Design of a Reconfigurable Quality Assurance Phantom for Verifying the Spatial Accuracy of Radiosurgery Treatments for Multiple Brain Metastases

Alban C. Cobi
Department of Mechanical Engineering,
Massachusetts Institute of Technology,
77 Massachusetts Avenue,
Cambridge, MA 02139
e-mail: acobi@mit.edu

Luke A. **Gray**
Department of Mechanical Engineering,
Massachusetts Institute of Technology,
77 Massachusetts Avenue,
Cambridge, MA 02139
e-mail: lagray@mit.edu

Elizabeth R. Mittmann
Department of Mechanical Engineering,
Massachusetts Institute of Technology,
77 Massachusetts Avenue,
Cambridge, MA 02139
e-mail: emittman@mit.edu

Steven B. Link
Department of Mechanical Engineering,
Massachusetts Institute of Technology,
77 Massachusetts Avenue,
Cambridge, MA 02139
e-mail: sblink@mit.edu

Nevan C. Hanumara
Department of Mechanical Engineering,
Massachusetts Institute of Technology,
77 Massachusetts Avenue,
Cambridge, MA 02139
e-mail: hanumara@mit.edu

Yulia Lyatskaya
Department of Radiation Oncology,
BWH/DFCI/HMS,
Boston, MA 02115
e-mail: ylyatskaya@bwh.harvard.edu

Ellen Roche
Department of Mechanical Engineering,
Massachusetts Institute of Technology,
77 Massachusetts Avenue,
Cambridge, MA 02139
e-mail: etr@mit.edu

Alexander H. Slocum
Department of Mechanical Engineering,
Massachusetts Institute of Technology,
77 Massachusetts Avenue,
Cambridge, MA 02139
e-mail: slocum@mit.edu

Piotr Zygmanski
Department of Radiation Oncology,
BWH/DFCI/HMS,
Boston, MA 02115
e-mail: pzygmanski@bwh.harvard.edu

Radiation therapy frequently involves highly customized and complex treatments, employing sophisticated equipment, that require extensive patient-specific validation to verify the accuracy of the treatment plan as part of the clinical quality assurance (QA) process. This paper introduces a novel, reconfigurable QA phantom developed for the spatial validation of radiosurgery treatments of multiple brain metastases (MBM). This phantom works in conjunction with existing electronic portal imaging detector (EPID) technology to rapidly verify MBM treatment plans with submillimeter accuracy. The device provides a $12 \times 12 \times 12 \, cm^3$ active volume and multiple, independently configurable markers, in the form of 3 mm diameter radiopaque spheres, which serve as surrogates for brain lesions. The device is lightweight, portable, can be setup by a single operator, and is adaptable for use with external beam radiotherapy (EBRT) techniques and stereotactic linear accelerators (LINACs). This paper presents the device design and fabrication, along with initial testing and validation results both in the laboratory, using a coordinate measuring machine (CMM) and under simulated clinical conditions, using a radiosurgery treatment plan with 15 lesions. The device has been shown to place markers in space with a 0.45 mm root-mean-square error, which is satisfactory for initial clinical use. The device is undergoing further testing under simulated clinical conditions and improvements to reduce marker positional error. [DOI: 10.1115/1.4044402]

1 Introduction

There are over 30×10^6 cancer patients worldwide and this number keeps growing [1]. Twenty to forty percent of these patients will have their cancer spread to the brain (brain metastases) [2]. Lung, breast, renal, and colorectal cancers are most common to metastasize into the brain [2]. The single most common form of treatment for brain metastases is external beam radiation therapy (EBRT), which uses ionizing radiation to kill cancer cells [1], which are commonly small in nature (on the order of mm to cm) and often present in multiple locations (multiple brain metastases (MBM)) [3]. Due to the delicate nature of irradiating the brain, advanced EBRT techniques such as volumetric modulated arc therapy (VMAT) are used to target MBM lesions, while sparing healthy tissue [1,4]. These advanced treatment techniques are preferred because they are based on highly conformal treatment plans, capable of precisely targeting millimeter-sized lesions in a single fraction [5].

To target cancerous lesions, ionizing radiation is delivered to patients using a medical linear accelerator (LINAC) in a hospital setting. Modern LINACs generate microwaves using a magnetron or klystron. The microwaves accelerate electrons, inside the LINAC's gantry, to 6–15 MeV. The accelerated electrons strike a tungsten foil, causing it to emit Bremsstrahlung X-rays. The X-rays are then shaped into a beam by a sequence of coarse apertures and dynamically filtered to ensure optimal dose across the entire irradiation volume [6,7]. Finally, the beam passes through a multileaf collimator (MLC), which dynamically shapes it while
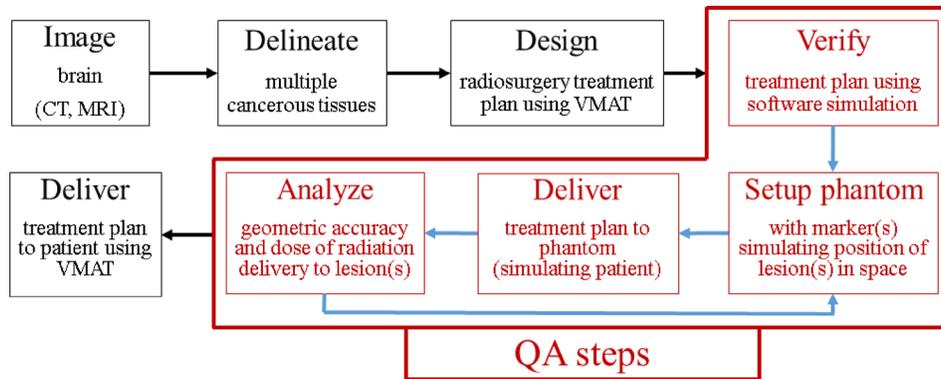
Fig. 1 Flowchart of brain radiosurgery process highlighting QA steps

the gantry rotates about the patient. In this way, the X-ray flux is modulated in three dimensions (3D) to conform to the exact shape of the tumor being targeted [1,8,9].

While the precision of modern LINACs has enabled high-precision treatments, it has not eliminated all of the errors associated with EBRT treatments [4]. Common sources of error associated with EBRT treatments stem from image-guided patient setup, patient movements during treatment, and spatial inaccuracies in the beam caused by LINAC gantry and MLC motions as well as MLC-shaped beam modeling by the medical personnel [10,11]. In a recent radiotherapy report, however, it is observed that errors can arise from a number of different sources, some of which are hospital-specific and some systemic [1]. These sources include decline in LINAC precision over lifetime (10 yr), insufficient training, experience or competence of hospital staff, stressed or fatigued personnel, bad procedural design, overreliance on automation, poor communication and lack of team work, hierarchal department structure that discourages ground-up initiative, working conditions, availability of resources, and presence of distractions and frequent or sudden procedural changes [1]. The cumulative effect of these errors can lead to large differences between intended and delivered dose, and grave consequences for the patient [1,12].

To reduce the sources of error during the radiosurgery process, extensive quality assurance (QA) is required, which has driven up the cost of MBM treatments [13–16]. Specifically, QA is required to verify the patient-specific treatment plan in order to validate the geometric accuracy and dosage before irradiating the patient [12,17]. Figure 1 shows the typical steps in the radiosurgery process using VMAT, with the critical steps of the QA process highlighted in the box.

Medical personnel use various hardware and software-based QA systems daily, monthly, and yearly to calibrate equipment and verify treatment plans [1,12,18,19]. In detail, the following state-of-the-art hardware and software-based techniques are currently used for QA of MBM treatments plans: (a) film two-dimensional (2D) dosimetry [17], (b) 2D detector arrays (diode or ionization chamber arrays) [17], (c) electronic portal imaging detector (EPID) transmission dosimetry [17,20,21], (d) LINAC log-file-based computational dosimetry [19,22], and (e) radiosensitive gel 3D dosimetry [17,19]. Film dosimetry is a hardware-based technique that uses radiosensitive film to show the radiation beam pattern formed by the MLCs [22]. It has a high spatial resolution but only makes 2D measurements, can be time-consuming, and prone to errors due to film calibration and film scanning [17,22]. Commercial detector array QA systems are hardware-based systems that use an array of detectors to measure spatial accuracy and dosage. They have insufficient spatial resolution in the plane of the detector for MBM treatments and cannot provide reliable 3D dose data directly in the locations where the tumors are present [17,21–24]. Devices with multiple detector arrays at various locations can measure dose for all MBM lesions but they are very time-consuming and not feasible for routine verification. EPID systems are hybrid systems that

perform computations based on data from a physical detector aligned with the LINAC's X-ray beamline. They measure the beam shape formed by the MLCs and predict dose with high spatial resolution in 3D but do not verify dose directly at the location of the lesions [25,26]. Similarly, log-file QA systems are software-based systems that perform dose computations solely using MLC position data from sensors on the LINAC and are prone to algorithmic or modeling errors [22]. Radiosensitive gel 3D dosimetry is the most advanced hardware-based technique that uses a volume of radiosensitive gel that polymerizes upon radiation [27]. It provides both high spatial resolution and makes dose measurements in 3D, but it requires optical tomography scanning, which adds cost and is inefficient in routine QA [17,19,27].

In general, patients mimicking physical devices, also known as phantoms, are preferred for QA because they provide a direct measurement of dose and spatial accuracy of treatment plan prior to irradiating the patient [17]. To ensure accurate and repeatable results, phantoms are placed on the treatment couch and aligned and calibrated to the coordinate system of the treatment room. Optical line lasers mounted on the walls of the treatment room and cone beam computed tomography (CBCT) imaging are used to automatically align the phantom with the line lasers [10]. The LINAC coordinate system is defined by carefully positioned line lasers that intersect at the desired origin, called the isocenter of the LINAC [12]. Figure 2 shows a sample LINAC setup on a Varian Edge STx LINAC (Varian, Palo Alto, CA) used for radiosurgery with our phantom on the treatment couch.

Due to the complex and experimental nature of radiotherapy treatments and QA, there is limited published data on the cost-effectiveness of specific QA systems [14–16]. In 2003, a cost-analysis technique called activity-based costing (ABC) was applied to radiotherapy to quantify costs associated with radiotherapy treatments [16]. The technique allocates cost to
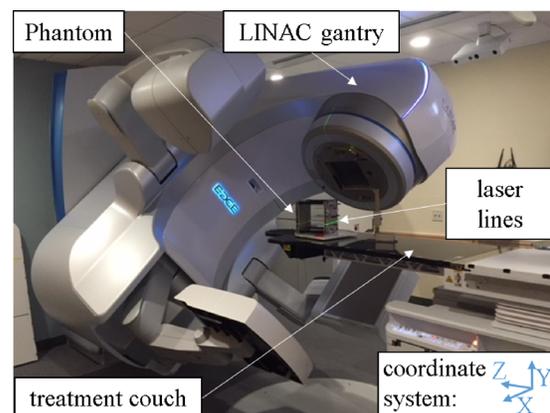

Fig. 2 LINAC setup labeled with major components

activities based on resource consumption. In a recent study, comparing costs of the Leuven University Hospital in Belgium from 2000 to 2009, the ABC technique concluded that the QA costs had the largest rise from 7% to 17% of total costs associated with radiotherapy. It is suggested that the sophistication of new QA equipment has led to this spike in QA costs [14]. Despite these conclusions drawn by the ABC cost-analysis technique, the lack of real data on specific QA systems makes it is difficult to quantify and compare time and cost associated with specific QA systems. Thus, with only anecdotal data, there is no standard on treatment plan QA, leaving radiation oncologists to use QA systems based on needs, experience, and comfortability with each device.

Hardware-based phantoms that use EPID, film dosimetry, detector arrays, or gel dosimetry technology are currently used for MBM treatments but they have the aforementioned limitations. Specifically, film dosimetry and gel dosimetry phantoms require scanners, which add extra time, cost, and potential for errors to the QA process. In contrast, detector array systems can be used stand-alone but require extensive time for routine QA due to their complexity. Additionally, detector array-based phantoms have integrated electronics, which poses a potential risk of damage with complicated treatment plans. Last, in the present commercial detector array, or EPID QA systems, simulation of treatment with nonzero couch angles is not possible, which is a serious limitation of such systems in end-to-end tests. In general, there is no "one size fits all" commercially available brain phantom due to the complexity and time it takes to configure current phantoms [1,3]. As a consequence, in practice, a combination of commercially available or custom-made, in-house QA systems is typically used to calibrate the LINAC and verify the treatment plan.

Because of these limitations, the gold standard is to employ a simple-to-use patient-specific phantom with fiducials, detectors, or dosimeters placed at the treatment locations and directly measure dose with sufficient spatial accuracy to validate the entire treatment plan in a single fraction. In this context, both high spatial resolution and an efficient QA system for a direct measurement of targeting of MBM lesions in a single treatment are essential for maintaining high QA of the MBM treatments.

We conjecture that a higher standard of QA, yet to be offered by any prior device, could be accomplished by enabling "full rehearsal." In other words, an end-to-end QA procedure in which the phantom exactly replicates the topology of a patient's lesions and allows all mechanical, optical, and radiative components of the LINAC setup to act in the exact same sequence, as they will during treatment of the actual patient, would provide a higher standard of QA. To achieve this, we propose a highly accurate mechanical phantom that is capable of placing physical markers to simulate lesions, in 3D space. Used with existing EPID technology, available on all modern LINACs, the QA system will verify both, dosage delivered and spatial accuracy of the radiation beam at the direct location of each lesion, thus verifying the total accuracy of the treatment plan.

This paper presents the design, fabrication, and testing of a manual, easily customizable, and reconfigurable proof-of-concept phantom capable of representing multiple, patient-specific MBM targets in 3D space. The phantom has sufficient resolution to validate the spatial accuracy of treatment plans in a single step, which represents an advance over existing phantoms that must be used in a piecewise manner, requiring multiple irradiations to the same treatment plan, or in conjunction with other equipment. The phantom relies on manual setup to combat an overreliance on automation. The phantom will reduce error from lack of training by offering a fast and easy procedure for placing markers. In addition, the phantom is customizable and reconfigurable to adapt to any patient and any EBRT technique and increase the efficiency and verification of the pretreatment QA process. In practice, it is able to accurately position of up to 20 fiducial markers in a $12 \times 12 \times 12 \, \text{cm}^3$ active volume, with minimal obstruction and attenuation of the irradiating beam by the structure of the phantom. The phantom will thereby allow for "full rehearsal," a higher standard of QA.

2 Methodology

The design process utilized a coarse-to-fine methodology beginning with broad functional requirements, determined from the clinical needs and the limitations of current QA methodologies, then reduced to target specifications. Various distinct designs for the strategies were presented to the clinicians and peer-reviewed prior to the fabrication and testing of the current prototype.

2.1 Functional Requirements. The phantom's requirements, outlined in Table 1, are comprised of both technical and, equally important, usability requirements as specified by the clinical team. Requirements nos. 1–4 are hard requirements that stem from the clinicians' experience with various state-of-the-art phantoms and radiotherapy techniques as well as the knowledge of radiobiological response of normal brain and brain metastases to treatment. Specifically, the 0.2 mm positional error requirement (no. 2) was chosen because it is in agreement with previous studies and it is one order of magnitude smaller than the maximum acceptable error of 1 mm associated with LINAC-based radiosurgery. It has been shown in radiosurgery literature that the actual treatment plan delivered to the patient can have errors stemming from CBCT-based patient setup, patient movements, and LINAC gantry sag and MLC motions alone that can exceed 1 mm [28,29]. Therefore a phantom that is accurate down to 0.2 mm is sufficient.

The 5 deg increment rotation in requirement no. 5 was not a hard requirement but is very important in enabling QA for

Table 1 Requirements for phantom

No.	Requirement	Requirement details
1	Independent marker placement in active volume	To represent the MBMs in various locations in the brain, the position of each marker, up to 12 markers, must be independently adjustable in 3D space using a Cartesian coordinate system (X, Y, Z), in an active treatment volume of no less than $12 \times 12 \times 12 \, \text{cm}^3$.
2	Marker resolution and positional error	Markers can have discrete 5 mm resolution in the X and Y axes, but must have continuous position control in the Z-axis (cranio-caudal axis). The total allowed marker positional error was estimated to be up to 0.2 mm in any arbitrary direction from the desired position of the marker.
3	Radiolucent and resistant to radiation	The entire phantom, excluding the markers, must be radiolucent—transparent to megavoltage radiation—with the markers able to be clearly seen in X-ray images. In addition, every component of the phantom must be made of materials that have long-term resistance to radiation damage, e.g., radiation can warp and embrittle certain plastics.
4	Setting and locking the X, Y, and Z axes	The X, Y, and Z axes must have clear labels and provide a means for the operator to easily set the marker positions. Once set, at the desired (X, Y, Z) coordinates, all the markers must be locked in place and not move during simulation of the patient treatment.
5	Phantom rotation and position features	The entire phantom must rotate manually at discrete 5 deg increments, relative to the LINAC, while on the treatment couch, and must have crosshair features to align with the room laser positioning system.
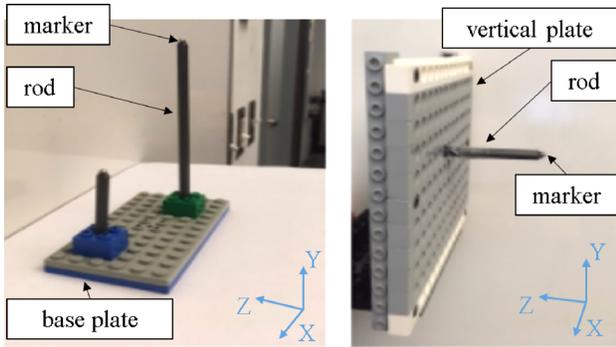
**Fig. 3 Initial prototypes displaying matrix of discrete rods with attached markers**

patient-specific treatment plans with multiple treatment couch irradiation angles. Allowing the phantom to precisely be positioned at different angles, while the couch is aligned with the Z-axis, makes delivery of all treatment arcs possible due to couch-EPID clearance. A common issue associated with nonzero couch angles is that the EPID detector collides with the couch during gantry motions. This precludes EPID for treatment plans requiring arcs with a rotated couch. To our knowledge, no commercial phantoms are able to be discretely or continuously rotate about their Y-axis while on the treatment couch to allow for rapid QA at various couch angles.

The main goal of the phantom was to serve as a reconfigurable QA device able to simulate multiple brain lesions at custom locations, in 3D space, thus allowing clinicians to validate the treatment plan in a single fraction. With this goal in mind, the team set out to build a proof-of-concept device that is robust and compatible with high-definition MLC LINACs (2.5 mm MLC leaf width) equipped with EPID and CBCT with 1 mm accuracy. In this context, robust is defined as capable of being setup by a single operator in under 30 min, and lasting for up to the lifetime of the LINAC (10 yr).

2.2 Concept. Using the requirements listed in Table 1, various low fidelity prototypes were constructed out of LEGO^{TM} bricks (LEGO, Billund, Denmark). Shown in Fig. 3, the most promising concept for marker placement that emerged comprised two discreetly adjustable axes and one continuously adjustable axis. At left, the markers are placed extending upward (in the Y-direction) from a base plate, as a function of the addition on discrete segments (blocks), while at right, they slide in and out (in the Z-direction) of a vertical plate with more flexibility in the Z-direction. Additional concepts with electromechanical actuation were considered, but a prepositioned, mechanical solution provided a more robust, less complex path toward meeting the radiolucent and resistance to radiation requirement. Additional feedback from LINAC operators indicated that they would be satisfied with a "set and forget" system and that minimizing complexity of operation and storage was important. The design at right was selected primarily to decrease complexity, by using a single sized rod. Following additional analysis regarding constraints, trade-offs, and with further input from LINAC operators, all of the individual components of the phantom underwent rigorous engineering design and error budgeting to address the functional requirements.

3 Design and Fabrication

3.1 Phantom Design Overview. The full phantom design is shown in Fig. 4, with individual modules labeled on the CAD rendering. On the left, the phantom is shown as it would sit on the LINAC couch. On the right is the Z-set fixture, which is used for setting the position of the markers along the Z-axis. Each module was engineered to minimize its contribution to the overall error
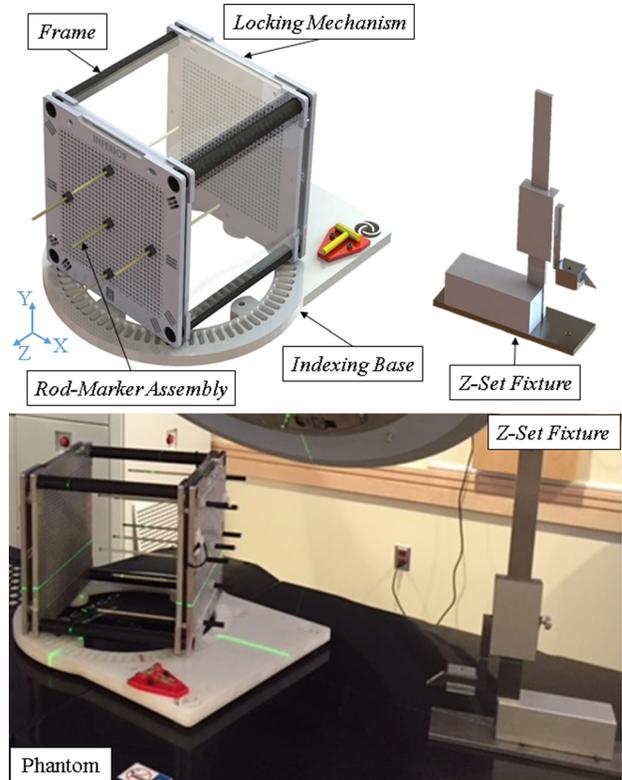

**Fig. 4 A CAD rendering (top) and the fabricated (bottom) phantom assembly and Z-set fixture**

budget, which was used to predict marker positional error, described on a per-module basis.

3.2 Module Design. *Frame*: The frame, shown in Fig. 5, consists of two walls (bulkheads), structurally connected with carbon fiber (CF) tubes. A clear acrylic shell surrounds the frame to protect the markers and is scribed with crosshair marks to meet requirement no. 5 (Table 1). The crosshairs are used in conjunction with the treatment room's line lasers, marking the center of the LINAC (isocenter), and the room coordinate system to align the phantom's center with the LINAC's isocenter. The CF tubes were chosen to maximize bending moment, while minimizing material that could interfere with imaging, so as to meet requirement no. 3.

Each bulkhead consists of a pair of 6.35 mm thick, precision ground flat aluminum plates that are press fit together. The two
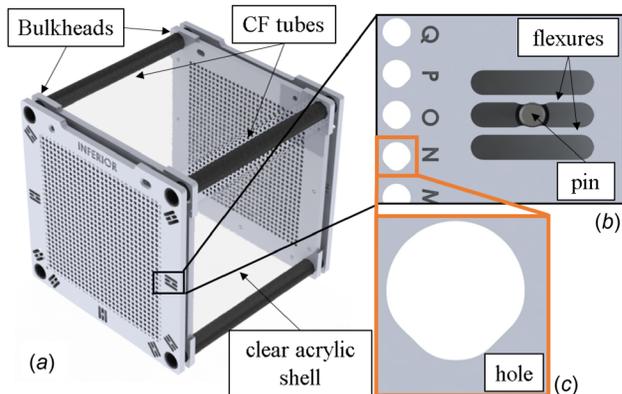
Fig. 5 A CAD rendering of (a) the phantom's frame, (b) pin and flexures for bulkhead plate alignment, and (c) irregular shaped hole geometry
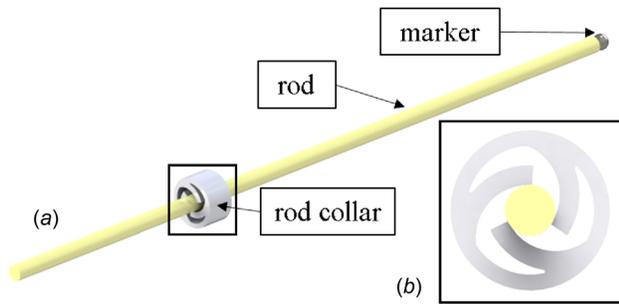
Fig. 6 A CAD rendering of the rod-marker assembly (left) and rod collar (right)

plates of each bulkhead are aligned with nine pins and spacers that project from one plate and engage flexures on the mating plate, thus using elastic averaging to avoid errors and stresses due to over constraint [30,31]. The plates are separated by 6.54 mm; which accommodates the *locking mechanism*. To meet requirement no. 2, each plate is perforated with holes spaced 5 mm apart on each axis on a 29×29 grid. Each hole's geometry is nominally circular with two flats to enable exact constraint of each *rod-marker assembly*, when positioned in the hole. Last, projecting from the bottom of each bulkhead are two hemispherical contacts. These couple with the V-shaped grooves on the *indexing base* to provide the rotational positioning to meet requirement no. 5. The bulkheads are the most intricate part of the frame, designed to contribute only a few microns of marker positional error, as a function of tolerances in the hole diameter, and plate misalignment. This error was assumed to be negligible.

Rod-Marker Assembly: Each independent assembly, as shown in Fig. 6, consists of a 3 mm tungsten-carbide marker that is bonded to the end of a 150 mm alumina rod fitted with a custom polycarbonate rod collar that sets the rod's Z-axis position. The rod collar has a three-prong, internal flexure feature [32,33] that provides a controlled interference fit with the rod. The outer collar diameter was kept under 10 mm, large enough to be manually manipulated by an operator, but small enough to only block the four directly adjacent holes on the bulkhead. The primary source of error in the rod-marker assembly was the straightness of the alumina rods, specified by the manufacturer (CoorsTek, Golden, CO), which contributed 0.1 mm of expected error at maximum insertion. Other errors due to bending under gravity, marker outer diameter tolerance, marker-to-rod assembly, marker sphericity, and thermal expansion were at least one order of magnitude less.

Locking Mechanism: The locking mechanism shown in Fig. 7, is comprised of a Durometer 50 silicone elastomer sheet, with the same hole pattern as the bulkheads, which sits inside a polycarbonate frame. This assembly is placed into the space between the walls of the bulkhead and constrained laterally with a vertical sliding fit between the bulkhead's alignment pins and spacers. With the rods passing through both, the bulkhead and the locking mechanism, depressing the polycarbonate frame causes the elastomer to engage and push down on all inserted rods, simultaneously positioning them against the bulkheads' holes' flats, whereupon the locking mechanism is secured in the depressed position with two locking pins. Based on the permitted frame travel, the clearance in the elastomer's holes, and the stiffness of the elastomer, approximately 3 N of normal force is applied to each rod in the Y direction, resulting in a pull-out force of 25 N in the Z direction, which is sufficient to retain the rods during configuration and usage of the phantom. The primary source of error associated with the locking mechanism is bending of the rods caused by the locking force, which contributes a total of 0.0004 mm marker position error in the Y direction. Error due to deformation and lashing of the elastomer sheet in the Z direction is negligible.

Indexing Base: This important feature of the phantom, shown in Fig. 8, allows the operator to simulate patient treatment at
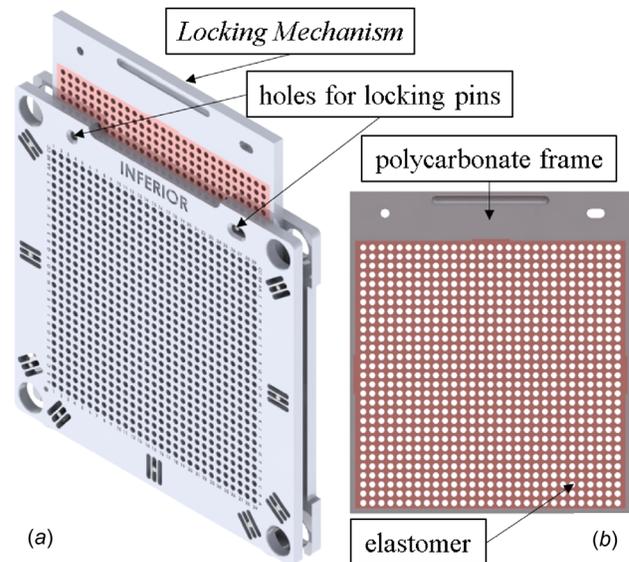
Fig. 7 A CAD rendering of (*a*) bulkhead showing locking mechanism and locking pins and (*b*) features of locking mechanism
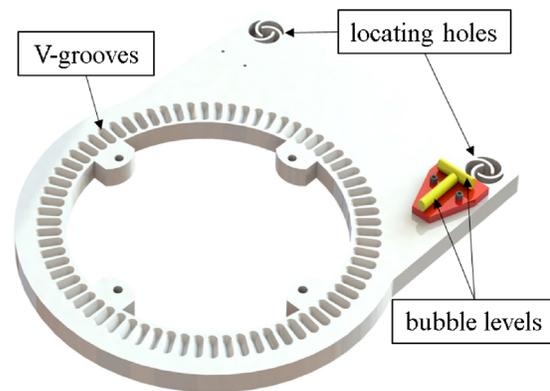
Fig. 8 A CAD rendering of indexing base for phantom to be rotated at 5 deg increments

different couch angles. To accomplish this, the phantom sits on an indexing base that is equipped with two locating holes that grip the pegs protruding from the LINAC treatment couch bar. The base then allows the phantom to be positioned and rotated in discrete 5 deg increments, to enable simulation of a multitude of treatment plans. The base is made of 12 mm thick Delrin chosen for durability and radio-transparency. Machined V-grooves engage with the hemi-spherical contacts in the base of the phantom frame. Two are full hemispheres and two are half hemispheres, for a total of six points of contact; thus, the frame is exactly constrained to the base using a split-groove kinematic coupling [34,35]. This design provides reliable angular engagement, without wobble. Two bubble levels provide the operators a visual confirmation that the phantom is nominally level. Because the phantom position is subsequently zeroed to the LINAC's isocenter, the base does not contribute to marker positional error.

Z-Set Fixture: This module is used to position the rod collar along the length of each rod, thus setting each marker's Z-axis coordinate. This fixture, as shown in Fig. 9, consists of a commercially available vertical height gauge, accurate to 0.01 mm (iGaging Digital Electronic Height Gauge with magnetic base, 6 in.), a base with a conical hole that engages the marker, and a machined block connected to the gauge's arm that holds each rod. When using this device, the operator sets gauge at the desired Z-axis coordinate for the marker using the scale on the
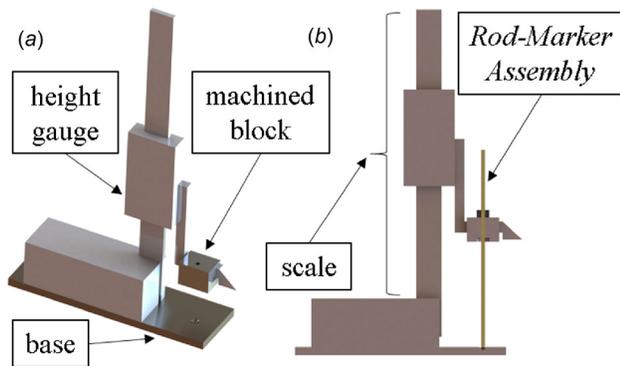
Fig. 9   Z-set fixture for rod-marker assembly

Table 2   Theoretically calculated errors and values

Type	Value
Theoretical sum of error in *X*-direction	0.24 mm
Theoretical sum of error in *Y*-direction	0.25 mm
Theoretical sum of error in *Z*-direction	0.13 mm
Total maximum theoretical expected error	0.37 mm

device, inserts a rod-marker assembly into the hole on the machined block, and pushes the rod-marker assembly down until it hits the base of the gauge, leaving the collar in the correct position. The rod-marker assembly is then transferred to the appropriate (X, Y) position hole on the phantom and the process is repeated until all rods are placed to map the multiple lesion centers. The sources of error in the Z-set fixture arise from the gauge manufacturer's specified accuracy, surface contamination during use, flatness of the base plate, and tolerance on the hole of the machined block, which when added up, contribute a total of 0.085 mm error in the Z direction.

3.3   Fabrication. The phantom was completely fabricated and assembled in-house at MIT. The bulkhead plates were CNC milled using a Haas VF2 vertical mill (Haas Automation Inc., Oxnard, CA). The locking mechanism frame, elastomer, and rod collars were cut on an OMAX 5555 Waterjet. The indexing base's outline was cut on an OMAX waterjet with the grooves and other features postmachined on a Prototrak Mill. The custom parts of the Z-set fixture were manually milled on the Prototrak Mill and then bonded to the height gauge with Loctite 495 Instant Adhesive. The frame was assembled on a surface plate using gauge

blocks for perpendicular alignment and adhered to the CF tubes with Loctite EA E-30CL Epoxy. Prior to assembly, all surfaces were cleaned with acetone and precision ground flat stones to get rid of small debris. Finally, the acrylic shell was adhered to the phantom using Loctite EA E-30CL Epoxy and postassembly the crosshairs were CNC scribed with the Prototrak mill.

3.4   Error Budget. The reconfigurable phantom's primary performance metric is the ability to accurately position each marker; therefore, the errors for each module, discussed in the preceding section, were summed to arrive at the total maximum theoretical expected error, defined as the Euclidian distance from the center of each spherical marker to the theoretical true position of the center of each marker, shown in Table 2. This did not take into account human errors during phantom setup and the manufacturing process, in particular the bulkhead assembly process and the frame assembly process, which were later shown to be significant contributors to error (Figs. 5–9).

3.5   Phantom Setup. This section describes the procedural flow for planning, setup, and use of the phantom, during the typical QA steps outlined in Fig. 1. Figure 10 shows a procedural flowchart for setting up the phantom prior to irradiation.

After the steps in Fig. 10 are executed, the clinicians exit the treatment room and enter the control room where they can safely irradiate the phantom and view the live treatment with EPID imaging in cine (video) mode. EPID imaging enables the clinicians to see the MLC-shaped treatment beams enveloping the markers as a function of delivery time and dose. Each frame of the EPID video can then be fed into an image processing algorithm, where the theoretical position of the MLC apertures and markers can be compared to the true positions to compute the errors. The clinicians can use this data to adjust the treatment plan accordingly. At the present stage of development, image processing to compute errors was not pursued.

4   Experimental Setup

Laboratory Validation: First, the bulkhead's plates were checked for misalignment, which would cause the inserted rods to angle out of perpendicularity. The assembled frame was placed on a precision surface plate and, using an iGaging Digital Electronic Height Gauge with 0.01 mm accuracy, each plate's height was measured relative to its mating plate. Then the frame was rotated 90 deg and the measurement repeated. This identified the *X–Y* skew between plates, from which the angular misalignment of an inserted rod could be calculated, along with the consequent Abbe error in marker position when at maximum Z insertion.

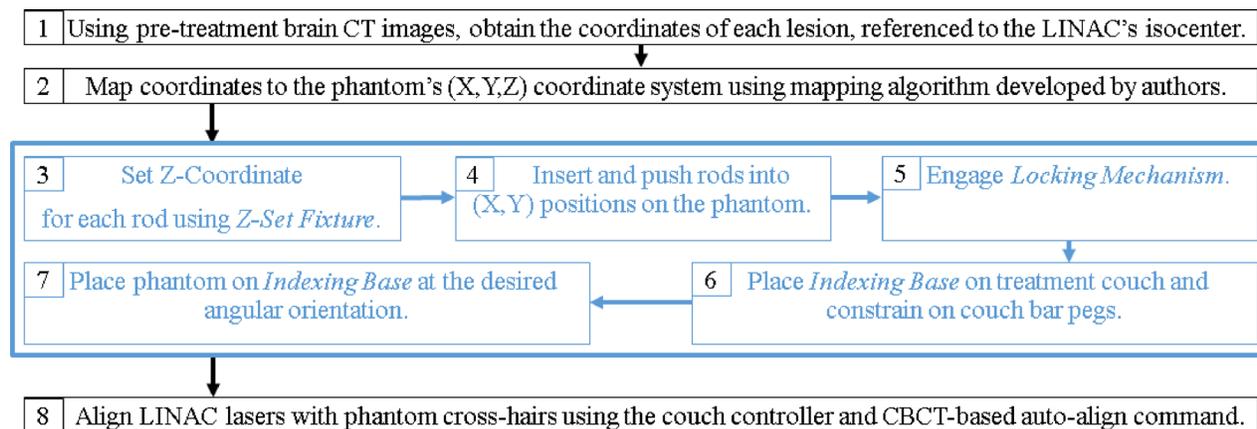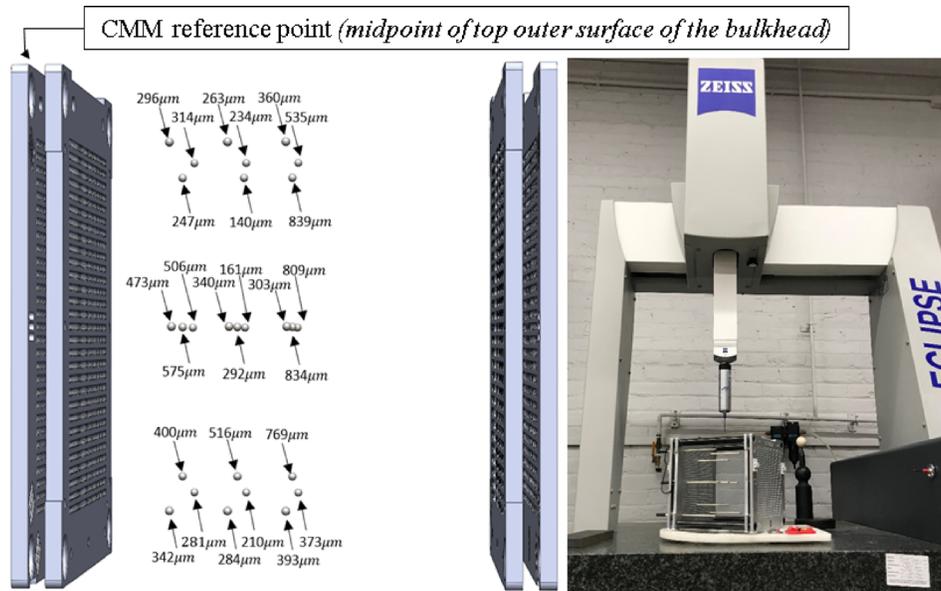Second, measurements using a ZEISS eclipse coordinate measuring machine (CMM) (ZEISS, Oberkochen, Germany) were

Fig. 10   Flowchart showing steps of radiotherapy QA process, highlighting steps for setting up the phantom with markers

Fig. 11 (Left) The 27 marker positions probed with a CMM with the error indicated. Nine (X, Y) locations were used with the markers inserted to three different Z positions. (Right) Test setup in the CMM.

performed to evaluate the fully assembled phantom's marker positional error, with markers placed in a test configuration. As mentioned earlier, the error in position of each marker was defined as the absolute distance from the center of each spherical marker to the theoretical true position of the center of each marker. The error at 27 positions, with different X, Y, and Z coordinates, within the phantom's active area were tested to attain an overall characterization of the marker placement accuracy within the device. The CMM's reference point was the midpoint of the top outer surface of the bulkhead into which the rods were inserted. Since the CMM operates with a threshold trigger force of 300 mN, which could cause deflection in the cantilevered rods, multiple measurements were taken around each marker and the center position calculated.

Clinical Validation: The practical use of the phantom was evaluated in a clinical setting at Brigham and Women's Hospital in Boston, MA. A mock patient treatment plan with 15 lesions was selected to irradiate the phantom at 0 deg couch angle with the gantry rotating around the Z-axis (cranio-caudal axis). The phantom was configured and positioned as described in Sec. 3.5. The plan was delivered on Varian Edge STx LINAC (Varian, Palo Alto, CA) and marker placement was recorded using EPID (Varian, Palo Alto, CA) in cine (video) mode.

5 Results and Discussion

5.1 Laboratory Validation Results. The largest measured misalignments between the two plates of each bulkhead, measured with the iGaging Digital Electronic Height Gauge on a surface block, were 0.16 mm and 0.07 mm along the X and Y axes, respectively. The resulting Abbe errors solely due to this bulkhead plate misalignment, at the tip of a fully inserted rod-marker assembly, were calculated as 1.39 mm and 0.77 mm in the X and Y axes, respectively. This shows that bulkhead misalignment introduced a larger marker positional error than predicted.

Figure 11 shows the distribution of the marker positions within the active error of the phantom and the errors measured at each position using the CMM. A maximum error of 0.834 mm and a root-mean-square error of 0.45 mm were computed over the 27 different measurements, with a standard deviation of 0.20 mm. The average error across all of the positions was measured to be

0.41 mm. The native uncertainty of the CMM is an order of magnitude lower at 0.025 mm. The largest error measurements were observed at the halfway point between the plates, where the Z-coordinate was maximum. This is consistent with the expectation that Abbe Error from misaligned plates should dominate and grow with the length of the cantilever.

The height gauge measurements point to large errors in bulkhead plate misalignment as a primary reason that the 0.2 mm marker positional error requirement was not met. While it was predicted that elastic averaging of the machined flexures would result in micron-level alignment between the two plates [30,31], we hypothesize this misalignment could be due to several causes. The potential causes of this are predicted to be human errors during machining or contamination errors during the assembly of the bulkhead plates. Each bulkhead was machined separately and required 12 h of machining time. Due to this large machining time, we believe errors in edge-finding or shifting of the work stop could have caused the alignment features and consequently the 29×29 grid of holes to be offset. Alternatively, while care was taken to clean the work area in the machine shop, it was possible particles contaminated the bulkhead plate assembly process.

The CMM measurements confirmed that the phantom has larger than predicted marker positional errors. We hypothesize these errors, which vary spatially, are either caused by the manufacturing process or the phantom setup process. The bulkhead assembly and the frame assembly were the two processes identified in manufacturing, which had the potential for the largest errors. First, the measured bulkhead plate misalignment was sufficient to offset the markers by a maximum of 1.39 mm in the X-axis. Second, the manufacturing assembly process was performed with clamps, straightedges, and tooling blocks on a surface plate. This could be improved by developing a standard operating procedure and custom fixtures to support the components during assembly and curing. Furthermore, the design should be reviewed with respect to design for assembly and assembly performed in a clean room to reduce error from potential particle contamination.

The use of the locking mechanism, the Z-set fixture, and handling of the rod-marker assembly were identified as the steps of the phantom setup (steps 3–5 in Fig. 10) that could introduce large human errors. First, while the locking mechanism locks all rods
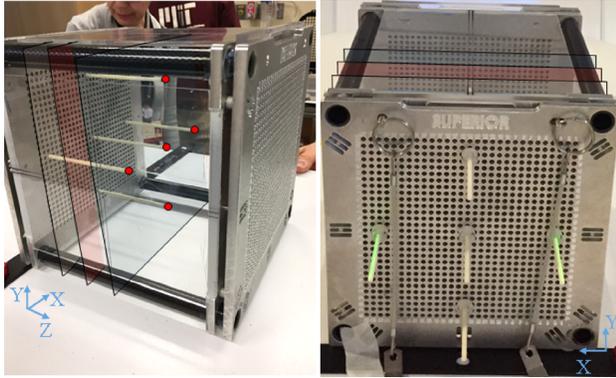
Fig. 12 The phantom setup with five markers in (left) isometric view and (right) view of X–Y plane

securely, it was noted that the rods are free to move prior to engaging the locking mechanism. Any sudden impulse or movement of the phantom during this setup process could cause the rods to shift prior to locking them. Second, failure to lock the manual height gauge of the Z-set fixture or slight movement during use could cause errors in the Z-axis. Finally, accidental shifting of the rod collar on the rod-marker assembly or bumping of the rods or phantom prior to locking the locking mechanism could cause the rods to shift in the Z-axis. These issues will be addressed by redesigning the Z-set fixture and locking mechanism in future prototypes. An automated Z-set fixture could help reduce errors from manually setting and locking the Z coordinate of each marker. Furthermore, solutions will be considered to lock each rod-marker assembly independently to reduce shifting during phantom setup.

5.2 Clinical Validation Results. The phantom was configured with 15 markers, with 5 markers positioned in each of 3 different X–Y planes. Treatment beams were designed to irradiate five markers (on the same X–Y plane) at a time. Figure 12 shows the phantom setup with five markers.

An example of EPID image is shown in Fig. 13, representing a single frame taken in cine (video) mode during the radiation of the mock VMAT plan with the MLC dynamically adjusting to restrict the beam to the specified treatment regions. In Fig. 13(a),

the darker regions correspond to areas of high X-ray flux, while the brighter regions are either collimated by the MLC (and not irradiated) or attenuated by the tungsten-carbide markers. Inside each of the five dark treatment apertures can be seen a white spot corresponding to a marker. To better visualize the low resolution results in Fig. 13(a), the image was processed using PERFRACTION (Sun Nuclear Corporation, Melbourne, FL) to show the locations of the markers inside the treatment beams in Fig. 13(b). This processed frame provides a means to readily confirm the actual locations of MLC apertures with respect to the actual marker positions. These actual positions of the MLC apertures and markers can be compared to ideal positions in the treatment planning software (TPS) as shown in Fig. 13(c) and positional errors between the two can be determined (Fig. 13(d)).

At the present stage of development, extensive image analysis software for precise determination of MLC aperture positions and quantifying errors in position and dosimetry was not pursued, but will be the focus of future clinical research. However, this visualization of the markers, demonstrates their potential use as stationary reference points for the dynamically conforming MLC, as a function of delivered dose and LINAC gantry angle. Most importantly, the results show the phantom fulfills its main goal of precisely representing multiple targets in 3D space for treatment plan QA. The basic outline of how these results can be used in a clinical setting, pending further software development to accompany this phantom, is provided below. The procedure to determine the positional errors of the MLC apertures and overall treatment plan is as follows:

(1) Export EPID video frames and use software to automatically determine the actual positions of all markers in the EPID frames as a function of LINAC gantry angle and time.
(2) Use EPID frames to obtain the actual locations of the MLC apertures, with respect to the reference points, as a function of LINAC gantry angle.
(3) Compare the actual MLC aperture locations and the marker locations to the ideal locations defined in the TPS.
(4) Calculate the positional and LINAC gantry errors for each lesion. It is important to note that, radiation oncologists will not accept positional errors significantly larger than 1 mm in targeting each lesion.
(5) Provide feedback for the clinical decision involving various treatment scenarios: Adjust treatment plan, improve CBCT-
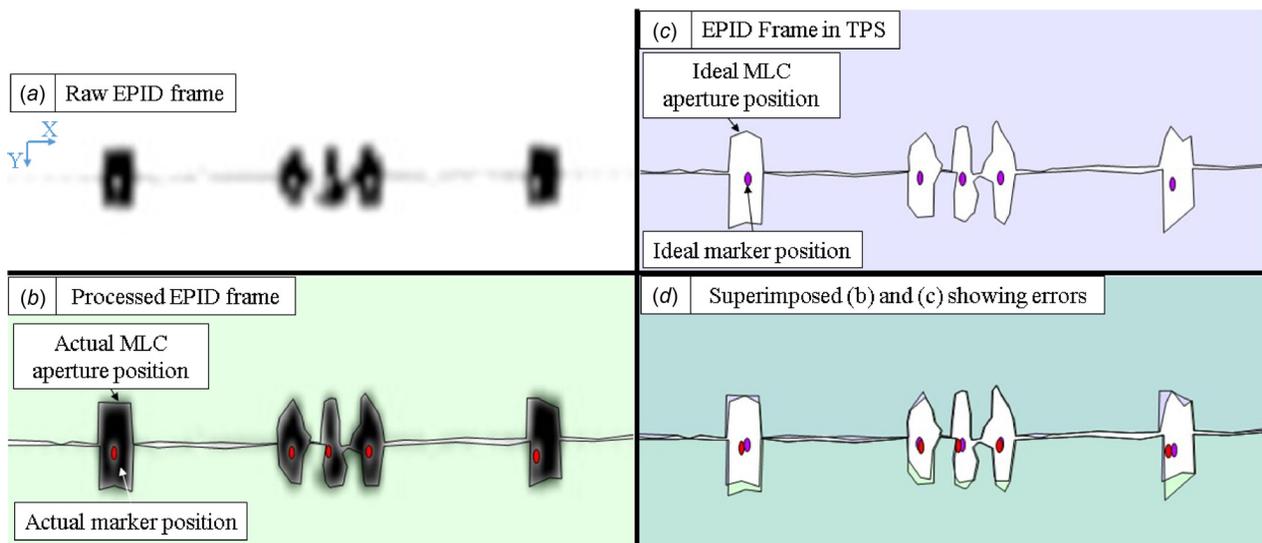
Fig. 13 A frame of EPID from cine (video) mode shown as (a) raw frame, (b) processed frame to show the actual marker positions and MLC aperture positions, (c) frame in TPS to show ideal marker and MLC aperture positions, and (d) the processed and TPS frames superimposed showing positional errors

based setup, repair or calibrate LINAC gantry and MLC, and address potential human errors accordingly to account for these errors.

The procedure to determine dosimetry errors is as follows:

(1) Export EPID video frames and use software to automatically determine the actual positions of all markers in the EPID frames as a function of LINAC gantry angle and time.
(2) Use EPID frames to obtain the actual locations of the MLC apertures, with respect to the reference points, as a function of LINAC gantry angle.
(3) Compute dosage to each lesion using an independent dose calculation algorithm (similar to commercially available software such as PERFRACTION by Sun Nuclear Corporation), using the known true positions of the lesions, represented by the markers of the phantom.
(4) Adjust treatment plan parameters accordingly to account for dosimetric errors.

In addition, considering the need for robustness and operation by a single operator, the clinical team evaluated the setup efficiency of the phantom in the actual clinical environment by timing a single trial of the mock treatment plan. The QA steps of the treatment plan with 15 markers took 20 min to perform. For comparison, it usually takes the clinicians about 15–30 min to setup and warm up the electronics of the commercially available Delta4 Phantom (ScandiDos, Uppsala, Sweden), up to an additional 5 min for each lesion targeted (due to phantom shifts on the treatment couch), and 10 min to take it down and store it elsewhere, due to its size and weight (27 kg) and the need for a transport cart (added cost of equipment as this is specialized cart). Comparatively, our phantom can validate the treatment plan for all lesions simultaneously, weighs less than 5 kg and can be stored inside the treatment room due to its compact size and weight. The team believes that the setup time of the phantom could be reduced by as much as 50% if parts of the phantom setup such as the Z-set fixture and the coordinate mapping algorithm are automated in future versions. In addition, time could also be reduced by allowing clinicians to have more experience using the phantom.

6 Conclusions and Future Work

Overall, we can conclude that this proof-of-concept phantom accomplished its primary objective of serving as a reconfigurable QA phantom able to simulate multiple brain lesions at custom locations, in 3D space. Furthermore, we showed that a reconfigurable and purely manual phantom with sufficient accuracy for patient-specific QA can be achieved. Based on the laboratory and clinical validation, the phantom met the requirements of securely placing fiducial markers arbitrarily within the treatment volume (#1, 4), while not degrading the image (#3), and successfully aligning with the treatment room's coordinates (#5). However, the maximum measured marker positional error was approximately four times greater than the requirement (#2) and predicted by the error budget. The possible causes of this were discussed and it is predicted that this can be addressed in future iterations. In practice, radiotherapy errors up to 1 mm are generally accepted by surgeons so requirement no. 2, set by the clinical team may be more stringent than necessary.

The primary potential users of this phantom, clinicians at Brigham and Women's Hospital in Boston, MA indicated that it possesses the required clinical features and, specifically, is readily configurable within the QA steps of MBM treatment plans using VMAT. Furthermore, the clinical tests showed that the phantom has several important advantages over existing commercially available phantoms. First, the phantom is able to be setup with multiple markers by a single person in 20 min. Second, the

phantom can be imaged with multiple markers simultaneously in a single fraction. Third, the images could be used by clinicians to determine positional and dosimetry errors associated with the treatment plan based on the actual locations of the markers and MLC apertures. And finally, the clinicians indicated that one of the most important features of the phantom was the indexing base, giving the clinicians the ability to perform an end-to-end test with all couch angles as in the patient treatment plan, a procedure that is not feasible with other devices. We believe, the phantom introduced in this research shows promise as being the first reconfigurable QA phantom for MBM treatment plan verification in a single pass. As with any medical device, a further usability study should be undertaken to ensure reliable operation by LINAC operators.

The device is currently undergoing further clinical tests and the team has already explored options for improving the accuracy of future devices. Future work will focus on adding design for assembly features, developing standard operating procedures, and designing custom fixtures to reduce assembly errors and redesigning the Z-set fixture and locking mechanism, to reduce phantom setup errors. Postredesign, end-to-end repeatability studies with the CMM and in a clinical setting will be performed to help determine potential errors introduced in the manufacturing process and during the phantom setup. Further usability tests will guide the designs of these modules to make the phantom easy to use and intuitive with the ultimate goal of having this phantom utilized for monthly, and annual treatment plan QA. This versatility will streamline procedural design and ultimately help in increasing the efficiency of the QA process and decreasing the possibility of mistreatment associated with MBM radiosurgery.

As technologies used for radiation therapy are in constant evolution, we believe a relatively inexpensive, simple, highly reliable mechanical device that may provide process stability and gain widespread adoption because of low cost barrier to entry.

References

[1] Maliwal, N., 2017, "Radiotherapy: Technologies and Global Markets," BCC Research, Wellesley, MA, Report No. HLC176B.
[2] Wong, J., Hird, A., Kirou-Mauro, A., Napolskikh, J., and Chow, E., 2008, "Quality of Life in Brain Metastases Radiation Trials: A Literature Review," Curr. Oncol., 15(5), pp. 25–45.
[3] Andrevska, A., Knight, K. A., and Sale, C. A., 2014, "The Feasibility and Benefits of Using Volumetric Arc Therapy in Patients With Brain Metastases: A Systematic Review," J. Med. Radiat. Sci., 61(4), pp. 267–276.
[4] Bhide, S. A., and Nutting, C. M., 2010, "Recent Advances in Radiotherapy," BMC Med., 8(1), p. 25.
[5] Clark, G. M., Popple, R. A., Young, P. E., and Fiveash, J. B., 2010, "Feasibility of Single-Isocenter Volumetric Modulated Arc Radiosurgery for Treatment of Multiple Brain Metastases," Int. J. Radiat. Oncol. Biol. Phys., 76(1), pp. 296–302.
[6] Khan, F. M., 2014, Khan's the Physics of Radiation Therapy, Lippincott Williams & Wilkins (LWW), Philadelphia, PA.
[7] Anderson, R., Lamey, M., MacPherson, M., and Carlone, M., 2015, "Simulation of a Medical Linear Accelerator for Teaching Purposes," J. Appl. Clin. Med. Phys., 16(3), pp. 359–377.
[8] Yu, C. X., 1995, "Intensity-Modulated Arc Therapy With Dynamic Multileaf Collimation: An Alternative to Tomotherapy," Phys. Med. Biol., 40(9), pp. 1435–1449.
[9] Brewster, L., Mohan, R., Mageras, G., Burman, C., Leibel, S., and Fuks, Z., 1995, "Three Dimensional Conformal Treatment Planning With Multileaf Collimators," Int. J. Radiat. Oncol. Biol. Phys., 33(5), pp. 1081–1089.
[10] Ezzell, G. A., 2017, "The Spatial Accuracy of Two Frameless, Linear Accelerator-Based Systems for Single-Isocenter, Multitarget Cranial Radiosurgery," J. Appl. Clin. Med. Phys., 18(2), pp. 37–43.

[11] Huang, Y., Chin, K., Robbins, J. R., Kim, J., Li, H., Amro, H., Chetty, I. J., Gordon, J., and Ryu, S., 2014, "Radiosurgery of Multiple Brain Metastases With Single-Isocenter Dynamic Conformal Arcs (SIDCA)," Radiother. Oncol., **112**(1), pp. 128–132.

[12] Schell, M. C., Bova, F. J., Larson, D. A., Leavitt, D. D., Lutz, W. R., Podgorsak, E. B., and Wu, A., 1995, "Stereotactic Radiosurgery," Report of Task Group 42, AAPM, Inc. Woodbury, New York, Report No. 54.

[13] Mariotto, A. B., Robin Yabroff, K., Shao, Y., Feuer, E. J., and Brown, M. L., 2011, "Projections of the Cost of Cancer Care in the United States: 2010-2020," J. Natl. Cancer Inst., **103**(2), pp. 117–128.

[14] Van De Werf, E., Lievens, Y., Verstraete, J., Pauwels, K., and Van Den Bogaert, W., 2009, "Time and Motion Study of Radiotherapy Delivery: Economic Burden of Increased Quality Assurance and IMRT," Radiother. Oncol., **93**(1), pp. 137–140.

[15] Van De Werf, E., Verstraete, J., and Lievens, Y., 2012, "The Cost of Radiotherapy in a Decade of Technology Evolution," Radiother. Oncol., **102**(1), pp. 148–153.

[16] Lievens, Y., Van Den Bogaert, W., and Kesteloot, K., 2003, "Activity-Based Costing: A Practical Model for Cost Calculation in Radiotherapy," Int. J. Radiat. Oncol. Biol. Phys., **57**(2), pp. 522–535.

[17] Low, D. A., Moran, J. M., Dempsey, J. F., Dong, L., and Oldham, M., 2011, "Dosimetry Tools and Techniques for IMRT," Med. Phys., **38**(3), pp. 1313–1338.

[18] Klein, E. E., Hanley, J., Bayouth, J., Yin, F.-F., Simon, W., Dresser, S., Serago, C., Aguirre, F., Ma, L., Arjomandy, B., Liu, C., Sandin, C., and Holmes, T., 2009, "Task Group 142 Report: Quality Assurance of Medical Accelerators," Med. Phys., **36**(9), pp. 4197–4212.

[19] Zygmanski, P., Wagar, M., Maryanski, M., and Hacker, F., 2018, "Current Limitations and Emerging Solutions for Quality Assurance (QA) of Single-Isocenter VMAT Treatment of Multiple Brain Metastases (MBM)," Med Phys., **45**(6), pp. E283–E283.

[20] Sumida, I, Yamaguchi, H., Kizaki, H., Koizumi, M., Ogata, T., Takahashi, Y., and Yoshioka, Y., 2012, "Quality Assurance of MLC Leaf Position Accuracy and Relative Dose Effect at the MLC Abutment Region Using an Electronic Portal Imaging Device," J. Radiat. Res., **53**(5), pp. 798–806.

[21] Bäck, A., 2015, "Quasi 3D Dosimetry (EPID, Conventional 2D/3D Detector Matrices)," J. Phys. Conf. Ser., **573**(1), p. 012012.

[22] Rowshanfarzad, P., Sabet, M., Barnes, M. P., O'Connor, D. J., and Greer, P. B., 2012, "EPID-Based Verification of the MLC Performance for Dynamic IMRT and VMAT," Med. Phys., **39**(10), pp. 6192–6207.

[23] Niroomand-Rad, A., Blackwell, C. R., Coursey, B. M., Gall, K. P., Galvin, J. M., McLaughlin, W. L., Meigooni, A. S., Nath, R., Rodgers, J. E., and Soares, C. G., 1998, "Radiochromic Film Dosimetry: Recommendations of AAPM Radiation Therapy Committee Task Group 55," Med. Phys., **25**(11), pp. 2093–2115.

[24] Baldock, C., Deene, Y. D., Doran, S., Ibbott, G., Jirasek, A., Lepage, M., Oldham, M., and Schreiner, L. J., 2011, "Topical Review: Polymer Gel Dosimetry," Phys. Med. Biol., **55**(5), pp. 1–87.

[25] Schyns, L. E., Persoon, L. C., Podesta, M., Van Elmpt, W. J., and Verhaegen, F., 2016, "Time-Resolved Versus Time-Integrated Portal Dosimetry: The Role of an Object's Position With Respect to the Isocenter in Volumetric Modulated Arc Therapy," Phys. Med. Biol., **61**(10), pp. 3969–3984.

[26] Bawazeer, O., Herath, S., Sarasanandarajah, S., and Deb, P., 2015, "World Congress on Medical Physics and Biomedical Engineering, June 7–12, 2015, Toronto, Canada," IFMBE Proc., **51**, pp. 553–556.

[27] Watanabe, Y., Warmington, L., and Gopishankar, N., 2017, "Three-Dimensional Radiation Dosimetry Using Polymer Gel and Solid Radiochromic Polymer: From Basics to Clinical Applications," World J. Radiol., **9**(3), pp. 112–125.

[28] Rosca, F., Zygmanski, P., Lorenz, F., Hacker, F., Chin, L., Friesen, S., Petsuksiri, J., Shanmugam, L., and Ramakrishna, N., 2005, "An MLC-Based Linac QA Procedure for the Characterization of Radiation Isocenter and Room Lasers' Position," Med. Phys., **32**(6), p. 2040.

[29] Xu, Y., Silverman, J. S., Du, K. L., Das, I. J., and Kondziolka, D., 2017, "Patient Positioning Accuracy in Stereotactic Radiosurgery With Mask Fixation and Cone Beam CT," Int. J. Radiat. Oncol., **99**(2), pp. E740–E741.

[30] Teo, T. J., and Slocum, A. H., 2017, "Principle of Elastic Averaging for Rapid Precision Design," Precis. Eng., **49**, pp. 146–159.

[31] Willoughby, P., 2005, "Elastically Averaged Precision Alignment," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA.

[32] Kharadi, F. H., Jadhav, M. S., Kanhurkar, S. D., Pereira, P. A., Bhojwani, V. K., and Phadkule, S., 2015, "Selection of High Performing Geometry in Flexure Bearings for Linear Compressor Applications Using FEA," Int. J. Sci. Technol. Res., **4**(1), pp. 170–173.

[33] Awtar, S., 2004, "Synthesis and Analysis of Parallel Kinematic XY Flexure Mechanisms," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA.

[34] Slocum, A., 2010, "Kinematic Couplings: A Review of Design Principles and Applications," Int. J. Mach. Tools Manuf., **50**(4), pp. 310–327.

[35] Ryan Vallance, R., Morgan, C., and Slocum, A. H., 2004, "Precisely Positioning Pallets in Multi-Station Assembly Systems," Precis. Eng., **28**(2), pp. 218–231.