

Automating linear accelerator quality assurance

Tobias Eckhause

Department of Radiation Oncology, University of Michigan, Ann Arbor, Michigan 48109-5010

Hania Al-Hallaq

Department of Radiation Oncology and Cellular Oncology, The University of Chicago, Chicago, Illinois 60637

Timothy Ritter

Ann Arbor VA Medical Center, Ann Arbor, Michigan 48109

John DeMarco

Department of Radiation Oncology, Cedars-Sinai Medical Center, Los Angeles, California, 90048

Karl Farrey

Department of Radiation Oncology and Cellular Oncology, The University of Chicago, Chicago, Illinois 60637

Todd Pawlicki and Gwe-Ya Kim

UCSD Medical Center, La Jolla, California 92093

Richard Popple

Department of Radiation Oncology, University of Alabama Birmingham, Birmingham, Alabama 35249

Vijeshwar Sharma

Karmanos Cancer Institute, McLaren-Flint, Flint, Michigan 48532

Mario Perez

Royal North Shore Hospital, Sydney, NSW 2065, Australia

SungYong Park

Karmanos Cancer Institute, McLaren-Flint, Flint, Michigan 48532

Jeremy T. Booth

Royal North Shore Hospital, Sydney, NSW 2065, Australia

Ryan Thorwarth and Jean M. Moran^{a)}

Department of Radiation Oncology, University of Michigan, Ann Arbor, Michigan 48109-5010

(Received 17 May 2015; revised 5 September 2015; accepted for publication 9 September 2015; published 25 September 2015)

Purpose: The purpose of this study was 2-fold. One purpose was to develop an automated, streamlined quality assurance (QA) program for use by multiple centers. The second purpose was to evaluate machine performance over time for multiple centers using linear accelerator (Linac) log files and electronic portal images. The authors sought to evaluate variations in Linac performance to establish as a reference for other centers.

Methods: The authors developed analytical software tools for a QA program using both log files and electronic portal imaging device (EPID) measurements. The first tool is a general analysis tool which can read and visually represent data in the log file. This tool, which can be used to automatically analyze patient treatment or QA log files, examines the files for Linac deviations which exceed thresholds. The second set of tools consists of a test suite of QA fields, a standard phantom, and software to collect information from the log files on deviations from the expected values. The test suite was designed to focus on the mechanical tests of the Linac to include jaw, MLC, and collimator positions during static, IMRT, and volumetric modulated arc therapy delivery. A consortium of eight institutions delivered the test suite at monthly or weekly intervals on each Linac using a standard phantom. The behavior of various components was analyzed for eight TrueBeam Linacs.

Results: For the EPID and trajectory log file analysis, all observed deviations which exceeded established thresholds for Linac behavior resulted in a beam hold off. In the absence of an interlock-triggering event, the maximum observed log file deviations between the expected and actual component positions (such as MLC leaves) varied from less than 1% to 26% of published tolerance thresholds. The maximum and standard deviations of the variations due to gantry sag, collimator angle, jaw position, and MLC positions are presented. Gantry sag among Linacs was 0.336 ± 0.072 mm. The standard deviation in MLC position, as determined by EPID measurements, across the consortium was 0.33 mm for IMRT fields. With respect to the log files, the deviations between expected and actual positions for parameters were small (<0.12 mm) for all Linacs. Considering both log files

and EPID measurements, all parameters were well within published tolerance values. Variations in collimator angle, MLC position, and gantry sag were also evaluated for all Linacs.

Conclusions: The performance of the TrueBeam Linac model was shown to be consistent based on automated analysis of trajectory log files and EPID images acquired during delivery of a standardized test suite. The results can be compared directly to tolerance thresholds. In addition, sharing of results from standard tests across institutions can facilitate the identification of QA process and Linac changes. These reference values are presented along with the standard deviation for common tests so that the test suite can be used by other centers to evaluate their Linac performance against those in this consortium. © 2015 American Association of Physicists in Medicine. [<http://dx.doi.org/10.1118/1.4931415>]

Key words: quality assurance, linear accelerator, EPID, multileaf collimator, quality control

1. INTRODUCTION

Treatment delivery systems are capable of complex simultaneous motions which have led to increased quality assurance (QA) requirements. Multiple simultaneous motions along different axes increase flexibility in treatment delivery and can decrease the total time a patient is on a treatment table. These increasingly complex systems require new QA tests and procedures. AAPM Task Group 142 (TG 142)¹ provides guidance tolerance values for individual components of linear accelerators based on the treatment delivery type [conformal, IMRT/volumetric modulated arc therapy (VMAT), SRS]. Users develop their own tests with their own equipment or purchase QA equipment and software to meet that guidance. The clinical physics community faces an ongoing challenge to keep abreast of the latest safety and quality practices as new technologies are developed and clinically implemented. With respect to the types of QA methods, methods based on portal images²⁻⁶ and dynamic log files⁷⁻⁹ have been used to assess the performance of an accelerator. During treatment, the linear accelerator control system records mechanical component positions and monitor unit (MU) delivery. Linear accelerator (Linac) log files can provide supplemental information for evaluating standard delivery methods as well as more recent delivery techniques such as MLC tracking¹⁰ or 4π .¹¹ Electronic portal imaging devices (EPIDs) integrated into the accelerator are used for patient monitoring, including transit dosimetry,¹² and the digital output format lends itself to automated image processing. Such methods can be valuable for monitoring the performance of the overall system.

Our hypotheses are that we can increase the efficiency of QA with an automated and integrated QA program, and that we can determine expected performance values for a specific type of linear accelerator. This hypothesis was tested by comparing the results of a QA test suite across a consortium of eight institutions. Each center delivered the same test suite at regular intervals on their Linac using a standard phantom created for this collaboration. Collected data were compared within and across institutions. The goal was to assess Linac performance over time and assess the type of information learned from the log file and the EPID. The consortium allows us to test the portability of these tools and to validate a simplified, consistent QA method for current and new treatment delivery systems.

2. MATERIALS AND METHODS

We formed a multi-institutional collaboration which included eight Varian TrueBeam linear accelerators. The TrueBeam control system has more automation than previous generations of Varian Linacs.¹³⁻¹⁵ The developed test suite emphasized the mechanical requirements of AAPM TG 142.¹ The same suite of QA test fields was delivered on each Linac for four consecutive weeks and at monthly intervals for six months. The sensitivity of each test was assessed by introducing deliberate deviations in the delivered fields and by aligning a simple, standard phantom in the field with EPID images as a reference. Software was developed with MATLAB (The Mathworks, Inc., Natick, MA) to automate the analysis of the log files and EPID images. The content of the log files was validated using a combination of mechanical checks and analysis of EPID images. Each institution delivered the test suite on their TrueBeam linear accelerator with one institution measuring on two Linacs. The use of the test suite was parallel to each institution's standard Linac QA program.

2.A. Linear accelerators

The analysis techniques in this work require the trajectory log files, which list Linac component positions as a function of time, and an integrated EPID. The specific tools developed for this study were designed for use on the Varian TrueBeam radiotherapy delivery system (Varian Medical Systems, Palo Alto, CA). The test fields are in DICOM-RT format and could be utilized on other Linacs with similar delivery capabilities with appropriate conversions for any hardware differences (such as MLC type and leaf width). The complete test suite is also deliverable without modification on a Varian iX accelerator equipped with factory-installed VMAT capabilities. The analysis tools and software for this project, however, were customized to the Varian TrueBeam Linac. For photon delivery, the tests included static fields, static gantry intensity modulated radiation therapy (SG-IMRT) fields where only the MLC is in motion, and VMAT fields where the MLC and gantry are in simultaneous motion during delivery. The gantry angle motion can accelerate and decelerate during delivery. In this study, we used 6 MV photon beams. Among the eight accelerators, five have the HD120 MLC and three have the Millennium MLC. Both models have 120 leaves, 60

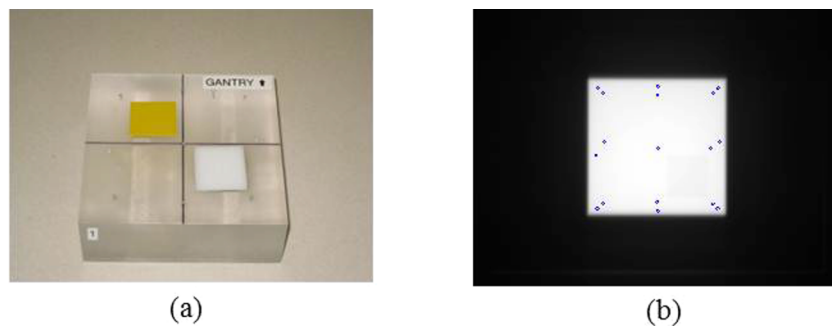


FIG. 1. (a) The phantom contains small spherical fiducials. The two pieces of plastic (upper left and lower right) create contrast for measuring image quality. (b) An EPID image of the phantom. The locations of the fiducials (marked with circles) are determined with automated analysis software.

in each bank. Each bank of the HD120 model contains 32 narrow central leaves projecting to 2.5 mm at isocenter and 28 outer leaves projecting to a width of 5.0 mm (except for the outermost leaves). The Millennium MLC contains 40 central leaves in each bank projecting to 5.0 mm and 20 outer leaves projecting to 10.0 mm width. DICOM-RT plan files were developed for a standard machine and each MLC type (Millennium or HD120). The files were provided to each institution with substitution of local machine ID. Files could be used for delivery in clinical mode or file mode.

The TrueBeam system determines the positions of mechanical components using the primary resolvers²⁵ and encoders which are part of the feedback system that controls the motors that drive motion. The recorded gantry angle, in particular, is determined by a digital encoder. The control system records these values at 20 ms intervals and saves them, along with the expected (planned) values, to a trajectory log file at the end of each delivery. The log file records couch, collimator, gantry axis, MLC leaf positions, and MUs. Analysis tools were developed to validate trajectory log file data directly against EPID measurements and to evaluate the performance of multiple accelerators over time. The recorded values can be compared with baseline or expected values. The trajectory log files are typically several megabytes, resulting in several gigabytes per week for a typical patient volume. Consequently, the TrueBeam system periodically purges trajectory files based on file age and total number of files. Participants were instructed to ensure that the trajectory log files created during delivery of the QA tests were recovered for analysis before deletion.

For the log file data to be used as a supplemental part of a QA program, independent checks of the validity of the recorded axes' positions and MU are necessary. In this work, EPIDs were used to verify the log file data.

The TrueBeam gantry includes an integrated EPID perpendicular to the treatment beam. The model AS1000 EPID has a resolution of 1024×768 with a pixel size of 0.392×0.392 mm. All images in this study were acquired at a source-detector-distance of 150 cm. Each image was flood- and dark-field corrected and the values of bad pixels were replaced with averages over neighboring pixels.

2.B. Phantom

A simple phantom was created to establish the geometry of the EPID images, to standardize the tests, and to enable

automated analysis. The phantom was constructed of Lucite into which 2-mm diameter steel balls were embedded (Fig. 1). The steel balls serve as fiducials to allow the determination of the field position against a fixed coordinate system. The fiducials are arranged in a 3D pattern with a central ball located at isocenter. Scribe marks were made on the central axis of the phantom (top and sides) to aid in the setup and alignment of the phantom in the treatment room. Two different types and thicknesses of plastic (yellow and white squares in Fig. 1) were attached to the Lucite to determine contrast to noise ratio (outside of the scope of this work). The manufacturing tolerance was a 0.025 mm. CT imaging with 1 mm slice thickness was used to verify the phantom density and to verify the depth of the fiducials.

2.C. Analysis algorithms and software

We developed software in MATLAB to automate the analysis of both the log files and images. For the log file analysis, we define deviation as the actual value minus the expected value. The magnitudes of those deviations were then compared with a subset of tolerance limits from Tables II to IV of the AAPM TG-142 report.¹ When specified, the more stringent tolerances for SRS/SBRT were used.

In addition to performing QA checks, the developed software can display time series or integrated information on component positions. Leaf positions and leaf position deviations can be displayed. Similarly, collimator angle, gantry angle, and couch positions and deviations can be plotted. Figure 2 shows some examples of such plots for delivery of a sliding window field.

Each participant submitted the log files and EPID images associated with delivery of the QA test suite described in Sec. 2.D. For each Linac component, the deviations were calculated for each snapshot. The maximum and standard deviation of the difference between the expected and actual value were then determined for each parameter. The precision of the reported data in the log file is orders of magnitude below the tolerance limits defined in the TG-142 report. The validation is limited to the tools used for the independent assessment. For example, gantry angles have a precision of 0.007° and MLC leaf positions, $3 \mu\text{m}$. These positions and angles must be validated in order to be part of a QA program. By comparing fiducial positions and field edge positions, leaf and collimator

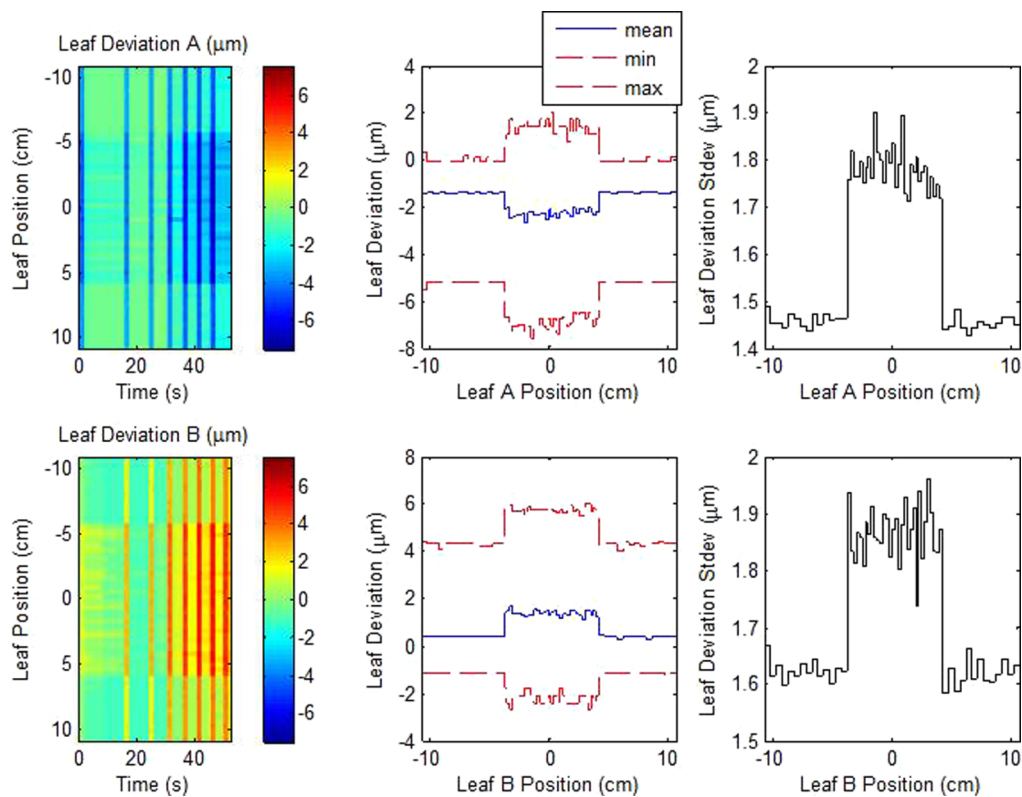


FIG. 2. A sample display of the leaf position deviation during delivery of a sliding window test. Positions deviations versus time are shown in the color-scaled display on the left. The maximum, minimum, and average deviations, in microns, are shown in the center and the standard deviation of the position deviation is shown on the right. The statistics are computed for the 2627 snapshots, acquired at 20 ms intervals during delivery of the field. The blue vertical stripes in leaf bank A and the red vertical stripes in leaf bank B indicate larger deviations which occur during leaf motion. The maximum reported leaf deviations are approximately $6 \mu\text{m}$.

positions and angles can be calculated with respect to a fixed location in space.

In this study, EPID images of the treatment field with the phantom in place are used to determine leaf and collimator positions with respect to a fixed phantom. Any offset in the setup of the phantom could be accounted for by the image analysis. The images of the phantom with its fiducials serve as part of the validation of the information in the trajectory log file. To facilitate automation, we have also created software to identify image files associated with test fields and automatically analyze those images. The images are examined to locate (i) phantom fiducial markers, (ii) collimator defined field edges, and (iii) MLC leaf defined field edges. The image analysis algorithms are customized for both types of MLCs used in this study.

The phantom fiducials are located using the MATLAB implementation of the Canny edge detection algorithm. The threshold for edge detection is adjusted until all of the 17 steel balls in the phantom are detected in the image. For jaw-defined fields, the field edges are the interpolated position where the intensity is the average of in-field and out-of-field intensities. The field edges are then fit to a line, and jaw positions and collimator rotations are determined from the field edge lines. Several of the tests evaluate the accuracy of leaf positions. For each leaf pair, an intensity profile in the direction of leaf motion and the average is calculated across the central third of the width of each leaf. For fields with larger leaf gaps, the leaf edge

position is calculated and reported as the peak in the intensity gradient parallel to the leaf. For fields where leaf pairs have a narrow gap, as in a picket fence test, the leaf positions are calculated instead by locating the local peak in the intensity profile.^{2,16} The leaf edge positions from each member of the leaf pair are then identified as the peak position \pm half of the FWHM. The EPID area confined the images to the central 36 leaf pairs of the Millennium MLC and the central 52 leaf pairs of the HD MLC, and only those leaves are included in the imaging analysis. Figure 3 shows the leaf positions detected in a picket fence test image.

We also use the image analysis to automatically calculate shifts in the radiation field at the cardinal gantry angles. Sag is due to gravitational torque on the gantry. We quantify the sag as the change in the field center with respect to the phantom center (at isocenter) between a gantry angle of 0° and 180° (using the IEC standard). In moving from 0° to 180° , the field shifts along the gantry rotation axis away from the gantry.¹⁷ The location of the phantom is determined from the location of the central ball bearing on the EPID image. The field position is determined from the location of the field edges.

We also calculate a shift in the radiation field in the vertical direction. The vertical shift in field position is defined as the difference between radiation field positions (with respect to the phantom) at gantry angles of 90° and 270° . The phantom and field positions are determined in the same manner as is done to determine gantry sag.

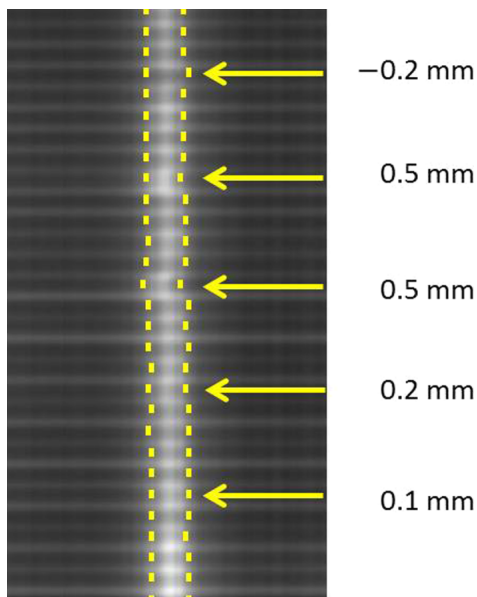


FIG. 3. A magnified image of a single picket in a picket fence test. The leaf positions recorded in the log files are validated using EPID images. The dashes indicate the individual leaf positions, as calculated from the image. Deliberate jogs in leaf positions are indicated by arrows.

2.D. Test fields and phantom

The test suite was developed to evaluate collimator and MLC deviations for conformal, IMRT, and VMAT delivery techniques. All of the MLC IMRT tests (Table V of TG 142) and collimator rotation and jaw position tests identified in the TG-142 report are included. Dosimetric, mechanical, safety, and respiratory gating tests are not part of the current test suite. The test fields were designed in ARIA 11 (Varian Medical Systems, Palo Alto, CA) and exported to each center in DICOM-RT format. All fields use a 6 MV photon beam. There were 21 delivery fields and 4 setup fields.

The EPID on each Linac was calibrated in the manner specified by the manufacturer, by acquiring a flood and a dark field image. To account for differences in the calibration settings and off-axis profiles, the first field delivered in the test suite is a flood field. The EPID response to this field is used to normalize EPID images for consistent analysis of data from each Linac, regardless of institutional factors. EPID images are normalized using the software developed in MATLAB so that this flood field creates an image which is radially symmetric about the imager center. Each delivery of the test suite begins with the calibration flood field, which is delivered with no obstruction (table retracted) between the gantry head and the imager.

After the flood field was delivered and the collimator and gantry set to 0° , a central fiducial is aligned to isocenter in the manner used in regular practice at each institution, either using laser alignment or using crosshairs. Rotational alignment of the phantom is accomplished by aligning scribe marks. The height of the phantom is positioned such that a central fiducial was near gantry isocenter based on either the laser or source-surface distance indicators. The EPID is extended with a source-detector distance of 150 cm.

Table I lists the fields and the corresponding procedures that are evaluated for each field or set of fields. The last three test fields, which test variable gantry speed and variable dose rate, were downloaded from the Varian website (available at myvarian.com).

After the flood field, a set of $10 \times 10 \text{ cm}^2$ fields is delivered at cardinal gantry angles. Picket fence tests at the cardinal gantry angles are used to test leaf position and leaf velocity accuracy. MLC transmission images are acquired at gantry angles 0° and 90° . Comparisons are made to baseline images. The fields with staggered leaf patterns test static leaf position accuracy at gantry angles of 0° and 90° . The remaining tests combine leaf motion with gantry rotation at both fixed and variable speeds, followed by a sliding window field with a variable dose rate.

3. RESULTS



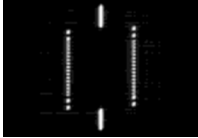
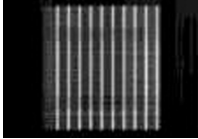
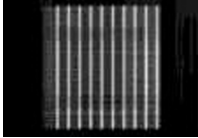
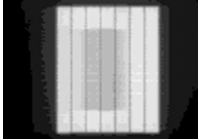
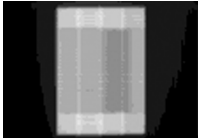
The same set of test fields was delivered on each Linac. Test suite delivery and analysis were efficient for all consortium participants, with data acquisition and analysis of EPID images and trajectory log files taking approximately 30 min to complete. In that time, all of the MLC IMRT tests (Table V) identified in the TG 142 report¹ are accomplished along with additional VMAT tests. The expected values of the Linac components and delivered MU, as recorded in the log files, were identical for each delivery of the field for every participating center. For measurements of leaf positions, and collimator jaw positions, the accuracy with which the parameters were validated was below the TG-142 action threshold. A summary of the results is shown in Table II.

EPID images are used to validate the leaf positions recorded in the log files. To do so, deliberate deviations in leaf positions were introduced into a picket fence test. Figure 3 shows an EPID image of the MLC and the calculated leaf positions, with intentional deviations in leaf position indicated by the arrows. We are able to distinguish leaf position deviations exceeding 0.3 mm. Although leaf positions are recorded with a precision of $1 \mu\text{m}$, the automated image analysis method cannot validate to this level of precision. It is important to note that the trajectory log files provide data based on the encoders. For routine QA, the data must be validated with a measurement such as with the EPID using the phantom as a reference.

The TG-142 tolerance threshold for segmental and dynamic IMRT error is deemed acceptable if 95% of the leaf deviations are less than 3.5 mm and the maximum error root-mean-square (RMS) is less than 3.5 mm, based on a histogram of time-dependent leaf position errors. Since our analysis is based on validated leaf positions, we use the more stringent tolerance limit of 1 mm in a picket fence test.

Deviations (the actual minus expected value) in the mechanical positions as recorded in log files are less than the action threshold, often by orders of magnitude. The behavior of the TrueBeam system can be contrasted with that of other systems, where deviations are significantly larger.¹⁴ Trajectory log files show that leaf position deviations remain well below tolerance limit of 1 mm. No leaf deviation exceeding

TABLE I. Fields delivered as part of the test suite and the corresponding procedures tested.

Field description	Test Image	Procedure(s) tested	TG 142 table [Klein <i>et al.</i> (Ref. 1)]	EPID analysis	Log file analysis
1. Jaw-defined field ^a		Gantry sag, collimator rotation, and jaw position	Table II (monthly)	Field edge, angle; phantom position	Jaw and collimator
2. MLC defined static pattern ^a		Leaf positions	Table V (monthly)	Leaf edge positions	Leaf position
3. Interleaf static MLC pattern ^a		MLC transmission	Table V (annual)	Image-to-image comparison	Leaf position
4. Picket fence test—static gantry ^a		Leaf position (IMRT)	Table V static gantry quantitative (monthly)	Leaf positions for all pickets and cardinal gantry angles	Leaf position and velocity
5. Picket fence test while gantry is rotating		Leaf position (IMRT) in ARC mode	Extension to Table V – arc delivery (monthly)	Leaf positions for all pickets	Leaf position and gantry angle
6. Sliding window with variable gantry speed		Leaf position (IMRT) and variable travel speed	Extension to Table V – arc delivery, + checking limits	Image-to-image comparison	Leaf position, leaf velocity, and gantry angle
7. Sliding window with variable gantry speed and variable dose rate		Leaf position accuracy (IMRT), travel speed, and MU/angle	Extension to Table 5 – arc delivery, + checking limits	Image-to-image comparison	Leaf position, gantry angle, and MU/angle

^aDelivered at multiple gantry angles.

0.11 mm was recorded except during events where the beam was held off. During events where the beam was held off, the leaf positions satisfy the criteria that the magnitude of leaf position deviation is less than 3.5 mm RMS for 95% of error counts, as listed in the report of TG 142. The largest leaf deviations during delivery of IMRT fields were observed in the picket fence test at a gantry angle of 270°. As seen in Fig. 4, the maximum deviations vary little over time but may differ between Linacs. The three Linacs with smaller deviations (1–3) had the Millennium MLC and the others had the HD120 MLC.

The recorded collimator angle showed deviations in the log files that were a maximum 0.0057°. The collimator angle was also validated by locating phantom fiducials in EPID images. The angle of the field edges was compared with the angle of

the phantom. The standard deviation of the difference in field angle from the phantom angle was 0.14° and the maximum difference was 0.57° across all institutions and all deliveries. The accuracy with which the collimator angle has been validated is limited due to setup variations.

The largest deviation, as a ratio of tolerance limits, was seen for the gantry angle. The largest deviations in gantry angle occurred during acceleration and deceleration of the gantry during delivery of VMAT fields. Although the gantry angle was not mechanically validated as a standard part of the test suite, measurements of gantry angle form a standard part of independent QA programs at each institution. All gantry angle deviations recorded in the log files are below tolerance thresholds. The maximum gantry angle deviation occurred during acceleration and deceleration of the gantry.

TABLE II. Summary of Linac behavior. The first column describes the Linac component parameter tested. The accuracy with which the data have been validated is listed in the second column. The log file deviation is the average across institutions of the maximum deviation of that parameter.

Linac performance variable	Validation level	Log file deviation	TG-142 action threshold	Note
Static MLC position	0.3 mm	0.009 mm	1 mm	Validation from EPID images
Dynamic MLC position	0.3 mm	0.098 mm	1 mm	Validation from EPID images
Gantry angle	0.5°	0.26°	1°	Not mechanically validated for all deliveries
Gantry sag	0.2 mm	Not applicable	1 mm	Validation from EPID of average sag of 0.35 mm from EPID
Interleaf leakage	Not applicable	Not applicable	99% agreement 2%/1 mm	Used gradient compensation (Ref. 18)
Collimator angle	0.6°	0.087°	1°	Validation limited by phantom alignment

Automated measurements of gantry sag, where the field displacement is assessed along the gantry rotation axis between 0° and 180°, are possible with the use of the phantom and the EPID. Gantry sag was determined by locating the fixed phantom from the EPID data as determined by automated image analysis methods similar to the method by Rowshanfarzad *et al.*¹⁹ The field positions along the gantry axis at 90° and 270° are intermediate between those at 0° and 180°. The sag varied between 0.21 and 0.49 mm with consistent results over time, but with larger variation between Linacs.

The beam central axis lies less than 0.8 mm from the imager center in all cases for all deliveries at all gantry angles, and the standard deviation of that distance varies between 0.21 and 0.39 mm depending on gantry angle, which is similar to that determined by Gao *et al.*²⁰

Phantom setup variations, while not a component of the Linac, were recorded as well. The lateral and longitudinal displacements of the phantom from field center were -0.07 ± 0.45 mm and -0.06 ± 0.46 mm, respectively. Those variations may arise from human setup uncertainties, or deviations of laser or optical crosshairs from the beam axis. This demonstrates that the phantom was easy to use in a reproducible

way and that each institution kept their laser and crosshair alignment well within TG-142 tolerances.

4. DISCUSSION

Tolerance thresholds to initiate an investigation and possible corrective actions are suggested in the TG-142 report, but other action criteria are left to the institution.¹

A change in Linac behavior, such as poorer performance in positioning of a leaf, may lead the physicist to adjust the equipment even if the deviation is below the clinical threshold to stop treatment. This is analogous to a level 1 action as described in TG-142 report. The criteria to initiate this action depend on the typical behavior of a Linac and whether or not the Linac is used for stereotactic treatments. The establishment of criteria for investigative action can be facilitated by the collection of a larger set of data on Linac behavior which can be achieved by sharing across institutions. The use of standard QA tools and algorithms can facilitate not only expedited QA procedures but also comparison of overall Linac performance.

In some instances, behavior that was acceptable but differed from other accelerators was identified such as for variations in the field shift along the couch. In AAPM TG 119, the authors performed the same tests on IMRT plans with different planning and dosimetry systems.²¹ The multi-institutional results were used to establish confidence intervals.²¹ This same approach could be used to redesign a QA program with an emphasis on the most sensitive tests. Glide-Hurst *et al.* demonstrated support of quality through intercomparison of commissioning results for leaf reproducibility, dynamic leaf gap, output, and other characteristics of TrueBeam accelerators at three institutions.¹³ In this work, all collaborators used the same phantom and delivery suite for their QA. The current technique was efficient with further time savings possible by minimizing redundancy in the tests or removal of tests that are not required monthly. Other fields could perhaps be modified based on the performance of the Linacs. In this way, guidance could be developed based on performance. For example, the picket fence test results showed that the variations in the leaf position accuracy were largest at a gantry angle of 270°. With additional data and adequate evaluation, the suite could potentially be reduced to measure at the more sensitive gantry angle weekly and then at all cardinal angles

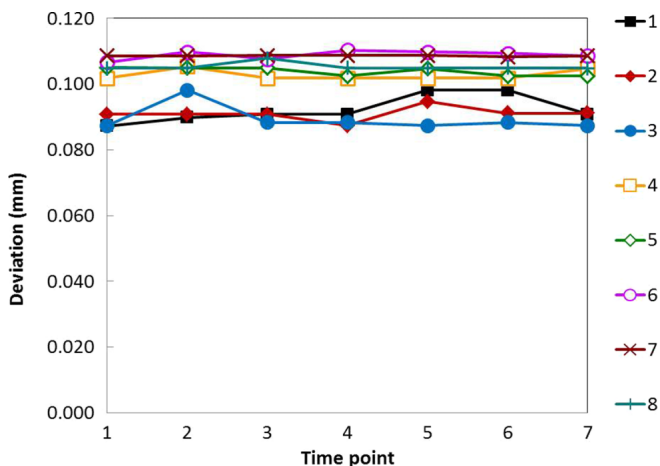


FIG. 4. The maximum leaf deviation for all linear accelerators. Linacs 1–3, denoted with filled markers, were equipped with Millennium MLCs. Linacs 4–8 were equipped with HD120 MLCs. The first four time points represent delivery at weekly intervals followed by monthly intervals, except at one institution, where the initial intervals are monthly, followed by weekly measurements. Note the stable response over time.

monthly. Alternatively, the physicist could choose to retain testing at the cardinal gantry leveraging the efficiency gains achieved with the test suite and rapid analysis tools (~1 min).

Agnew *et al.* validated log files with EPID measurements to assess how their use might fit into a modified quality control program for IMRT.²² Trajectory log files and EPID measurements were acquired daily for a year to assess the overall MLC performance. With the trajectory log files, they found the MLC positioning accuracy on two TrueBeam accelerators to be 0.004 ± 0.011 mm and 0.004 ± 0.003 mm for static and rotational deliveries, respectively. With their EPID measurements, they found positioning accuracy of 0.000 ± 0.039 mm and 0.000 ± 0.025 mm for static and rotational measurements, respectively. Our results show similar accuracies for trajectory log files, with the exception that we observe larger standard deviations in leaf position during rotational delivery. They also demonstrated the fact that trajectory log files require validation and may provide results different from imaging measurements. Unlike Agnew, we did not find the same disagreement of accuracy between EPID and trajectory log files, possibly since our measurements were made monthly rather than daily. However, it is important to highlight that the trajectory log file alone is not a measurement but rather a supplemental record of the Linac behavior. For any use of log files in a quality assurance or quality control program, we strongly agree that log files must be validated regularly for other uses such as evaluating patient delivery.

The standard deviations in Linac parameters presented in Table III include deviations due to differences between the eight linear accelerators and deviations due to changes over time, except interleaf transmission, where the baseline for measurements is specific to each linear accelerator. The standard deviations of leaf and collimator positions, as determined by EPID measurements, are dominated by deviation

between accelerators. When a baseline is established separately for each accelerator, the average deviations are 0.00 ± 0.026 mm for static MLC, 0.00 ± 0.0221 mm for IMRT MLC, and 0.02 ± 0.0417 mm for VMAT MLC positions. These latter values are similar to other published deviations for individual linear accelerators and show deviations that are approximately a factor of 10 smaller than deviations across institutions.²²

As new delivery techniques are introduced clinically, it may be beneficial to have additional information on Linac performance beyond what can be garnered from transit dosimetry measurements combined with cone beam CT information.²³ Trajectory log files can also be used to reconstruct delivered doses for patient treatments. Sun *et al.* used trajectory log files to verify the treatment delivery for 127 IMRT patients and 51 standard treatments.²⁴ They developed software to automate the analysis and found the use of the log files to be beneficial in confirming the treatment delivery as a more efficient process than a physics weekly check with inspection of parameter positions by a physicist.

Across institutions, we found that the observed deviations have been stable over time and consistent between institutions with most deviations far below tolerances. The performance of all MLCs was well below the published guidance of TG 142. Differences between the two MLCs from the EPID measurements were small and may be due to small differences in how the leaf identification algorithm works for the different size leaves. The focus in our work was on evaluating the overall performance of the TrueBeam accelerator with respect to published guidance and to provide data on expected Linac performance as a reference for the broader medical physics community. For example, there has been significant value to the radiotherapy community to have a standard set of phantoms and expected QA results for the IMRT commissioning tests designed and made available as part of AAPM TG 119.²¹

TABLE III. Summary of average values and deviations across eight linear accelerators for selected Linac performance parameters.

Linac parameter	Method	Field	Average from EPID	Average from logs
Gantry sag	Images of phantom at gantry angles of 0° and 180°	1	0.336 ± 0.072 mm	Not applicable
Field shift on couch axis	Images of at gantry angles of 90° and 270°	1	-0.109 ± 0.539 mm	Not applicable
Collimator angle deviation	Field edge angle against aligned phantom	1	$-0.035^\circ \pm 0.179^\circ$	$0.0009^\circ \pm 0.0016^\circ$
Collimator position deviation	Field edge position against phantom center	1	-0.06 ± 0.30 mm	< 1 μm mean and standard deviation
Static MLC position, mean leaf deviation	Field edge	2	Gantry = 0°: 0.00 ± 0.224 mm; gantry = 90°: 0.02 ± 0.228 mm	Gantry = 0°: 0.005 ± 0.007 mm; gantry = 90°: -0.004 ± 0.012 mm
Interleaf MLC transmission	Image-to-baseline using gradient compensation with 1% and 1 mm	3	$99.9\% \pm 0.3\%$ pixels passed	Not applicable
IMRT MLC position, gantry 270°		4	0.00 ± 0.333 mm	-0.003 ± 0.042 mm
VMAT MLC position, deviation	EPID image field edge and log file	5	0.02 ± 0.282 mm	-0.004 ± 0.012 mm
VMAT maximum gantry angle deviation	Log file	6	Not applicable	$0.08^\circ \pm 0.027^\circ$

Delivery and analysis of the EPID images and trajectory log files took approximately 30 min to complete. The estimated time to perform the same set of tests using traditional (nonautomated) methods is between 1.5 and 2 h, excluding analysis time. Meaningful investigation of the results and comparison to baselines would add significantly more time depending on the software used and the type of analysis performed. Additionally, the use of the EPID to quickly acquire a dense, 2D set of data has distinct advantages over competing methods. For example, leaf positions are quantitatively analyzed from the picket fence test rather than only scoring the test results as pass or fail based on visual inspection by the physicist. The test suite allows users to quickly investigate how their machine performs against the baselines established by other users of the same equipment. Action levels and control limits can potentially be determined with an adequate number of measurements over time. Individual tests can also be performed after machine maintenance to confirm proper functioning of equipment prior to releasing equipment for clinical use.

5. CONCLUSIONS

In this work, we present a test suite and automated analysis tools, and use them to compare the performance of multiple accelerators at different institutions. In all instances, the measured results were well below the thresholds reported in the AAPM TG 142 report.¹ We have demonstrated consistent results in performance based on automated analysis of trajectory log files and EPID images. The analysis, together with the test suite, can be used in a QA program. Mechanical deviations involving the MLC and collimator in particular are directly recorded and deviations between actual and expected values can be directly compared to tolerance or action levels. We have presented statistical information about the actual performance of a group of TrueBeam Linacs and MLCs.

This collaborative partnership points toward broadening our approach to quality control as a community where instead of individual medical physicists reinventing the QA wheel each time a new technique becomes available, we use common tools to permit a meaningful evaluation of overall beam performance. TrueBeam and other Linac users will be able to compare their QA results with the performance achieved within this consortium. This approach can be used for other equipment as well. All but one of the fields in this test suite can be used in their current format for EPID QA measurements for an earlier generation Linac model by the same manufacturer. The test suite can also be modified to work with other manufacturers' Linacs since the tests are all based on those recommended by the TG 142 report. Linacs which are outliers compared to the published results can be quickly identified for corrective action, thereby improving the delivery accuracy. In addition, we have shown improvements through the development of an efficient QA process in a consortium, with data acquisition and analysis of EPID images and trajectory log files taking approximately 30 min to complete. Future work will include the extension of the test

suite to encompass more of the monthly tests in the TG 142 report.

ACKNOWLEDGMENTS

This work was supported by a grant from Varian Medical Systems. The authors thank Wayne Keranen and Thanos Etmektzoglou of Varian Medical Systems for providing technical information in support of the project.

^{a)}Author to whom correspondence should be addressed. Electronic mail: jmmoran@med.umich.edu

¹E. E. Klein, J. Hanley, J. Bayouth, F. F. Yin, W. Simon, S. Dresser, C. Serago, F. Aguirre, L. Ma, B. Arjomandy, C. Liu, C. Sandin, and T. Holmes, "Task Group 142 Report: Quality assurance of medical accelerators," *Med. Phys.* **36**, 4197–4212 (2009).

²J. Chang, C. H. Obcemea, J. Sillanpaa, J. Mechalakos, and C. Burman, "Use of EPID for leaf position accuracy QA of dynamic multi-leaf collimator (DMLC) treatment," *Med. Phys.* **31**, 2091–2096 (2004).

³J. I. Prisciandaro, C. M. Frechette, M. G. Herman, P. D. Brown, Y. I. Garces, and R. L. Foote, "A methodology to determine margins by EPID measurements of patient setup variation and motion as applied to immobilization devices," *Med. Phys.* **31**, 2978–2988 (2004).

⁴M. Mamalui-Hunter, H. Li, and D. A. Low, "MLC quality assurance using EPID: A fitting technique with subpixel precision," *Med. Phys.* **35**, 2347–2355 (2008).

⁵P. M. McCowan, D. W. Rickey, P. Rowshanfarzad, P. B. Greer, W. Ansbacher, and B. M. McCurdy, "An investigation of gantry angle data accuracy for cine-mode EPID images acquired during arc IMRT," *J. Appl. Clin. Med. Phys.* **15**, 187–201 (2014).

⁶P. Rowshanfarzad, M. Sabet, M. P. Barnes, D. J. O'Connor, and P. B. Greer, "EPID-based verification of the MLC performance for dynamic IMRT and VMAT," *Med. Phys.* **39**, 6192–6207 (2012).

⁷J. G. Li, J. F. Dempsey, L. Ding, C. Liu, and J. R. Palta, "Validation of dynamic MLC-controller log files using a two-dimensional diode array," *Med. Phys.* **30**, 799–805 (2003).

⁸D. W. Litzenberg, J. M. Moran, and B. A. Fraass, "Verification of dynamic and segmental IMRT delivery by dynamic log file analysis," *J. Appl. Clin. Med. Phys.* **3**, 63–72 (2002).

⁹D. W. Litzenberg, J. M. Moran, and B. A. Fraass, "Incorporation of realistic delivery limitations into dynamic MLC treatment delivery," *Med. Phys.* **29**, 810–820 (2002).

¹⁰P. R. Poulsen, B. Cho, A. Sawant, D. Ruan, and P. J. Keall, "Dynamic MLC tracking of moving targets with a single kV imager for 3D conformal and IMRT treatments," *Acta Oncol.* **49**, 1092–1100 (2010).

¹¹P. Dong, P. Lee, D. Ruan, T. Long, E. Romeijn, Y. Yang, D. Low, P. Kupelian, and K. Sheng, "4pi non-coplanar liver SBRT: A novel delivery technique," *Int. J. Radiat. Oncol., Biol., Phys.* **85**, 1360–1366 (2013).

¹²W. van Elmpt, L. McDermott, S. Nijsten, M. Wendling, P. Lambin, and B. Mijnheer, "A literature review of electronic portal imaging for radiotherapy dosimetry," *Radiother. Oncol.* **88**, 289–309 (2008).

¹³C. Glide-Hurst, M. Bellon, R. Foster, C. Altunbas, M. Speiser, M. Altman, D. Westerly, N. Wen, B. Zhao, M. Miften, I. J. Chetty, and T. Solberg, "Commissioning of the Varian TrueBeam linear accelerator: A multi-institutional study," *Med. Phys.* **40**, 031719 (15pp.) (2013).

¹⁴J. Li, R. D. Wiersma, C. J. Stepaniak, K. J. Farrey, and H. A. Al-Hallaq, "Improvements in dose accuracy delivered with static-MLC IMRT on an integrated linear accelerator control system," *Med. Phys.* **39**, 2456–2462 (2012).

¹⁵R. A. Popple and I. A. Brezovich, "Dynamic MLC leaf sequencing for integrated linear accelerator control systems," *Med. Phys.* **38**, 6039–6045 (2011).

¹⁶T. Losasso, "IMRT delivery performance with a varian multileaf collimator," *Int. J. Radiat. Oncol., Biol., Phys.* **71**, S85–S88 (2008).

¹⁷W. Du, S. Gao, X. Wang, and R. J. Kudchadker, "Quantifying the gantry sag on linear accelerators and introducing an MLC-based compensation strategy," *Med. Phys.* **39**, 2156–2162 (2012).

¹⁸J. M. Moran, J. Radawski, and B. A. Fraass, "A dose gradient analysis tool for IMRT QA," *J. Appl. Clin. Med. Phys.* **6**, 62–73 (2005).

- ¹⁹P. Rowshanfarzad, M. Sabet, D. J. O'Connor, P. M. McCowan, B. M. McCurdy, and P. B. Greer, "Detection and correction for EPID and gantry sag during arc delivery using cine EPID imaging," *Med. Phys.* **39**, 623–635 (2012).
- ²⁰S. Gao, W. Du, P. Balter, P. Munro, and A. Jeung, "Evaluation of IsoCal geometric calibration system for Varian linacs equipped with on-board imager and electronic portal imaging device imaging systems," *J. Appl. Clin. Med. Phys.* **15**, 164–181 (2014).
- ²¹G. A. Ezzell, J. W. Burmeister, N. Dogan, T. J. LoSasso, J. G. Mechalakos, D. Mihailidis, A. Molineu, J. R. Palta, C. R. Ramsey, B. J. Salter, J. Shi, P. Xia, N. J. Yue, and Y. Xiao, "IMRT commissioning: Multiple institution planning and dosimetry comparisons, a Report from AAPM Task Group 119," *Med. Phys.* **36**, 5359–5373 (2009).
- ²²A. Agnew, C. E. Agnew, M. W. Grattan, A. R. Hounsell, and C. K. McGarry, "Monitoring daily MLC positional errors using trajectory log files and EPID measurements for IMRT and VMAT deliveries," *Phys. Med. Biol.* **59**, N49–N63 (2014).
- ²³W. van Elmpt, S. Nijsten, S. Petit, B. Mijnheer, P. Lambin, and A. Dekker, "3D *in vivo* dosimetry using megavoltage cone-beam CT and EPID dosimetry," *Int. J. Radiat. Oncol., Biol., Phys.* **73**, 1580–1587 (2009).
- ²⁴B. Sun, D. Rangaraj, G. Palaniswaamy, S. Sridhar, O. Wooten, D. Yang, S. Mutic, and L. Santanam, "Initial experience with TrueBeam trajectory log files for radiation therapy delivery verification," *Pract. Radiat. Oncol.* **3**, e199–e208 (2013).
- ²⁵Varian TrueBeam Technical Reference Guide, Vol. 1, pp. 54–58. Resolvers provide rotational position of a motor calibrated to a hard stop.