Development of a Monte Carlo multiple source model for inclusion in a dose calculation auditing tool

Austin M. Faught^{a)}

*Department of Radiation Physics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA
The University of Texas Graduate School of Biomedical Sciences at Houston, Houston, TX 77030, USA
Department of Radiation Oncology, University of Colorado School of Medicine, Denver, CO 80045, USA*

Scott. E. Davidson

Department of Radiation Oncology, The University of Texas Medical Branch, Galveston, TX 77555, USA

Jonas Fontenot

Department of Physics, Mary Bird Perkins Cancer Center, Baton Rouge, LA 70809, USA

Stephen F. Kry

*Department of Radiation Physics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA
The University of Texas Graduate School of Biomedical Sciences at Houston, Houston, TX 77030, USA*

Carol Etzel

*Department of Epidemiology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA
Consortium of Rheumatology Researchers of North America (CORRONA), Inc., Southborough, MA 01772, USA*

Geoffrey S. Ibbott and David S. Followill

*Department of Radiation Physics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA
The University of Texas Graduate School of Biomedical Sciences at Houston, Houston, TX 77030, USA*

(Received 14 February 2017; revised 11 June 2017; accepted for publication 14 June 2017;
published 1 August 2017)

Purpose: The Imaging and Radiation Oncology Core Houston (IROC-H) (formerly the Radiological Physics Center) has reported varying levels of agreement in their anthropomorphic phantom audits. There is reason to believe one source of error in this observed disagreement is the accuracy of the dose calculation algorithms and heterogeneity corrections used. To audit this component of the radiotherapy treatment process, an independent dose calculation tool is needed.

Methods: Monte Carlo multiple source models for Elekta 6 MV and 10 MV therapeutic x-ray beams were commissioned based on measurement of central axis depth dose data for a $10 \times 10 \text{ cm}^2$ field size and dose profiles for a $40 \times 40 \text{ cm}^2$ field size. The models were validated against open field measurements consisting of depth dose data and dose profiles for field sizes ranging from $3 \times 3 \text{ cm}^2$ to $30 \times 30 \text{ cm}^2$. The models were then benchmarked against measurements in IROC-H's anthropomorphic head and neck and lung phantoms.

Results: Validation results showed 97.9% and 96.8% of depth dose data passed a $\pm 2\%$ Van Dyk criterion for 6 MV and 10 MV models respectively. Dose profile comparisons showed an average agreement using a $\pm 2\%/2 \text{ mm}$ criterion of 98.0% and 99.0% for 6 MV and 10 MV models respectively. Phantom plan comparisons were evaluated using $\pm 3\%/2 \text{ mm}$ gamma criterion, and averaged passing rates between Monte Carlo and measurements were 87.4% and 89.9% for 6 MV and 10 MV models respectively.

Conclusions: Accurate multiple source models for Elekta 6 MV and 10 MV x-ray beams have been developed for inclusion in an independent dose calculation tool for use in clinical trial audits. © 2017 American Association of Physicists in Medicine [<https://doi.org/10.1002/mp.12426>]

Key words: dose calculations, Monte Carlo, quality assurance, source model

1. INTRODUCTION

The Imaging and Radiation Oncology Core Houston (IROC-H) QA Center (formerly known as the Radiological Physics Center) is one of six National Cancer Institute (NCI) funded, quality assurance (QA) offices that provides QA core and auditing services to institutions participating in NCI's National Clinical Trial Network (NCTN). IROC-H has developed several programs over the years as a means to efficiently provide dosimetric and QA services to the clinical trial

community and to ensure NCI that the institutions participating in clinical trials deliver comparable and consistent radiation doses. IROC-H's QA programs are comprised of on-site evaluations and remote auditing tools. The on-site evaluations consist of a review of quality control procedures, measurement of basic beam dosimetry data, a review of patient dose calculations, and interviews of personnel that perform physical measurements on the therapy machines. The remote auditing tools are used to assist in the review of patient dose calculations, measure reference beam output with optically

stimulated luminescence dosimeters (OSLD) and thermoluminescent dosimeters (TLD), and evaluate advanced treatment procedures with end-to-end anthropomorphic QA phantoms. The anthropomorphic QA phantoms are designed to test the entire treatment process beginning with imaging of the patient and continuing through treatment planning, setup, and delivery of the prescription dose. This is done by comparing measurements from the phantom's TLD and radiochromic film to the institutions' treatment planning system (TPS) dose calculations.¹

Measurement based comparisons have provided acceptable assurance evaluating an institution's ability to accurately deliver dose for conventional treatment procedures. However, patient dose calculations in the lung or near bony anatomy using newer treatment delivery technologies require the use of heterogeneity correction dose algorithms that IROC-H is currently not able to fully verify with its current QA tools. IROC-H has published results detailing the associated uncertainty with TLD² and radiochromic film³ as well as the results from the remote anthropomorphic phantom audit program that present the ability of institutions to deliver their intended treatment plans.⁴⁻⁹ These publications show a varying degree of agreement between institution reported doses and measured dose in the phantoms. There is reason to believe that these plan variations could be from the beam commissioning process, delivery of the treatment, and the accuracy of the dose calculation algorithms^{3,10-12} used by the TPS. The average institutional pass rate for the H&N phantom is 80% with a relatively large 7%/4 mm acceptance criterion.¹³ Importantly, these failures arise most commonly from a dosimetric mismatch between the calculated dose and that actually delivered – only a small minority of failures have been attributable to setup errors.^{13,14} The observed differences between institutions are cause for concern that variations and inaccuracies in the delivery of radiation therapy between institutions could negatively impact patient safety and compromise the conclusions drawn from NCI supported multi-institutional clinical trials. Currently IROC-H has no means by which to check the dose calculations made by the TPS for IMRT and heterogeneity corrected treatments to identify and correct any errors resulting from a planning system's dose calculation algorithms. In order to evaluate the actual dosimetry and judge the accuracy of the TPS predicted dose distributions, a trusted independent dose calculation tool is needed. A completed calculation tool will be capable of providing a direct comparison of the dose calculations and beam modeling of institutions' TPS as well as serving as a secondary means of comparison in the anthropomorphic phantom auditing program. The latter use will be an additional means of comparison to help isolate potential sources of disagreement between institutional calculated doses and those measured from the anthropomorphic phantoms.

Interest in the Monte Carlo (MC) technique as means of trusted independent dose calculations have been motivated by its superior accuracy compared to traditional, deterministic convolution algorithms.¹⁵ While different codes utilizing the technique have been developed,¹⁶⁻²⁰ all operate under the

similar idea that with known interaction probability distributions of electrons and photons modeled radiation may be transported with a high degree of statistical certainty. For this reason, IROC-H chose to develop a Monte Carlo, measurement based, multiple source model for Elekta 6 MV and 10 MV therapeutic x-ray beams in which output is matched to standard dosimetry data as determined by IROC-H. This calculation method has the advantage of being independent of the complexities within the treatment head.¹⁵ Parameters to the analytical models describing the multiple sources are derived by minimizing the difference between simulated and measured data.²¹⁻²⁵ Differences in accelerator design between manufacturers lead to dosimetric differences that necessitate separate Monte Carlo models for each manufacturer in order to achieve suitable model accuracy. The developed multiple source models for Elekta 6 MV and 10 MV beams will be integrated with previously developed Varian 6 MV and 10 MV models²⁶ and eventually Varian FFF 6 MV and FFF 10 MV models into a dose calculation tool that will be capable of accurately calculating doses for over 90% of all beams monitored by IROC-H.

2. MATERIALS AND METHODS

2.A. Source model

The source model used in the IROC-H analytical model consisted of three sources, primary, extra-focal, and electron contamination, corresponding to bremsstrahlung photons created in the target, scattered photons created within the treatment head, and electron contamination from the treatment head, respectively, as initially described by Davidson *et al.*²⁷ The energy spectra for the primary and extra-focal sources were described by the product of a Fatigue-Life and Fermi function, referred to as a Fatigue-Fermi Distribution (FFD) [Eq. (1)], and the source distributions were implemented without change from Liu *et al.*²⁸ The electron contamination source spectrum and source distribution were modeled as described in Fippel *et al.*²⁹

$$f(E) = \left(\frac{\sqrt{\frac{E-\mu}{\beta}} + \sqrt{\frac{\beta}{E-\mu}}}{2\gamma(E-\mu)} \right) \left(\frac{e^{-\left(\frac{1}{2}\right)\left(\frac{\sqrt{\frac{E-\mu}{\beta}}\sqrt{\frac{\beta}{E-\mu}}}{\gamma}\right)^2}}{\sqrt{2\pi}} \right) \left(\frac{1}{1 + e^{\frac{E-E_F}{kT}}} \right) \quad (1)$$

Such that $E > \mu$; $\gamma, \beta > 0$

In Eq. (1), E is the photon energy, E_F the cut-off energy, and μ , β , and γ shape the photon spectrum. This function was chosen for the model for its ability to fit the photon spectra from numerous linac manufacturers determined from the Monte Carlo code BEAM and studied by Sheikh-Bagheri and Rogers³⁰ without being overly parameterized.

Determination of the primary and extra-focal spectrum parameters was done in an optimization process that compared 10×10 cm² field size percent depth dose (PDD) data

from measurements to calculations based on summing dose matrices from a series of mono-energetic source calculations in 0.25 MeV increments. All calculations were performed with the Dose Planning Method¹⁶ Monte Carlo code. The extra-focal source spectrum was assigned the same parameters as the primary source but linearly scaled in relative fluence and energy by parameters determined during the optimization process.

The second step of the commissioning process was based on measured data from a $40 \times 40 \text{ cm}^2$ field size, and was executed with the intention of modeling off-axis effects such as increased off-axis fluence and decreased mean energy off-axis, collectively contributing to the off-axis Horn Effect. Using the optimized spectra from the first step and an off-axis correction for half-value layer (HVL) formulated by Taylor *et al.*³¹ and implemented without change, calculations were run for $1600 \times 1 \times 1 \text{ cm}^2$ beamlets, making up an open $40 \times 40 \text{ cm}^2$ field. Each beamlet's contribution to the total dose was adjusted based on a piecewise linear function such that the calculated dose profiles were matched to measured dose profiles at a depth of d_{max} .

A projected fluence map at isocenter was formed with dimensions determined by the jaw settings. This map was divided into $0.5 \times 0.5 \text{ mm}^2$ fluence segments. The fluence through each fluence segment was determined from MLC position pulled from the DICOM plan file for each plan segment, the transmission through the MLC leaves, transmission through the rounded leaf ends, and the leakage between adjacent leaves. Each group of fluence segments exposed to the primary source during the plan segment was assigned a fluence based on the number of monitor units per plan segment. The transmission through the MLC assigned additional fluence to the fluence segments as a percentage of the monitor units in an amount that varied along the leaf length with the least amount being at the rounded tips of the leaves. A piecewise linear function was used to weight the fluence in this region to model the rounded tip. The fluence segments that this function was applied to corresponded to the projected width of the leaves at isocenter and an effective tip length of 5 mm. This resulted in a more effective modeling of the penumbra caused by the shape of the leaves. Interleaf leakage was modeled by assigning additional fluence, expressed as a percentage of the monitor units, to the fluence segments. This was done along the fluence segments alongside the leaves in a 1 mm wide region with length corresponding to the interleaf regions. A final, composite fluence map was calculated by summing the $0.5 \times 0.5 \text{ mm}^2$ fluence segments from all beam segments determined in the treatment plan. The MC calculation was then run by segmenting the fluence map into beamlets of similar monitor units. In this way, a pseudoexplicit transport of the particles was performed rather than an explicit transport through the MLC leaves.

2.B. Validation and benchmarking

Validation of the newly designed Elekta source models was performed through a comparison of calculated dosimetry

values to measured beam data. Specifically that data consisted of depth dose data and dose profiles at depths of 1.6 cm (d_{max}), 5.0 cm, 10.0 cm, and 20.0 cm for the 6 MV model and 2.0 cm (d_{max}), 5.0 cm, 10.0 cm, 20.0 cm, and 25.0 cm for the 10 MV model. These comparisons were performed for the following field sizes: $3 \times 3 \text{ cm}^2$, $5 \times 5 \text{ cm}^2$, $10 \times 10 \text{ cm}^2$, $15 \times 15 \text{ cm}^2$, $20 \times 20 \text{ cm}^2$, to $30 \times 30 \text{ cm}^2$. Measurements were performed in a water phantom and calculations were performed in a simulated $50 \times 50 \times 50 \text{ cm}^3$ water phantom. Agreement between measured and the calculated Elekta source model data was evaluated as suggested by Van Dyk *et al.*³² with a modified and more restrictive criteria of $\pm 2\%$ of the maximum dose for depth dose data and flat areas of the dose profiles ($< 3\%/ \text{mm}$ dose fall off) and $\pm 2 \text{ mm}$ distance to agreement for high gradient regions of the dose profiles. For dose profile comparison, the analysis was performed out to an off-axis distance corresponding to 5% of the maximum central axis dose. To mimic the volume averaging effect of the ion chamber used to collect measurement data (Wellhofer CC13, CNMC, Nashville, TN, USA), calculated dose profiles were convolved with a Gaussian function. Validation calculations were performed using 10^6 particles per cm^2 yielding an uncertainty of 0.5% of the dose maximum.

Upon completion of the validation portion, we next undertook benchmarking of the models. The benchmarking of the validated 6 MV and 10 MV source models was designed to be done in a step-by-step process with increasingly difficult treatment planning and computational challenges including heterogeneous media and modulated treatment plans. IROC-H's anthropomorphic head and neck phantom³³ as seen in Fig. 1, was used in creating, delivering, and comparing to

FIG. 1. The anthropomorphic, hollow, plastic shell to IROC-H's head and neck phantom with the polystyrene insert removed and opened. The polystyrene insert in the RPC's head and neck phantom is opened up to reveal a transverse, cross-sectional view. The insert contains a primary PTV, secondary PTV, and critical structure made of solid water that may be distinguished from the insert in a CT scan for treatment planning purposes. Also pictured are the holes to house the thermoluminescent dosimeters (TLD) for point dose comparisons. The orthogonal slits intersecting in the primary PTV are for sagittal films while an axial film may be placed between the two halves to the insert shown in this cross-sectional image. [Color figure can be viewed at wileyonlinelibrary.com]

calculation a highly modulated, nine co-planar beam, 107 segment, IMRT plan to the homogenous phantom. The total MU was 1806 and 1879 for the 6 MV and 10 MV deliveries, respectively. Next, benchmarking was performed using IROC-H's heterogeneous lung phantom (Fig. 2) described by Followill *et al.*³⁴ with a nine co-planar beam, 3D conformal radiotherapy (CRT) plan. As a final test, benchmarking was performed with a 6 co-planar beam, IMRT plan on the lung phantom. The total MU delivered for the IMRT lung plans were 1214 and 1248 for 6 MV and 10 MV deliveries, respectively. All treatment plans were designed to meet the dose prescription guidelines established originally by the RTOG (Table I) and each plan was delivered three times to verify the consistency of the measurements. All deliveries were performed on an Elekta Inifinity with MLCi2 multi-leaf collimator.

Point dose comparisons were made between calculated doses and measured TLD doses of the target structures. The measured doses were determined from the small volume of TLD powder contained in the TLD capsule, and the calculated values were determined from a corresponding contour of the TLD in the CT scan as described previously by Molineu *et al.*³³

Two dimensional dose distributions were evaluated using the gamma index technique.³⁵ Agreement between calculated and measured dose was evaluated in the film planes intersecting the phantom target using a criteria of $\pm 3\%$ of the target TLD dose and ± 2 mm distance to agreement. Exclusion of selected regions of the film were performed in areas where the film was altered to allow for proper assembly of the dosimetry tools within the phantom. These regions included cutout regions along the central longitudinal axis and central lateral axis of the sagittal and coronal films in the lung phantom that allow for positioning the film along a single central axis, a small cutout in the right, postero-lateral corner of the

TABLE I. The treatment planning guidelines for IROC-H's anthropomorphic phantoms are presented for the IMRT Head and Neck Phantom (left) and the Lung Phantom (right).

Head and neck phantom guidelines	Lung phantom guidelines
6.6 Gy delivered to 95% of the primary PTV	6.0 Gy delivered to 95% of the PTV
5.4 Gy delivered to 95% of the secondary PTV	Less than 4.1 Gy delivered to the spinal cord
Less than 1% of the PTVs may receive less than 93% of the Rx dose	Less than 1.8 Gy delivered to 40% of the lung
OAR is to receive a dose less than 4.5 Gy	Less than 3.6 Gy delivered to the entire heart
Normal tissue dose must be less than 110% of Rx dose	Less than 4.5 Gy delivered to 50% of the heart

axial film for the head and neck phantom to allow for proper film placement in the dosimetry insert, and a cutout in the sagittal film of the head and neck phantom to allow for the placement of TLD in the critical structure region.

The Monte Carlo dose calculation resolution matched that of the CT voxel size in the scans of the anthropomorphic phantoms which was $0.518 \times 0.518 \times 1.25 \text{ mm}^3$ in the head and neck phantom and $1.27 \times 1.27 \times 1.25 \text{ mm}^3$ in the lung phantom. The estimated single voxel standard error of the mean from the Monte Carlo dose calculations was 1.1% using 12 million particles per square centimeter.

3. RESULTS

The seven optimized parameters, determined in the Elekta source model commissioning, that describe the photon energy spectra, fluence contributions, and volume averaging of the ion chamber leading to an exaggerated penumbra for dose profiles are reported in Table II for the Elekta 6 MV and 10 MV models. The first three parameters, γ , μ , and β describe the spectrum shape, peak energy location, and relative scale, respectively. The fourth and fifth parameters in the table scale, with respect to the primary source, the spectrum of the extra-focal source and relative fluence of the extra-focal source, respectively. The sixth parameter is the relative contribution of electron contamination for each model, and the final parameter is the standard deviation of the Gaussian function used to mimic partial volume averaging of the ion chamber used to collect measurement data.

An empirically determined field size dependent output correction was implemented in addition to the normal output factor due to the models inability to fully model field size dependent scatter conditions. Each model, Elekta 6 MV [Eq. (2)] and 10 MV [Eq. (3)], had its own correction implemented.

$$y = 1.15 - \frac{2.35 \text{ cm}}{6.12 \text{ cm} + x} \quad (2)$$

$$y = 1.13 - \frac{2.07 \text{ cm}}{6.23 \text{ cm} + x} \quad (3)$$

FIG. 2. The outer shell of IROC-H's thorax phantom with the lung insert removed. Within the insert are slits for the placement of radiochromic film and holes for TLD. Rods containing TLD capsules are also inserted into the shell for point dose measurements corresponding to the location of the spinal cord and heart, the representative critical structures for this treatment. [Color figure can be viewed at wileyonlinelibrary.com]

TABLE II. Optimized parameters for the source models as determined during the initial commissioning of the models. The first three parameters describe the shape and location of the spectra. The fourth through sixth parameters relate the relative contribution and energy scale of the three sources, and the final parameter is the standard deviation of the Gaussian kernel convolved with the calculated dose profiles to mimic the volume averaging effect of an ion chamber at the penumbra of dose profiles.

Parameter	Value Elekta 6 MV	Value Elekta 10 MV
Fatigue-life distribution shape parameter, γ	1.79	1.15
Fatigue-life distribution location parameter, μ	-0.0163	-0.0165
Fatigue-life distribution scale parameter, β	3.69	3.95
Primary spectrum to extra-focal spectrum reduction factor	1.71	3.10
Extra-focal fluence relative to the primary fluence	0.1101	0.1901
Electron contamination contribution (relative to the primary photon contribution)	0.002	0.005
Standard deviation of Gaussian used to convolve the MC dose profile to match the measured dose profile during validation (in mm)	1.8	1.8

In the above equations, y is a unit-less output correction and x is the length of the equivalent square field size.

A comparison between measured and calculated depth dose data for the 6 MV and 10 MV models was performed for field sizes ranging from $3 \times 3 \text{ cm}^2$ to $30 \times 30 \text{ cm}^2$. Van Dyk criteria³² agreement is summarized in Table III. The average percentage of points passing the established criterion was 97.9% and 96.8% for 6 MV and 10 MV models respectively.

A comparison of calculated dose profiles to measured profiles was performed for both Elekta 6 MV and 10 MV models for field sizes ranging from $3 \times 3 \text{ cm}^2$ to $30 \times 30 \text{ cm}^2$ at depths of d_{max} , 5 cm, 10 cm, and 20 cm. An additional comparison at a depth of 25 cm was performed for the Elekta 10 MV model. Agreement between measured and calculated results are presented in Table III. The average percentage of points passing the established criteria were 98.0% and 99.0% for 6 MV and 10 MV models, respectively, showing excellent agreement. Representative examples of depth dose comparisons and dose profile comparisons for 6 MV and 10 MV models are presented in Figs. 3 and 4, respectively.

Results comparing TLD measurements in the anthropomorphic head and neck phantom to calculated doses by the multiple source 6 MV and 10 MV models are shown in Table IV. TLD capsules were located in the primary PTV (four capsules) and the center of the secondary PTV (two capsules). The averaged ratio of calculated dose to measured dose for the primary PTV and secondary PTV are reported. The average ratio of calculated dose to measurement was within 2% for both 6 MV and 10 MV models.

The point dose comparisons between TLD measurement in the heterogeneous lung phantom and the model calculations are shown in Table V for the 3D CRT and IMRT plans. The phantom contained two TLD capsules in the PTV. The

TABLE III. Van Dyk criteria³² agreement for Elekta 6 MV and 10 MV models for measured and calculated depth dose data and dose profile data. Dose profiles were measured and calculated in-plane and cross-plane at the various depths.

Field size (cm^2)	% Points passing			
	Elekta 6 MV model		Elekta 10 MV model	
	Depth dose	Dose profiles	Depth dose	Dose profiles
3×3	98.9	97.7	96.0	99.3
5×5	98.9	97.7	96.0	99.4
10×10	98.3	97.8	96.0	99.5
15×15	97.2	97.1	97.7	98.4
20×20	97.2	98.5	97.7	98.0
30×30	97.2	99.1	97.2	99.4

reported agreement is expressed as a ratio of calculated dose to measured dose. Agreement, expressed as a ratio of calculated dose to TLD measurement, in the PTV were all within 3% and averaged 1.7% and 1.8% for both 6 MV and 10 MV models, respectively, for the 3D CRT and IMT plans.

Agreement between model calculations and film measurements were also assessed using the gamma index technique³⁵ with a $\pm 3\%/2 \text{ mm}$ criterion. The average percentage of pixels passing the criterion for the head and neck phantom in the axial and sagittal film planes is reported in Table VI. The agreement in the sagittal planes averaged 87.3% and 87.2% for 6 MV and 10 MV models, respectively, and in the axial plane showed an average agreement of 87.2% and 93.8% respectively. An example of the gamma comparison conducted in the axial plane for the 10 MV model is shown in Fig. 5.

The multiple source model calculations were also compared to film measurements in the axial, coronal, and sagittal planes of the heterogeneous lung phantom using a $\pm 3\%/2 \text{ mm}$ gamma criterion for the 3D CRT plan. The average percent of data passing in each plane for both models are reported in Table VI. The 6 MV model showed 86.8%, 87.8%, and 86.9% of data passing the criterion for axial, coronal, and sagittal planes respectively. The 10 MV model showed better agreement with 90.2%, 89.5%, and 88.3% of data passing in the axial, coronal, and sagittal planes.

The same $\pm 3\%/2 \text{ mm}$ gamma criterion was used in the comparison of multiple source model dose calculations to measurements for the IMRT lung delivery. Results from the film comparison to calculation are reported in Table VI. Agreement for the 6 MV model, expressed as a percentage of pixels passing the criterion, was 85.2%, 90.0%, and 88.6% for axial, coronal, and sagittal planes, respectively. The 10 MV model exhibited passing rates of 91.2%, 90.6%, and 88.0% in axial, coronal, and sagittal planes respectively.

4. DISCUSSION

With the overall average agreement ratio being 1.012 for all energy and treatment conditions and no agreement greater

FIG. 3. Depth dose comparisons for 6 MV (left) and 10 MV (right) are presented for a $10 \times 10 \text{ cm}^2$ field size. The measured (line) and calculated dose ('*') were compared using a $\pm 2\%$ criterion. [Color figure can be viewed at wileyonlinelibrary.com]

FIG. 4. A comparison of calculated ('◇') and measured (line) dose profiles are presented for a $30 \times 30 \text{ cm}^2$ field at depths of d_{max} , 5 cm, 10 cm, 20 cm, and 25 cm (10 MV only). Calculated doses are compared to measured profiles in both the in-plane and cross-plane directions. [Color figure can be viewed at wileyonlinelibrary.com]

than $\pm 3\%$ for all TLD to dose calculation comparisons, the multiple source model exhibited excellent accuracy. In the head and neck phantom, the OAR is in close proximity to

the primary PTV. The steep dose gradient needed to meet IROC-H's planning criteria makes the TLD measurement more sensitive to positional changes from the phantom setup

TABLE IV. Point dose comparison between the multiple source models and measurement for an IMRT plan delivered to an anthropomorphic head and neck phantom.

Point dose location	Ratio of calculated/measured doses	
	6 MV model	10 MV model
Primary PTV	1.014	1.020
Secondary PTV	1.000	0.992

TABLE V. The point dose comparison between the multiple source model and TLD measurement for a 3D CRT plan and an IMRT plan delivered to a heterogeneous lung phantom.

Point dose location	Ratio calculated/measured			
	3D CRT plan		IMRT plan	
	6 MV model	10 MV model	6 MV model	10 MV model
PTV	1.030	1.016	1.004	1.021

TABLE VI. Comparison between film measurement in axial, coronal, and sagittal planes and Monte Carlo dose calculations for an IMRT head and neck plan delivered to the IROC-H head and neck phantom and a 3D conformal lung plan and a IMRT lung plan delivered to the heterogeneous lung phantom is reported. Agreement was assessed using a $\pm 3\%/2$ mm criterion and is reported as a percentage of data passing.

Model	Delivery technique	Averaged percent of pixels passing		
		Axial film plane	Coronal film plane	Sagittal film plane
Elekta 6 MV	IMRT H&N	87.3	N/A	87.2
	3D CRT lung	86.8	87.8	86.9
	IMRT lung	85.2	90.0	88.6
Elekta 10 MV	IMRT H&N	87.2	N/A	93.8
	3D CRT lung	90.2	89.5	88.3
	IMRT Lung	91.2	90.6	88.0

or the secure placement of the insert within the phantom itself. While the dose gradients near the OARs in the lung phantom are lower than the head and neck phantom, the lower dose to the TLD (1–2 Gy compared to 6.6 Gy) affects the accuracy of the reading process based on IROC-H reading practices for phantom TLD. It is also worth noting that IROC Houston does not place a pass/fail criterion on agreement with the OAR TLD for exactly this reason – the agreement for these TLD is very noisy. For the same reason our study restricted the comparison of measured TLD dose to calculated dose to PTV based TLD only.

All film planes for both phantom trials exhibited passing rates over 85% using a $\pm 3\%/2$ mm gamma criterion and an overall average passing rate of 88.3%. For IROC-H phantom audits the minimum passing rate is established as 85%, and so this performance was deemed acceptable. This is particularly reasonable because the gamma criterion used, $\pm 3\%/2$ mm, was more restrictive than the normal criterion used in phantom audits ($\pm 7\%/4$ mm for the head and neck phantom and $\pm 5\%/5$ mm for the lung phantom). This criterion was chosen to ensure suitable accuracy of the dose calculation tool while still accounting for measurement uncertainty associated with phantom setup, film registration, and dosimeter reading.

The accuracy of the source models combined with their development entirely from measurement data makes them well suited to serve as an independent quality assurance audit tool for clinically used dose calculation algorithms. This is a valuable tool in the peer review process of institutions joining NCTN studies as IROC-H studies have suggested that dose calculation algorithms are a source of variability among institution submitted results.^{10,11} Inclusion of the commissioned source models for Elekta 6 MV and 10 MV beams is especially important in order for the dose calculation tool to be useful across multiple manufacturers monitored by IROC-H. The benchmarking tests demonstrated the accuracy of the models in clinically realistic setups while also serving as a proof of concept in how the model might be used to recalculate dose distributions from IROC-H phantom audits to serve

FIG. 5. A gamma comparison (left) conducted in the axial plane of the IMRT H&N phantom is presented for the 10 MV beam model. A left to right dose profile comparison (right) corresponding to the white line in the gamma map is also shown. The Monte Carlo calculated dose is shown and compared to the measured dose from the film. The axial plane of comparison corresponds to the opening in the dosimetric insert of the phantom shown in Fig. 2. [Color figure can be viewed at wileyonlinelibrary.com]

as an additional means of comparison to isolate and correct any errors discovered during the audit. Finally the tool can be used as a direct means of comparison to the institution TPS dose calculation algorithms by comparing dose calculations in a common benchmark case.

Further studies to expand the models' use to newer, and increasingly popular, treatment techniques is warranted. Of particular interest would be refinement of the tool to handle Volumetric Arc Therapy (VMAT). The current study focused on static IMRT plans in an effort to focus on beam model accuracy. With VMAT increasingly being a treatment technique of choice for modulated plans, its inclusion in a dose calculation tool is an important next step. Additional work is under way to incorporate flattening filter free (FFF) models into the calculation tool. Due to differences in the spectra and off-axis effects, these additions will require entirely separate source models to be developed. Additional studies to benchmark the model against delivery with the Agility MLC are also warranted. Inclusion of VMAT capabilities and FFF source models would allow for the completed calculation tool to be used on the majority of treatment plans submitted to IROC-H for credentialing purposes.

5. CONCLUSION

Multiple source models of Elekta 6 MV and 10 MV therapeutic linear accelerators were developed in a two-step process consisting of spectrum optimization and off-axis modeling. The models were validated against open field dosimetry measurements in a water tank for field sizes ranging from $3 \times 3 \text{ cm}^2$ to $30 \times 30 \text{ cm}^2$. Using a refined criteria of acceptability ($\pm 2\%$ for central axis data, $\pm 2\%$ for low gradient regions of profiles, and $\pm 2 \text{ mm}$ for high gradient regions of profiles), depth dose comparison yielded an average passing rate of 97.9% and 96.8% for 6 MV and 10 MV models respectively. The average passing rate of the dose profile comparisons were 98.0% and 99.0% for 6 MV and 10 MV models, respectively. The models were then benchmarked against IROC-H's anthropomorphic phantoms in increasingly challenging tests consisting of a highly modulated IMRT plan delivered to the homogenous head and neck phantom, a 3D CRT plan delivered to the heterogeneous lung phantom, and an IMRT plan delivered to the heterogeneous lung phantom. The average percentage of pixels passing a $\pm 3\%/2 \text{ mm}$ gamma criterion were 87.4% and 89.9% for 6 MV and 10 MV models respectively.

Based on the agreement between measurement and calculated data, the developed multiple source models will serve as an excellent tool to monitor the dose calculation algorithms and heterogeneity corrections used by institutions audited by IROC-H.

ACKNOWLEDGMENTS

This work was supported by Public Health Services Grant Nos. CA 10953, CA 180803, and CA 81647.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to report.

^{a)} Author to whom correspondence should be addressed. Electronic mail: austin.faught@ucdenver.edu; Telephone: +1 (720) 848 5374.

REFERENCES

1. Ibbott G. MO-A-BRC-01: credentialing for clinical trials: the role of the RPC. Paper Presented at AAPM Annual Meeting, July 26–30 2009, Anaheim, CA; 2009.
2. Kirby TH, Hanson WF, Johnston DA. Uncertainty analysis of absorbed dose calculations from thermoluminescence dosimeters. *Med Phys.* 1992;19:1427–1433.
3. Davidson SE, Popple RA, Ibbott GS, et al. Technical note: heterogeneity dose calculation accuracy in IMRT: study of five commercial treatment planning systems using an anthropomorphic thorax phantom. *Med Phys.* 2008;35:5434–5439.
4. Alvarez P, Molineu A, Hernandez N, et al. SU-FF-T-403: evaluation of doses delivered by SBRT to the lung of an anthropomorphic thorax phantom. *Med Phys.* 2005;32:2043–2044.
5. Alvarez P, Molineu A, Hernandez N, et al. TU-E-224A-01: evaluation of heterogeneity corrections algorithms through the irradiation of a lung phantom. *Med Phys.* 2006;33:2214.
6. Alvarez P, Molineu A, Hernandez N, et al. MO-D-AUD-04: a comparison of heterogeneity correction algorithms within a lung PTV. *Med Phys.* 2007;34:2521.
7. Followill D, Molineu A, Alvarez P, et al. TH-D-224C-02: the state of radiotherapy physics through the eyes of a quality auditor. *Med Phys.* 2006;33:2283.
8. Molineu A, Hernandez N, Alvarez P, et al. SU-FF-T-148: IMRT head and neck phantom irradiations: correlation of results with institution size. *Med Phys.* 2005;32:1983–1984.
9. Kry SF, Alvarez P, Molineu A, et al. Algorithms used in heterogeneous dose calculations show systematic differences as measured with the radiological physics center's anthropomorphic thorax phantom used for RTOG credentialing. *Int J Radiat Oncol Biol Phys.* 2013;85:e95–e100.
10. Davidson SE, Ibbott GS, Prado KL, et al. Accuracy of two heterogeneity dose calculation algorithms for IMRT in treatment plans designed using an anthropomorphic thorax phantom. *Med Phys.* 2007;34:1850–1857.
11. Kry SF, Alvarez P, Molineu A, et al. Algorithms used in heterogeneous dose calculations show systematic error as measured with the Radiological Physics Center's anthropomorphic thorax phantom used for RTOG credentialing. Paper Presented at: ASTRO Annual Meeting, Boston MA; 2012.
12. Kerns JR, Followill D, Lowenstein J, et al. Agreement between institutional measurements and treatment planning system calculations for basic dosimetric parameters as measured by IROC-Houston. *Int J Radiat Oncol Biol Phys.* 2016;95:1527–1534.
13. Molineu A, Hernandez N, Nguyen T, et al. Credentialing results from IMRT irradiations of an anthropomorphic head and neck phantom. *Med Phys.* 2013;40:022101.
14. Kry SF, Molineu A, Kerns JR, et al. Institutional patient-specific IMRT QA Does not predict unacceptable plan delivery. *Int J Radiat Oncol Biol Phys.* 2014;90:1195–1201.
15. Chetty IJ, Curran B, Cygler JE, et al. Report of the AAPM Task Group No. 105: issues associated with clinical implementation of Monte Carlo-based photon and electron external beam treatment planning. *Med Phys.* 2007;34:4818–4853.
16. Sempau J, Wilderman SJ, Bielajew AF. DPM, a fast, accurate Monte Carlo code optimized for photon and electron radiotherapy treatment planning dose calculations. *Phys Med Biol.* 2000;45:2263–2291.
17. Salvat F, Fernandez-Varea JM, Baro J, et al. *PENELOPE, an algorithm and computer code for Monte Carlo simulation of electron-photon showers*. Barcelona, Spain: Centro de Investigaciones Energeticas, Medioambientales y Tecnologicas; 1996, 799.

18. Nelson WR, Hirayama H, Rodgers DWO. *The EGS4 Code System*. Menlo Park, CA: Stanford Linear Accelerator Center;1985, SLAC-265.
19. Briesmeister JF. *MCNP – a General Monte Carlo N-Particle Transport Code*. Los Alamos, NM: Los Alamos National Laboratory; 1993, LA-12625.
20. Baro J, Sempau J, Fernandez-Varea JM, et al. PENELOPE: an algorithm for Monte Carlo simulation of the penetration and energy loss of electrons and positrons in matter. *Nucl Instrum Methods*. 1995;100:31–46.
21. Aljarrah K, Sharp GC, Neicu T, et al. Determination of the Initial Beam Parameters in Monte Carlo Linac Simulation. *Med Phys.*, 2006;33: 850–858.
22. Deng J, Jiang SB, Pawlicki T, et al. Derivation of electron and photon energy spectra from electron beam central axis depth dose curves. *Phys Med Biol*. 2001;46:1429.
23. Faddegon BA, Blevins I. Electron spectra derived from depth dose distributions. *Med Phys*. 2000;27:514–526.
24. Siljämäki S, Tillikainen L, Helminen H, et al. TU-FF-A1-04: determining parameters for a multiple? Source model of a linear accelerator using optimization techniques. *Med Phys*. 2005;32:2113–2114.
25. Yang J, Li JS, Qin L, et al. Modelling of electron contamination in clinical photon beams for Monte Carlo dose calculation. *Phys Med Biol*. 2004;49:2657.
26. Davidson SE, Cui J, Kry S, et al. Modification and validation of an analytical source model for external beam radiotherapy Monte Carlo dose calculations. *Med Phys*. 2016;43:4842–4853.
27. Davidson S, Cui J, Followill D, et al. A flexible Monte Carlo tool for patient or phantom specific calculations: comparison with preliminary validation measurements. *J Phys*. 2008;102:012004.
28. Liu HH, Mackie TR, McCullough EC. A dual source photon beam model used in convolution/superposition dose calculations for clinical megavoltage x-ray beams. *Med Phys*. 1997;24:1960–1974.
29. Fippel M, Haryanto F, Dohm O, et al. A virtual photon energy fluence model for Monte Carlo dose calculation. *Med Phys*. 2003;30:301–311.
30. Sheikh-Bagheri D, Rogers DWO. Monte Carlo calculation of nine megavoltage photon beam spectra using the BEAM code. *Med Phys*. 2002;29:391–402.
31. Tailor RC, Tello VM, Schroy CB, et al. A generic off-axis energy correction for linac photon beam dosimetry. *Med Phys*. 1998;25:662–667.
32. Van Dyk J, Barnett RB, Cygler JE, et al. Commissioning and quality assurance of treatment planning computers. *Int J Radiat Oncol Biol Phys*. 1993;26:261–273.
33. Molineu A, Followill DS, Balter PA, et al. Design and implementation of an anthropomorphic quality assurance phantom for intensity-modulated radiation therapy for the Radiation Therapy Oncology Group. *J Radiat Oncol Biol Phys*. 2005;63:577–583.
34. Followill DS, Evans DR, Cherry C, et al. Design, development, and implementation of the Radiological Physics Center's pelvis and thorax anthropomorphic quality assurance phantoms. *Med Phys*. 2007; 34:2070–2076.
35. Low DA, Moran JM, Dempsey JF, et al. Dosimetry tools and techniques for IMRT. *Med Phys*. 2011;38:1313–1338.