utilized in order to best isolate dose rate effects; the simpler device geometry minimizes any possible spectral effects. It is important to note that the same diodes are used in both MapCHECK2 and ArcCHECK systems. Systematic effects of dose rate on AC QA and LF QA were assessed by delivering a 25x25arc onto the phantom using 70 and 570 MU/min.

Dose rate is dynamically determined based upon gantry, jaw, and leaf speed limitations. Thus, for deliveries with fixed jaws and leaves, delivered dose rate should be calculable as a function of MU, gantry excursion (gantry degrees subtended), and maximum gantry speed (6°/s). To prove this, 23 10x10 arcs (see Table 1) with varying gantry excursions and MU were delivered on the ArcCHECK phantom. 'Calculated' dose rate matched log file data for all cases. Next, percent

difference in calculated (Plan-CS) and measured diode dose was tracked for each of the 23 fields and analyzed as a function of 'calculated' dose rate.

2.2.2 Out-of-Field Discrepancy: Energy Effect

Percent difference in calculated versus measured diode dose was analyzed as a function of off-axis distance to determine the effects of off-axis softening on AC QA. Because off-axis spectra vary with field size, static 2x2, 10x10 and 25x25 arcs were assessed. 0% and 10% dose thresholds were utilized.

2.3.3. Delivery Complexity

The effect of beam segment complexity on calculated vs. measured dose difference was investigated for 42 fields delivered onto the ArcCHECK phantom. 100MU static fields were delivered at gantry 0° and -90° for 1x7, 7x1, 2x8, 8x2, 3x5, 5x3, 5x10, 10x5, 2x2, 3x3, 5x5, 10x10, 15x15, and 20x20 field sizes. 300MU arcs from -90° to 90° were delivered for the same field sizes. Two metrics, MPDD (mean percent dose difference) and MAPDD (mean absolute percent dose difference), were utilized. These metrics work by comparing any two of the five dose distributions described in Section 2 at each of 1386 ArcCHECK diode positions. Equation 1 shows how to calculate MAPDD for LF-MC vs. AC

$$MAPDD = Mean\left(\left|\frac{LFMC_i - AC_i}{AC_i} * 100\right|\right), \qquad \text{Eq. 1}$$

where i is the i-th diode (i = 1-1386) such that AC_i is greater than some threshold percentage of max(AC). MPDD is simply MAPDD without taking the absolute value.

LF QA and AC QA should agree with one another even for complex fields. To verify this is the case, we divided a 12x12 static field and 12x12 arc into 1/4/9/16/36 sub-fields (i.e. 1 12x12, 4 6x6, 9 4x4, 16 3x3, and 36 2x2 fields/arcs). Each static field was setup at gantry 0. Arcs were a full 358° CW or CCW. Plans were generated in Pinnacle on a 'prescription' geometry and then re-calculated and delivered on the ArcCHECK geometry. The prescription geometry, shown in Figure 1, consisted of three structures. A 20x20x20 solid water cube centered on the isocenter serves as the patient volume. A 10x10x10 'Target' ROI centered on isocenter was utilized for DVH matching purposes; all static fields and arcs had matching DVHs. Lastly, an isocenter-centered 10x10x1 (1cm perpendicular to gantry 0° CAX) 'Prescription' ROI was created. Sub-field beam weights were optimized to achieve a flat beam profile at 10cm depth for the full field; dose was optimized to the Prescription ROI using a 1Gy minimum DVH objective and 1.1Gy maximum DVH objective of equal priority. Static field beam weights were copied to their corresponding arc fields. Each set of sub-fields was delivered onto the ArcCHECK phantom. Sub-field measurements were combined such that ten combined-field measurements (5 static field and 5 arc) were created. Differences in AC, Plan-CS, Plan-MC, and LF-MC doses were assessed as a function of the size and number of segments delivered.

2.4 Comparing Log File and ArcCHECK Techniques

Plan, LF, and AC dose distributions are subject to varying sources of dose uncertainty and share various correlated factors. AC QA compares 'Plan' and AC doses; discrepancies stem from phantom-dependent inaccuracies (e.g. setup error, diode miscalibration/drift, diode sensitivity dependencies, and physical diode limitations), dose calculation inaccuracy, delivery error, and

variation in beam consistency (drift in beam output/symmetry). Comparing LF and AC doses, dosimetric uncertainty decreases w.r.t. Plan vs AC. This decrease in uncertainty stems from the fact that LF and AC doses both account for various delivery errors. For LF QA, LF and Plan doses are compared. Conventionally this means comparing LF-CS to Plan-CS to isolate delivery error. However, dose calculation accuracy should also be verified. By comparing LF-MC to Plan-CS, both delivery error and dose calculation accuracy are investigated. It is important to note that whereas AC accounts for variations in beam consistency, neither LF nor Plan doses do. This means that LF QA ignores variation in beam output.

AC and LF QAs were carried out for 11 clinical dual-arc VMAT patients (9 H&N, 2 low dose rate brain). Dose was determined at each of ArcCHECK's 1386 diode locations for each arc and each of the five aforementioned dose distributions. Various pairs of these five plans were compared. Plan pair comparisons included (1) Plan-CS vs. AC, (2) Plan-MC vs. AC, (3) LF-CS vs. AC, (4) LF-MC vs. AC, (5) Plan-MC vs. Plan-CS, (6) LF-CS vs. Plan-CS, (7) LF-MC vs. Plan-MC, and (8) LF-MC vs. Plan-CS. Three comparisons metrics were utilized. First, 2%/2mm and 1%/1mm global gamma pass rates were calculated using 10% (conventional) and 85% (in-field/target approximation) dose thresholds. Second, global percent dose difference was calculated at each diode position and averaged over all diodes (1386 diodes/arcs * 22 arcs). 'Global' indicates that percent differences have been scaled relative to the maximum dose value being compared to, e.g. for X vs. Y where X and Y are two 1386 element dose matrices, global percent diode dose difference equals

$$mean\left(\frac{\boldsymbol{X}-\boldsymbol{Y}}{\max(\boldsymbol{Y})} * 100\right). \qquad \text{Eq. 2}$$

This second metric was calculated in order to isolate and quantify systematic and statistical dose uncertainties due to delivery error, dose calculation difference, and the ArcCHECK phantom as well as to confirm the gamma analysis results.

3. Results and Discussion

3.1 Preliminary Validations

For the Elekta Infinity/Agility linac used in this study, Calibration Workflows resulted in leaf, leaf bank, and jaw positions agreeing with nominal positions within approximately 0.3mm each (0.1mm RMS). Despite the initial accuracy of these calibration processes, agreement with nominal positions can degrade over time. Monthly picket fence/phantom tests nominally ensure collimation components are within 1mm of post-calibration baseline positions, however they can be somewhat subjective and don't always directly isolate individual leaf position errors. For these reasons, it may be useful to perform Calibration Workflows more often if utilizing LF QA clinically, or to implement more objective collimator component QA testing. The necessary increased frequency of calibration or exploration of other QA methods were not investigated as part of this study

In regards to the added couch ROI, measured and calculated dose values (with and without the couch) agreed within 0.25% for all fields (0.10% on average).

Verification data comparing MC calculation to measurement is shown in Table 3. The SciMoCa model utilized herein met each of the aforementioned AAPM recommended criteria 100% of the time. Furthermore, 96.5% of out-of-field measurement datum matched calculation within 2%. For field sizes ≤ 20x20cm, 97% of out-of-field measurement data matched calculation with 1%. For

the larger 30x30cm field size, 44% of measurement data matched within 1%. On average, over all field sizes, calculated output factors deviated from measurement by 0.20 ± 0.17% and PDDs agreed within 0.65±0.19%. On average, beam width varied by 0.54 ± 0.38%. Penumbral width varied by 1.1 ± 0.5% of the field size. For field sizes less than or equal to 10x10cm this difference was within 0.5mm, however penumbral width varied by almost 3mm for the 30x30cm field; these absolute differences are displayed in parenthesis in Table 3. As hypothesized, deviation in penumbral width was larger for the profiles taken with the CC13 chamber (20x20cm and 30x30cm) than the EDGE diode (< 20x20cm). This larger deviation is expected to result from penumbral blurring and not poor beam modeling.

3.2 Strategies for Reducing the Computational Cost of Convolution-Superposition

Figure 2 displays change in pass rate ((LF-CS vs. AC) minus (Plan-CS vs. AC)) as a function of control point spacing and dose grid resolution for 10 (8 H&N, 2 Brain) of the 22 clinical VMAT arcs used in this study. To make the plot clearer, error bars were plotted only in one direction for each curve. For the dose grid resolution curve, 2° control point spacing is utilized. 2mm resolution offers a 1.2±0.5% increase in pass rate compared to 3mm. For all subsequent studies, a 2mm dose grid resolution was utilized. Increased control point spacing resulted in lower LF-CS pass rates. Yielding a negative change in pass rate, control point spacing ≥ 2° was deemed unacceptable. 1° LF-CS spacing was used throughout the rest of this study. Similarly, Barbeiro et al. found log files reconstructed with ~3 times as many control points as the TPS agreed better with film than those reconstructed with the TPS' discretization.[26]

Using 2mm dose grid resolution and 1° control point spacing, computational cost for CS was 12 seconds per control point. This corresponds to 18 minutes for a 358° arc planned using 4 degree control point spacing or 72 minutes for a log file plan reconstructed using 1 degree control point spacing. In comparison, MC dose calculation took ~5 minutes to calculate log file dose using every control point. This discrepancy in computational cost stems primarily from CS's computational cost being dependent on the number of control points, but also from the aforementioned differences in hardware.

3.3 Discrepancy in Calculated vs. Measured Diode Dose

3.3.1 In-Field Discrepancy: Dose Rate Effect

Figure 3 plots in-field reduction in diode measurement with decreasing dose rate w.r.t 570 MU/min for 3x3cm and 20x20cm MapCHECK2 fields, a 25x25cm ArcCHECK field, and the 23 10x10cm ArcCHECK fields in Table 1. For the MapCHECK2 measurements, low dose rate (35 MU/min) measurements were -1.2±0.4% and -1.6±0.4% lower than the high dose rate (570 MU/min) measurements for 3x3cm and 20x20cm fields respectively. The trend shown in Figure 3 is similar to Letourneau et al's, although slightly less in magnitude; Letourneau measured an under-response of -2.0% at 50MU/min.[9] Variance in dose rate effect with field size is believed to result from the reduced signal observed with smaller field sizes. Jursinic's 'hypothesized reaction scheme' would suggest smaller signals are less susceptible to the dose rate effect.[27] This would also explain why out-of-field diodes registered identical doses for low and high dose rates. Delivering a static 25x25cm arc so that all diodes were in-field, ArcCHECK measurement behaved similarly to MapCHECK2; measurement decreased with decreasing dose rate. Lastly, delivered dose rate was calculated from MU and gantry excursion for the 23 10x10cm arcs shown in Table 1. This 'calculated' dose rate matched the actual

log file determined dose rate for all cases. Subsequently, as displayed by the solid line in Figure 3, variation in in-field diode measurement as a function of dose rate agreed well with previous results.

To better understand the clinical effect of dose rate dependence on AC QA, MC and CS calculated doses were compared to AC measurement for each diode in an open 25x25cm arc for both 70 and 570 MU/min dose rates. Plan-CS vs. AC percent dose differences were -2.1 ± 0.5% and -0.7 ± 0.5% at 570 and 70MU/min dose rates respectively. In comparison, Plan-MC vs. AC percent dose differences were 0.3 ± 0.8% and 1.7 ± 0.8%. Interestingly, because MC doses were generally higher than CS doses and higher dose rates resulted in greater measurement, MC dose differences were smaller for high dose rates, while CS dose differences were smaller for low dose rates. In comparison to AC QA, LF QA should be invariant w.r.t. dose rate. To verify whether this is indeed the case, percent difference in 70MU/min vs. 570MU/min LF-MC dose was calculated for the same 25x25cm arc; per-diode dose difference was 0±0.85% and 0±0.21% when calculated with 1% and 0.25% MC dose uncertainties, respectively. This shows that whereas AC QA is sensitive to dose rate and could yield false positives/negatives, LF QA does not have this issue.

3.3.2 Out-of-Field Discrepancy: Energy Effect

Figure 4 displays percent difference in calculated (Plan-CS and LF-MC) versus AC measured diode dose as a function of off-axis distance for a 10x10cm 358° arc. For the in-field region, measurement and calculation agreed well. In the penumbral region, significant dose difference is seen. Penumbral differences are attributed to imperfect geometry (e.g. imperfect linac isocentricity, imperfect leaf calibration, and setup error) and physical diode limitations (e.g. finite size), as well as the two issues discussed in the next paragraph.

Out-of-field, calculation is much lower than measurement. This disagreement is comprised of two major components. First, the calculation model could be underestimating collimator scatter and head leakage at greater off axis distances. For CS, this is typical of many institutions and is known to be the case at our institution as well.[28-32] Second, diodes could be over-responding to the low energy patient-scattered photons. Rickner et. al theorizes 30% and 70% increases in diode sensitivity to 0.5MeV and 0.1MeV photons.[13] The typical energy range of patient-scattered photons is right in this range.[33] Thus, one can expect an approximate 50% over-response to low energy photons, and an approximate 30% net over-response after multiplying by the low energy component of the fluence. Given past literature, and the results of figure 4, a relative 10-30% over-response by out-of-field diodes is concluded for this 10x10 field. It is important to note this is for a 0% dose threshold. A 10% dose threshold is explored below to determine clinically relevant effects only.

Field size plays a critical role in determining the energy spectrum of photons at each diode. Here, the effect of field size on energy spectrum, and subsequently diode sensitivity was studied. Three static 200MU arcs (2x2cm, 10x10cm, 25x25cm) were delivered on the ArcCHECK phantom. Figures 5a and 5c plot histograms of percent difference in calculated versus measured diode dose for each field size and for CS and MC respectively. Figures 5b and 5d plot the same histograms with a 10% dose threshold applied. Histograms consisted of a tight primary peak and prolonged 'spectral' out-of-field distribution for both CS and MC.

As discussed previously, this 'spectral' distribution results from inaccurately calculated collimator and head leakage as well as diode over-response to patient-scattered radiation. Dose differences increased with decreasing field size. By shrinking the field, the distance from the field edge to distal diodes increases. This results in more medium to scatter off and thus a softer photon spectrum and greater diode over-response. As illustrated by the relative positions of their out-of-

field distributions as well as the integral area of these distributions, MC yielded better agreement with out-of-field measurement than CS. For the 2x2cm field size, a good portion of CS's spectral distribution is centered around -90%. This corresponds to a 10 times greater measured versus calculated dose. This massive difference is in part due to an underestimation of out-of-field dose by the CS algorithm, but also largely due to the extremely low dose – a low absolute dose will yield a larger relative dose difference. For the 2x2cm MC field, measurement is 70% higher than calculation. Given the diode over-response values mentioned in 3.3.2, 70% over-response is reasonable. It is important to note that some portion of this discrepancy will be due to calculation inadequacies as quantified in Section 3.1.

Despite the very large discrepancies seen in Figures 5a/c, when a standard 10% dose threshold is applied (Figures 5b/d), much of this out-of-field component is eliminated. For instance, for the 2x2cm field where the absolute quantity of patient scatter is minimal, there is a miniscule out-of-field distribution. Thus, diode energy dependence is going to result in very few failing diodes. However, for the 10x10cm field where the absolute magnitude of patient scatter is greater, the number of failing out-of-field diodes that are above the measurement threshold is more noticeable - energy effects are seen on the order of 3-5% of the maximum measured dose. Although systematic uncertainties due to energy and dose rate will cancel out to some degree, when compared to standard 2-3% gamma criteria, 3-5% over-response to lower energy radiation remains worrisome.

3.3.3 Delivery Complexity

Static fields/arcs with a wide variety of field sizes were delivered onto the ArcCHECK phantom and log files reconstructed using Monte Carlo. LF-MC calculated dose is compared to AC diode measurement using MPDD and MAPDD along with an 85% dose threshold. Figures 6a and 6b

display MPDD and MAPDD respectively as a function of field size ($FS_X \cdot FS_Y$). Smaller fields yielded greater MAPDD values, indicating greater LF-MC vs. AC disagreement for smaller field sizes. Negative and positive MPDD values were seen because an 85% dose threshold will include some field penumbra. Consequently, any small shift in diode position due to setup error or MLC position due to leaf calibration error can result in large differences between calculation and measurement. Depending on the position of the diode w.r.t. CAX (e.g. $\pm\hat{x}/\hat{y}$), direction of setup error, and direction of leaf calibration error, measured diode dose may increase or decrease w.r.t. calculation.

Delivering various combinations of abutting fields onto the ArcCHECK phantom, calculated (LF-MC, Plan-MC, Plan-CS) vs. AC measured dose difference was evaluated as a function of segment size and quantity. Table 4 displays MPDD values for each static and arc field arrangement using both 10% and 85% dose thresholds. As the number of segments increased, Plan-CS increasingly deviated from measurement (max 13%). This Plan-CS dose difference likely occurs due to the out-of-field dose mismatch seen between CS and AC in section 3.3.2. This discrepancy is seen not only for the 10%, but also the 85% dose threshold. This is because what is considered out-of-field for one segment may be considered in-field for the combined 12x12cm field. Plan-MC and LF-MC matched measurement within a few percent for all cases. LF-MC and Plan-MC deviated from measurement by 1.42 ± 1.2% and 1.56 ± 1.2% on average respectively. This shows improvement by accounting for delivery errors was only 0.14%.

3.4 Comparing LF QA and AC QA

Gamma pass rates were calculated for each of 22 clinical beams and each dose-pair shown in Table 5 using a 10% dose threshold. Comparing #3 to #1, we saw lower LF-CS than Plan-CS pass rates, suggesting even 1° control point spacing was insufficient. Comparing #4 to #2, reconstructive accuracy was evaluated for the full 25Hz log file; pass rates were identical or improved for 17/18

(2%/2mm) and 14/18 arcs (1%/1mm). High pass rates for #5 and #6 indicate both dose calculation differences and machine delivery errors were small. LF QA results (#7) were similar to AC QA results (#1/2).

Values are calculated for the same 22 beams for each plan pair in Table 6. Comparison #5 isolates differences in dose calculation algorithm; MC and CS agreed within 1-2%. Comparisons #6/7 isolate delivery errors for CS and MC respectively; the effect of delivery error was ~0±1%. Comparing #3 to #1 and #4 to #2, the effects of delivery error are again seen to be very small for this Agility system. #2/4/5 show MC dose was systematically 1.2% higher than CS dose when utilizing an 85% dose threshold. This is believed to result from slight MC vs. CS beam model differences stemming from how CS dose is normalized to a 10x10cm field at 10cm depth and MC is not. Rather, the MC beam model was simply designed to best match calculation to measurement for the greatest amount of PDDs and dose profiles. Accounting for both delivery error and calculation difference, comparison #8 is this study's standard LF QA comparison. Summing systematic differences linearly and statistical differences in quadrature, #8 equals #5 plus #7 for both dose thresholds. Eliminating phantom error sources resulted in substantially reduced uncertainty; values for #5-8 were ~0.8 ± 1.2% lower than #1-4. We conclude phantom-based uncertainty is ~0.8 ± 1.2%. It is important to note that phantom-dependent uncertainty includes not only measurement and setup uncertainties, but also any potential beam inconsistencies due to imperfect machine QA that the log file does not track. Gamma and statistical analyses both led to the same conclusions.

Summary

As discussed briefly in the introduction, there exist distinct advantages and disadvantages of the two QA systems. The key advantage of ArcCHECK is that it directly measures dose; physicists can trust measurement within the ranges of uncertainty set forth in this study. However, it is difficult to determine the cause of failing pre-treatment QA. In comparison to AC QA, LF QA is unaffected by phantom-dependent uncertainties and capable of isolating the cause of dose difference. The downside of LF QA is that it relies heavily on machine QA and accurate dose calculation. Log file reconstructed dose would be more accurate if BLD calibration was carried out more often. Reconstruction accuracy could be further improved by better maintaining beam output and symmetry whether through more frequent QA or tighter tolerances. Furthermore, LF QA is unable to account for dose calculation differences when utilizing the same algorithm as the TPS. Thus, LF QA should ideally utilize an independent dose calculation algorithm that has been thoroughly vetted as detailed in TG-106.[33] In the case of dose deviance between the two algorithms, ArcCHECK can be utilized as a second check. Assuming a non-MC TPS dose calculation engine, Monte Carlo is well suited for LF QA. Conventionally, MC is deemed accurate, but slow. However, reconstructing every log file sample with no increase in cost, MC-based LF QA is not only accurate, but exceptionally fast. Once properly setup, LF-MC QA is a fully automatable patient-specific QA technique that can quickly and accurately calculate per-fraction delivered dose. However, given the increased reliance on machine QA and dose calculation accuracy, LF QA should at least initially be used only as a supplement to current QA systems.

Conclusion

Utilizing control point spacing greater than 1° resulted in substantially degraded dosimetric accuracy. CS-based LF QA utilizing 1° control point spacing and MC-based LF QA utilizing the full 25Hz log file both yielded equivalent results to AC QA.

Calculation and ArcCHECK measurement differed by up to 1.5% in-field due to variation in dose rate and up to 5% out-of-field. For highly segmented experimental plans, despite CS calculation deviating by as much as 13% from measurement, Plan-MC and LF-MC doses generally matched AC measurement within 3% (mean 1.56% and 1.42% respectively). Carrying out AC QA and LF QA for 22 clinical VMAT arcs, phantom-dependent, calculation algorithm-dependent, and delivery error-dependent dose uncertainties were found to be 0.8 ± 1.2%, 0.2 ± 1.1%, and 0.1 ± 0.9% respectively. It follows that by eliminating phantom-dependent uncertainty, LF QA is theoretically more accurate than AC QA. However, without a gold standard to compare to, we merely conclude that LF QA and AC QA offer similar dosimetric accuracy.

Acknowledgements: The authors would like to thank Elekta AB for their financial support of this study.

Conflict of Interest: This work is partially supported by an Elekta Research Grant

Figure 1. (left) The 'prescription' geometry. The teal isodose line displays the flatness of the combined 12x12 field when (right) 36 abutting 2x2 fields are delivered

Figure 2. Difference in 2%/2mm pass rate ((LF-CS vs. AC) minus (Plan-CS vs. AC)) is plotted as a function of varying dose grid resolution (fixed 2° control point spacing) and control point spacing (fixed 2mm dose grid resolution). For visualization purposes, error bars are plotted only in one direction.

Figure 3. Reduction in in-field diode measurement with decreasing delivered dose rate is plotted for (1) 3x3cm and 20x20cm MapCHECK2 fields, (2) a 25x25cm ArcCHECK arc, and (3) the 23 10x10cm arcs shown in Table 1. Reduction is normalized to 570 MU/min.

Figure 4. Calculated vs. measured dose differences are plotted as a function of axial distance from isocenter for a 10x10 arc using Plan-CS and LF-MC calculated doses.

Figure 5. Histograms of percent diode dose difference (calculated vs. measured) for CS and MC with respect to ArcCHECK measurement for both 0% and 10% dose thresholds.

Figure 6. LF-MC and AC doses are compared for varying field sizes using an 85% dose threshold. a) Mean Percent Dose Difference (MPDD) and b) Mean Absolute Percent Dose Difference (MAPDD) are plotted against FSx*FSy for all the aforementioned fields in Section 2.3.3.

Bibliography

1. Dinesh Kumar M, Thirumavalavan N, Venugopal Krishna D, Babaiah M. QA of intensity-modulated beams using dynamic MLC log files. *J Med Phys* 2006;31(1):36-41. doi: 10.4103/0971-6203.25668

2. Defoor DL, Vazquez-Quino LA, Mavroidis P, Papanikolaou N, Stathakis S. Anatomy-based, patient-specific VMAT QA using EPID or MLC log files. *J Appl Clin Med Phys.* 2015 May 8;16(3):5283. doi: 10.1120/jacmp.v16i3.5283.

3. Tyagi N, Yang K, Yan D. Comparing measurement-derived (3DVH) and machine log file-derived dose reconstruction methods for VMAT QA in patient geometries. *J Appl Clin Med Phys*. 2014 Jul 8;15(4):4645. doi: 10.1120/jacmp.v15i4.4645.

4. Childress N, Chen Q, Rong Y. Parallel/Opposed: IMRT QA using treatment log files is superior to conventional measurement-based method. *J Appl Clin Med Phys.* 2015 Jan 8;16(1):4-7. doi:10.1120/jacmp.v16i1.5385

5. Agnew A, Agnew CE, Grattan MW, Hounsell AR, McGarry CK. Monitoring daily MLC positional errors using trajectory log files and EPID measurements for IMRT and VMAT deliveries. Phys Med Biol. 2014 May 7;59(9):N49-63. doi: 10.1088/0031-9155/59/9/N49.

6. Haga A, Magome T, Takenaka S, et al. Independent absorbed-dose calculation using the Monte Carlo algorithm in volumetric modulated arc therapy. *Radiat Oncol*. 2014 Mar 14;9:75. doi: 10.1186/1748-717X-9-75.

7. Butson MJ, Elferink R, Cheung T, et al. Verification of lung dose in an anthropomorphic phantom calculated by the collapsed cone convolution method. *Phys. Med. Biol.* 2000;45(11).

8. Saini AS, Zhu TC. Dose rate and SDD dependence of commercially available diode detectors. *Med Phys*. 2004 Apr;31(4):914-24.

9. Létourneau D, Gulam M, Yan D, Oldham M, Wong JW. Evaluation of a 2D diode array for IMRT quality assurance. *Radiother Oncol*. 2004 Feb;70(2):199-206.

10. Saini AS, Zhu TC. Energy dependence of commercially available diode detectors for in-vivo dosimetry. *Med Phys*. 2007 May;34(5):1704-11.

11. Griessbach I, Lapp M, Bohsung J, Gademann G, Harder D. Dosimetric Characteristics of a new unshielded silicon diode and its application in clinical photon and electron beams. *Med Phys.* 2005 Dec;32(12):3750-4. DOI: 10.1118/1.2124547

12. Yin Z, Hugtenburg RP, Beddoe AH. Response corrections for solid-state detectors in megavoltage photon dosimetry. *Phys. Med. Biol.* 2004 Aug;49(16):3691-3702.

13. Rikner G, Grusell E. General specifications for silicon semiconductors for use in radiation dosimetry. *Phys. Med. Biol.* 1987;32(9).

14. Czarnecki D, Zink K. Monte Carlo calculated correction factors for diodes and ion chambers in small photon fields. *Phys Med Biol.* 2013 Apr 21;58(8):2431-44. doi: 10.1088/0031-9155/58/8/2431

15. Chaswal V, Weldon M, Gupta N, Chakravarti A, Rong Y. Commissioning and comprehensive evaluation of the ArcCHECK cylindrical diode array for VMAT pretreatment delivery QA. *J Appl Clin Med Phys.* 2014 Jul 8;15(4):4832. doi: 10.1120/jacmp.v15i4.4832.

16. Nelms BE, Zhen H, Tomé WA. Per-beam, planar IMRT QA passing rates do not predict clinically relevant patient dose errors. *Med Phys.* 2011 Feb;38(2):1037-44. DOI: 10.1118/1.3544657

17. Zhen H, Nelms BE, Tome WA. Moving from gamma passing rates to patient DVH-based QA metrics in pretreatment dose QA. *Med Phys.* 2011 Oct;38(10):5477-89. doi: 10.1118/1.3633904.

18. Kawrakow I, Fippel M, Friedrich K. 3D electron dose calculation using a Voxel based Monte Carlo algorithm (VMC). *Med Phys.* 1996;23(4):445-457. DOI: 10.1118/1.597673.

19. M Fippel. Fast Monte Carlo dose calculation for photon beams based on the VMC electron algorithm. *Med Phys.* 1999 Aug;26(8):1466-75. DOI: 10.1118/1.598676.

20. Kawrakow I. Accurate condensed history Monte Carlo simulation of electron transport. I. EGSnrc, the new EGS4 version. *Med Phys.* 2000 Mar;27(3):485-98.

21. Kawrakow I, Fippel M. Investigation of variance reduction techniques for Monte Carlo photon dose calculation using XVMC. *Phys. Med. Biol.* 2000;45:2163-83.

22. Sikora M, Dohm O, Alber M. A virtual photon source model of an Elekta linear accelerator with integrated mini MLC for Monte Carlo based IMRT dose calculation. *Phys. Med. Biol.* 2007;52:4449-63.

23. Sikora M, Alber M. A virtual source model of electron contamination of a therapeutic photon beam. *Phys. Med. Biol.* 2009;54:7329-44.

24. Sikora M, Virtual Source Modelling of Photon Beams for Monte Carlo Based Radiation Therapy Treatment Planning, PhD thesis. http://bora.uib.no/handle/1956/5135?show=full

25. Smilowitz JB, Das IJ, Feygelman V et al. AAPM Medical Physics Practice Guideline 5.a: Commissioning and QA of Treatment Planning Dose Calculations – Megavoltage Photon and Electon Beams. *J Appl Clin Med Phys.* 2015 September;16(5):14-34. DOI: 10.1120/jacmp.v16i5.5768

26. Barbeiro AR, Ureba A, Baeza JA et al. 3D VMAT Verification Based on Monte Carlo Log File Simulation with Experimental Feedback from Film Dosimetry. *PLoS ONE*, *11*(11), e0166767. http://doi.org/10.1371/journal.pone.0166767

27. Jursinic PA. Dependence of diode sensitivity on the pulse rate of delivered radiation. *Med Phys.* 2013 Feb;40(2):021720. doi: 10.1118/1.4788763.

28. Huang JY, Followill DS, Wang XA, Kry SF. Accuracy and sources of error of out-of field dose calculations by a commercial treatment planning system for intensity-modulated radiation therapy treatments. *J Appl Clin Med Phys.* 2013 Mar 4;14(2):4139. doi: 10.1120/jacmp.v14i2.4139.

29. Howell RM, Scarboro SB, Kry SF, Yaldo DZ. Accuracy of out-of-field dose calculations by a commercial treatment planning system. *Phys. Med. Biol.* 2010;55(23):6999-7008. doi:10.1088/0031-9155/55/23/S03.

30. Joosten A, Matzinger O, Jeanneret-Sozzi W, Bochud F, Moeckli R. Evaluation of organ-specific peripheral doses after 2-dimensional, 3-dimensional and hybrid intensity modulated radiation therapy for breast cancer based on Monte Carlo and convolution/superposition algorithms: Implications for secondary cancer risk assessment. *Radiother. Oncol.* 2013 Jan; 106(1):33-41. doi: 10.1016/j.radonc.2012.11.012

31. Van den Heuvel F, Defraene G, Crijns W, Bogaerts R. Out-of-field contributions for IMRT and volumetric modulated arc therapy measured using gafchromic films and compared to calculations using a superposition/convolution based treatment planning system. *Radiother Oncol*. 2012;105:127–132. https://doi.org/10.1016/j.radonc.2011.12.030.

32. Kry SF, Bednarz B, Howell RM et al. AAPM TG 158: Measurement and calculation of doses outside the treated volume from external-beam radiation therapy. Med Phys. 2017 Oct;44(10):e391-e429. doi: 10.1002/mp.12462.

33. Scarboro SB, Followill DS, Howell RM, Kry SF. Variations in photon energy spectra of a 6 MV beam and their impact on TLD response. *Med Phys.* 2011 May;38(5):2619-28. DOI: 10.1118/1.3575419

34. Das IJ, Cheng C, Watts RJ et al. Accelerator beam data commissioning equipment and procedures: Report of the TG-106 of the Therapy Physics Committee of the AAPM. Med. Phys., 35: 4186–4215. doi:10.1118/1.2969070

Table 1. 10x10 arcs of varying gantry excursions and MU

Gantry Start/Stop/Excursion	**MU**
-90 / -60 / 30	10, 20, 35, 50, 100
-90 / -30 / 60	20, 40, 70, 100, 200
-90 / 30 / 120	20, 40, 80, 140, 200, 400
-90 / 90 / 180	20, 40, 60, 120, 210, 300, 600

Table 2. Dosimetric comparisons along with their significance. When two comparisons are listed it is because those two comparisons are being compared.

#	Comparison 1	Comparison 2	Significance
1	Plan-CS vs. AC	Plan-MC vs. AC	Which dose calculation algorithm matches AC better?
2	Plan-MC vs. Plan-CS	--	Dose difference due to calculation algorithm is isolated.
3	LF-CS vs. AC	Plan-CS vs. AC	Log files should match AC better. Is this the case?
4	LF-MC vs. AC	Plan-MC vs. AC	
5	LF-MC vs. Plan-MC	--	Machine delivery error is isolated.
6	LF-CS vs. Plan-CS	--	
7	LF-MC vs. Plan-CS	Plan-CS vs. AC	How do LF QA and AC QA compare?
		Plan-MC vs. AC	

Table 3. 6MV Elekta Agility Monte Carlo beam model validation results are shown. Absolute percent differences in output factor (OF), percent depth dose (PDD), penumbral width (PW), and beam width (BW) are shown for each field size. Due to insufficient measured profile length, differences in 40x40cm penumbra were not calculated.

Field (cm)	OF |%Diff|	PDD Mean(|%diff|)	PW Mean(|%diff|)	BW Mean(|%diff|)
1x1	0.3%	--	--	--
2x2	0.4%	0.64 ± 0.23%	1.9 ± 1.0% (0.4mm)	0.9 ± 0.5% (0.2mm)
3x3	0.5%	0.95 ± 0.20%	1.5 ± 1.1% (0.4mm)	1.0 ± 0.5% (0.3mm)
4x4	0.0%	--	--	--
5x5	0.3%	0.22 ± 0.12%	1.0 ± 0.3% (0.50mm)	0.8 ± 0.5% (0.4mm)
6x6	0.4%	--	--	--
8x8	0.3%	--	--	--
10x10	0.0%	0.51 ± 0.11%	0.5 ± 0.3% (0.50 mm)	0.6 ± 0.3% (0.6mm)
12x12	0.3%	--	--	--
15x15	0.1%	0.61 ± 0.11%	--	--
20x20	0.0%	0.75 ± 0.10%	0.7 ± 0.4% (1.5 mm)	0.1 ± 0.1% (0.2mm)
25x25	0.2%	0.75 ± 0.12%	--	--
30x30	0.1%	0.72 ± 0.17%	1.0 ± 0.3% (2.9mm)	0.14 ± 0.11% (0.4mm)
35x35	0.1%	0.73 ± 0.17%	--	--
40x40	0.0%	0.64 ± 0.25%	--	0.25 ± 0.11% (1.0mm)
All	0.20 ± 0.17%	0.65 ± 0.19%	1.1 ± 0.5%	0.54 ± 0.38%

Table 4. MPDD (mean percent dose difference w.r.t AC) values are displayed for various field complexities, static and arc fields, and for 10% and 85% dose thresholds.

Field(s) →		12x12cm	4 6x6cm	9 4x4cm	16 3x3cm	36 2x2cm
Threshold ↓				**Static Fields**		
LF-MC	10%	1.0 ± 11.0	-0.4 ± 11.1	0.2 ± 10.5	0.4 ± 9.4	4.5 ± 17.4
	85%	-1.1 ± 1.4	-2.7 ± 2.2	-1.2 ± 2.4	-0.8 ± 3.3	1.8 ± 4.8
Plan-MC	10%	3.4 ± 10.2	2.3 ± 12.2	0.9 ± 9.3	-0.5 ± 8.5	2.3 ± 16.6
	85%	1.4 ± 1.4	-0.1 ± 1.6	-0.5 ± 1.6	-1.9 ± 3.0	-1.0 ± 3.6
Plan-CS	10%	-0.2 ± 11.6	-2.1 ± 13.6	-3.6 ± 10.7	-6.7 ± 9.2	-11.4 ± 13.5
	85%	-1.5 ± 1.6	-4.3 ± 3.5	-4.6 ± 3.5	-6.9 ± 3.0	-7.6 ± 2.3
				Arcs		
LF-MC	10%	0.3 ± 6.2	-2.4 ± 6.6	-2.7 ± 8.2	-3.8 ± 10.3	-0.5 ± 21.5
	85%	0.2 ± 1.3	-1.8 ± 1.6	-0.9 ± 2.0	-0.1 ± 1.9	-1.6 ± 6.2
Plan-MC	10%	2.7 ± 5.9	-0.2 ± 6.6	-2.1 ± 8.0	-3.7 ± 10.2	-0.7 ± 21.6
	85%	2.8 ± 1.3	0.6 ± 1.6	-0.1 ± 1.9	-0.3 ± 1.9	-3.6 ± 3.2
Plan-CS	10%	-0.2 ± 11.6	-2.1 ± 13.6	-3.6 ± 10.6	-6.6 ± 9.2	-13.0 ± 24.8
	85%	-1.4 ± 1.6	-4.2 ± 3.5	-4.6 ± 1.9	-6.9 ± 3.0	-11.7 ± 2.9

Table 5. 2%/2mm and 1%/1mm gamma pass rates are shown for each dose-pair.

#	Comparison	2%/2mm	1%/1mm
1	Plan-CS vs. AC	97.5 ± 1.6%	81.6 ± 6.5%
2	Plan-MC vs. AC	96.5 ± 4.2%	78.1 ± 9.9%
3	LF-CS vs. AC	97.0 ± 2.4%	79.2 ± 8.1%
4	LF-MC vs. AC	97.3 ± 3.2%	78.9 ± 10.3%
5	Plan-MC vs. Plan-CS	99.0 ± 2.4%	85.0 ± 8.2%
6	LF-MC vs. Plan-MC	99.6 ± 0.4%	91.0 ± 2.7%
7	LF-MC vs. Plan-CS	96.8 ± 5.1%	76.3 ± 11.6%

Table 6. Global percent diode dose differences (mean ± σ) are displayed for various plan pair comparisons and for both 10% (standard) and 85% (target) dose thresholds.

	Comparison	10% Threshold	85% Threshold
#1	Plan-CS vs. AC	0.8 ± 1.8%	0.8 ± 2.6%
#2	Plan-MC vs. AC	1.0 ± 1.8%	2.1 ± 2.5%
#3	LF-CS vs. AC	1.0 ± 1.8%	0.8 ± 2.6%
#4	LF-MC vs. AC	1.0 ± 1.7%	2.1 ± 2.5%
#5	Plan-MC vs. Plan-CS	0.2 ± 1.1%	1.2 ± 1.7%
#6	LF-CS vs. Plan-CS	0.2 ± 1.0%	0.4 ± 1.4%
#7	LF-MC vs. Plan-MC	0.1 ± 0.9%	0.0 ± 1.3%
#8	LF-MC vs. Plan-CS	0.2 ± 1.3%	1.2 ± 2.1%