

# A Clinical Study of MLC-Based IMRT Lung Dose Calculation Accuracy on Plan Evaluation Parameters

Khaled Aljarrah<sup>1,2\*</sup>, Todd Pawlicki<sup>2</sup>, Andrzej Niemierko<sup>3</sup>, Loren K. Mell<sup>2</sup> and Steve B. Jiang<sup>2</sup>

<sup>1</sup>Department of Physics, Jordan University of Science & Technology, Irbid, Jordan,

<sup>2</sup>Department of Radiation Oncology, UC San Diego, La Jolla, CA, USA

<sup>3</sup>Department of Radiation Oncology, Massachusetts General Hospital, Boston, MA, USA

## Abstract

Intensity Modulated Radiation Therapy (IMRT) is widely accepted as an appropriate method to treat tumors at many different anatomic locations including lung. Dose calculation algorithms that have different degrees of accuracy are used to produce clinical IMRT treatment plans. In this study, Monte Carlo (MC) dose calculation was used to evaluate the reliability of plan evaluation parameters compared to a pencil beam (PB) dose calculation for IMRT of the lung. Twenty five lung IMRT cases were randomly selected for analysis. Plan evaluation parameters were calculated using PB and MC methods for the targets and organs at risk (OARs). Comparisons were made using dose-volume histograms, mean dose, and equivalent uniform dose. The following doses-volume histogram points were compared:  $D_{95}$ ,  $D_{95}$  of the GTV and PTV,  $V_{20}$  and  $V_{30}$  for the lungs,  $D_{33}$  for the heart and esophagus and  $D_{max}$  for the spinal cord. Mean dose differences were  $3.6 \pm 2.3\%$  and  $4.3 \pm 2.8\%$  for the GTV and PTV, respectively. The average EUD differences were  $4.1 \pm 2.4\%$  for the GTV and  $5.7 \pm 4.9\%$  for the PTV. Less than 2% differences were observed between the MC and PB algorithms for all OAR plan evaluation parameters. However, minimum and maximum differences for some plan evaluation parameters ranged from about  $\pm 20\%$ . There are appreciable differences in plan evaluation parameters between the PB and MC calculations for the targets. The mean dose and EUD have a weak but statistically significant inverse dependence on the number of fields, total MU, GTV volume and PTV volume for the targets. There can be large case-to-case differences between PB and MC for both the targets and OARs. Accurate MC calculations can remove those remaining systematic errors from treatment plans compared to PB calculations.

**Keywords:** IMRT; Lung; EUD; Dose calculation; Monte carlo; Pencil beam

## Introduction

It is well known that intensity modulated radiation therapy (IMRT) provides improved target coverage while sparing surrounding healthy tissue compared to three dimensional conformal radiotherapy. Delivering IMRT using a multileaf collimator (MLC) introduces some complexities (Oldham and Webb, 1997; Ma et al., 2000) such as leaf leakage and photon back scattered to the monitor chamber from the MLC leaves (Chui et al., 1994; Wang et al., 1996; Oldham and Webb, 1997; Hounsell, 1998; Jiang et al., 2001). Pencil beam algorithms (Boyer and Mok, 1985; Mohan et al., 1986; Ahnesjo et al., 1992) are often employed in IMRT treatment planning systems (TPS). Because pencil beam algorithms do not account for electronic disequilibrium in heterogeneous media (e.g., lung), the dose accuracy may be significantly affected (Wang et al., 1996; Ma et al., 1999; Wang et al., 2002; Ma et al., 2004; Yang et al., 2005).

Monte Carlo (MC) methods have been shown to be the most accurate method for radiotherapy dose calculations in homogeneous and heterogeneous geometries (Rogers et al., 1995; Ma et al., 1999; Siebers et al., 2000; McDermott et al., 2003; Rogers, 2006). The use of MC in radiation therapy has increased in the last two decades due in part to advancements in hardware and MC algorithms. At the same time, several MC codes have been developed for treatment planning (Wang et al., 1998; Ma et al., 1999), and for verification of IMRT dose distributions (Ma et al., 2003; Boudreau et al., 2005). Monte Carlo and pencil beam (PB) IMRT dose calculations have been previously compared for prostate and head and neck patients (Ma et al., 1999; Wang et al., 2002; Boudreau et al., 2005; Yang et al., 2005). It has been shown that pencil beam calculations overestimate the dose by 5-10% compared to MC calculations and measurements in slab lung phantoms (Boyer and Mok, 1985; Metcalfe et al., 1993; Butson et al.,

2000; Charland et al., 2003; Carrasco et al., 2004; Krieger and Sauer, 2005; Paelinck et al., 2005; Blazy et al., 2006). Similar results were found in anthropomorphic thorax phantom studies (Laub et al., 2001; Ma et al., 2003; McDermott et al., 2003; Davidson et al., 2007).

Previous studies were primarily devoted to implementing MC in treatment planning (Wang et al., 1998; Ma et al., 1999; Jelen et al., 2005). To our knowledge only two studies were devoted to evaluating lung IMRT dose calculation accuracy in a clinical setting. Both studies used the MC method to verify the IMRT lung dose calculation. One study used the MSKCC TPS (Wang et al., 2002) and the other study used the Helax PB, Helax-TMS's, and Helax-CC TPSs (Vanderstraeten et al., 2006). As found by Wang et al. (2002) the PTV dose calculated w Zhou SM, Das S, Wang Z, Marks LB ith MC using MCPAT code was underdosed by less than 10% compared to the MSKCC calculations. Vanderstraeten et al. (2006) reported that Helax-pencil beam results were unsatisfying for both the targets and organs at risk (OARs) comparing to MC using the MCDE code, Pinnacle-CS, Helax-TMS's, and Helax-CC. Both of these studies used a limited number of patients and fields in each treatment plan.

In comparison studies, patient dose distributions are usually

**\*Corresponding author:** Khaled Aljarrah, Department of applied Physics, Bio-Medical Physics Laboratory, Jordan University of Science & Technology (JUST), P.O. Box 3030, Irbid 22110, Jordan, Tel: +962-2-7201000 ext: 23512; Fax: +962-2-7201071; E-mail: [kjarrah@just.edu.jo](mailto:kjarrah@just.edu.jo)

**Received** March 30, 2010; **Accepted** May 06, 2010; **Published** May 06, 2010

**Citation:** Aljarrah K, Pawlicki T, Niemierko A, Mell LK, Jiang SB (2010) A Clinical Study of MLC-Based IMRT Lung Dose Calculation Accuracy on Plan Evaluation Parameters. J Cancer Sci Ther 2: 074-081. doi:10.4172/1948-5956.1000027

**Copyright:** © 2010 Aljarrah K, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

compared in terms of the mean target dose and cumulated dose volume histogram (cDVH) using dose volume points,  $D_x$  and  $V_x$ . A problem is that these dose points do not appropriately penalize target underdosing or critical structure overdosing. Furthermore, comparisons based on the mean target dose for relatively small dose nonuniformity or comparisons based on the minimum target dose for large dose heterogeneities are not adequate and are first order approximations as explained by Niemierko, (1997). To circumvent these problems, the Equivalent Uniform Dose (EUD) can be used to perform dose comparison and patient plan evaluation. The EUD is based on the assumption that two dose distributions are equal if they have the same radiobiological effect. EUD also describes the volume of interest in a single value which is more realistic and useful for comparison.

In this work, an expanded study using 25 lung cancer IMRT patients and different number of fields is performed to evaluate the effect of dose distribution accuracy on treatment plan evaluation parameters. Pencil beam and Monte Carlo algorithms are used to compare the target mean dose, dose-volume points, and the EUD for targets and OARs. In order to attribute any dose differences related to internal patient scatter versus linac head modeling, MC ion chamber and pencil beam doses were compared and reported for a homogenous water phantom.

## Material and Methods

Twenty five lung cancer IMRT patients, which had been previously treated at Massachusetts General Hospital (MGH) were randomly selected for this study. The tumor size and locations vary from patient to patient. The specific parameters for each patient are shown in Table 1. The clinical treatment plans were based on the patient CT scanned in supine position under normal free breathing condition. All patient plans used in this study were also used for treatment. Lungs, spinal cord, esophagus, and heart were contoured as OARs by the treating physician. The inverse treatment planning system CORVUS v6.2 (North American Scientific, Inc., Chatsworth, CA, USA) was used for IMRT plan optimization and final dose calculation. The CORVUS inverse treatment planning system, which is widely used throughout the world, utilizes a finite pencil beam algorithm for optimization

and dose calculation. From here onward we refer to the CORVUS TPS finite size pencil beam algorithm as PB. Each patient's plan was optimized using 6 MV photon beam for a Varian 21 EX linac equipped with the Millennium MLC. Five to 12 gantry angles were used with a beamlet size of 1x1 cm<sup>2</sup> for all treatment plans.

Once the plans were approved by the physician, the point dose for each plan was verified in a homogenous solid water phantom. The process was done by optimizing the plan to the water phantom, then measuring the dose to a point in the water phantom using a Farmer ionization chamber (0.6 cc active volume). The measurement point of the ion chamber for each patient was chosen in a homogenous dose region, which roughly corresponds to the target location in the patient plan. The ion chamber fluence perturbation correction factor was considered small because the measurement point is chosen in a homogenous dose region. Also, the combination of multiple-beam treatment plans reduce the ionization correction factor (Boudreau et al., 2005). Plans were accepted for treatment if the point dose was within 4% difference (PB versus measurements). For the purpose of comparison, the dose from PB, MC, and the measurements was normalized to the prescribed dose. Once a plan was approved for treatment, the patient's CT data as well as the water phantom's CT data and leaf sequence files were exported to the MC work station in order to re-compute the dose distribution to both the water phantom and patients.

## Monte carlo calculations

The Monte Carlo calculation was performed in two main steps. The first step was to model the linac treatment head. This process consisted of selecting the beam energy and beam full width half maximum incident on the target. A phase space file was generated below the jaws for 40x40 cm<sup>2</sup> field size, based on the linac modeling and the selected beam parameters (Aljarrah et al., 2006). The EGSnrc/BEAMnrc code system was used to model the Varian 21EX 6 MV photon beam (Rogers et al., 1995). The transport parameters for the BEAMnrc simulation were set as ECUT = 0.7 MeV, PCUT = 0.01 MeV, and the number of Bremsstrahlung splitting is set to 20. The dimensions and materials for the accelerator components were set based on the manufacturer's specifications. The parameters are

Patient number	MU	Number of fields	GTV size (cc)	PTV size (cc)	Tumor location
1	750	7	35.9	290.4	Right upper lobe near the mediastinum
2	585	5	12.0	159.4	Left upper lobe
3	422	5	17.5	73.8	Left lower lobe
4	778	6	224.9	523.9	Left upper lobe
5	565	5	28.9	92.7	Right upper lobe
6	608	5	32.0	102.8	Right upper lobe
7	764	8	101.9	392.2	Left lung near the lingual
8	650	5	85.7	395.8	Right lower lobe
9	1041	8	160.2	351.7	Right middle lobe near the mediastinum
10	480	5	87.2	289.7	Right upper lobe near the mediastinum
11	628	5	43.9	131.4	Right middle lobe near the mediastinum
12	429	5	52.6	224.8	Right lower lobe near the mediastinum
13	578	6	371.1	667.5	Right middle lobe near the mediastinum
14	559	5	16.2	96.9	Right middle lobe
15	1070	12	490.4	1189.3	Right upper lobe near the mediastinum
16	897	9	302.5	1049.3	Left upper lobe
17	743	5	114.2	114.2	Right upper lobe
18	581	5	107.1	428.7	Right upper lobe
19	594	5	98.0	261.8	Right upper lobe
20	653	5	84.2	215.8	Right lower lobe
21	662	7	181.2	181.2	Right middle lobe
22	1239	7	166.4	459.9	Lungs near the mediastinum and esophagus
23	804	5	32.9	200.1	Left upper lobe
24	582	8	119.4	416.2	Left lung near the lingual
25	894	9	290.7	756.6	Right lower lobe near the mediastinum

**Table 1:** Description of the 25 IMRT lung cases used in this study.

selected when there is matching between measurement and MC within 2% for the percentage depth dose (PDD) and profiles at depth of 1.5 cm, 5 cm, 10 cm, and 30 cm for field 4×4, 10×10, and 40×40 cm<sup>2</sup>.

In the second step, the generated phase space file was used as a source for the photon beam to calculate dose distribution in the patients and the water phantom as well. The MCSIM code, which is an EGS4/PRESTA user code, was used for dose calculation (Ma, 2004). The simulations were carried out using the following transport parameters: ECUT = 0.7 MeV, PCUT = 0.01 MeV, and ESTEPE = 0.04. The average leaf leakage was used as 1.8% based on ion chamber measurements and the transition to the area under the jaws was considered zero. In all patient calculations, we have kept the statistical uncertainty to be 2% or less so as not to significantly affect isodose lines, DVHs, or biological indices (Keall et al., 2000).

The absolute dose was calculated by converting the MC calculated dose per fluence to dose per MU at the linac calibration conditions in water (depth of 5 cm, 10×10 cm<sup>2</sup> field size, 100 cm SSD, and 100 cGy for 100 MU).

### Plan comparison and dose reporting

The dose to a point in the solid water phantom was calculated using MC and PB for each patient's plan and measured using the ion chamber. Both TPS-PB dose and chamber measurements were compared to MC predicted dose.

Patient IMRT dose distributions and plan evaluation parameters calculated with the MC method were compared to the PB calculation as the PB result minus the MC result expressed as a percentage. Dose-volume histograms (DVHs) were created for each patient, using MC simulations and the PB inverse planning system, for the following targets and OARs: GTV, PTV, ipsilateral and contralateral lungs, spinal cord, esophagus, and heart. The mean dose to each target and OARs was compared. Due to their clinical utility to predict OAR toxicity, dose volume points were used for comparison. For lung, volume dose points  $V_{20}$ , and  $V_{30}$  (the lung volume that receives at least 20 and 30 Gy, respectively) were used. For esophagus,  $D_{33}$ , heart  $D_{33}$ , (dose received by 33% of the volume), and spinal cord maximum dose ( $D_{max}$ ) were used. The spinal cord maximum dose was defined as spinal cord  $D_2$  (dose received by 2% of the spinal cord volume). To evaluate the difference of the stability of dose-volume points that denote target coverage, the dose volume points  $D_{95}$  and  $D_{98}$  (dose received by 95 and 98% of the volume, respectively) were compared.

The comparison based on the mean dose difference is unbiased when the dose is uniformly distributed across the region of interest. This, however, is not the situation for most IMRT dose distributions. The EUD is a promising function for IMRT dose distribution plan comparisons. EUD values are controlled by the parameter  $a$ , where EUD approaches maximum dose as ( $a \gg 1$ ), minimum dose as ( $a \ll 1$ ), and equal to the mean dose as  $a = 1$ . Previous studies show that EUD function can be effectively used to represent the characteristics of treatment plans seen with dose-volume constraints and therefore may be effective substitutes in IMRT treatment planning and plan evaluation. A strong correlation between dose-volume constraints and EUD was found in previous studies as well (Niemierko, 1997; Choi and Deasy, 2002; Wu et al., 2003; Zhou et al., 2004).

The equivalent uniform dose (Niemierko, 1997) was calculated from the MC and the TPS differential dose-volume histograms for each structure and compared. The following formula was used to calculate the EUD values:

$$EUD_a(d) = \left( \frac{1}{n} \sum_{i=1}^n d_i^a \right)^{\frac{1}{a}}$$

where  $d_i$  and  $v_i$  are the dose and the volume of the  $i^{\text{th}}$  voxel, and  $a$  is the tumor or structure control parameter that describes the effect of dose heterogeneity in the structure. Following Niemierko, the control parameters are -15 for tumor, 1 for lung, 12 for spinal cord, 5 for heart, and 8 for esophagus (Niemierko, 1997).

### Statistical analysis

The percent difference between the PB and MC algorithms were compared for all plan evaluation parameters previously mentioned. Pairwise comparisons were performed using the two-tailed paired Student's t-test.  $P$  values less than 5% were considered significant. A linear regression model was used to test correlations between the percent difference in plan evaluation parameters and treatment plan parameters.  $R^2$  values were used to evaluate goodness of fit of the linear models.

## Results

### Water phantom dose comparison

The average dose difference of the measurement point between PB and ion chamber and between MC and ion chamber was found as -0.3% (min: -2.6, max: 2.2) and -0.7% (min: -2.5, max: 1.6) as shown in the histogram of Figure 1. Both the PB and MC algorithms predicted the dose to the measurement point with an average of 0.4% (min: -1.5, max: 1.3).

### Plan evaluation parameter comparison

**Mean dose difference:** On average, the difference (MC vs. PB) in the GTV mean dose for all the plans is 3.6% (min: 0.0, max: 8.4) and 4.3% (min: 0.0, max: 10.5) for the PTV. Based on patient-to-patient comparison, the mean dose difference between PB and MC for the PTV is found to be higher than that for the GTV for most patients even though the difference in their mean averages is about 1%. For all the OARs, the mean dose difference between PB and MC vary within an average of 1%, but the mean dose difference in terms of individual patient's comparison shows a minimum dose difference as low as -2.6% in the case of esophagus and up to 2.8% in the case of the ipsilateral lung.

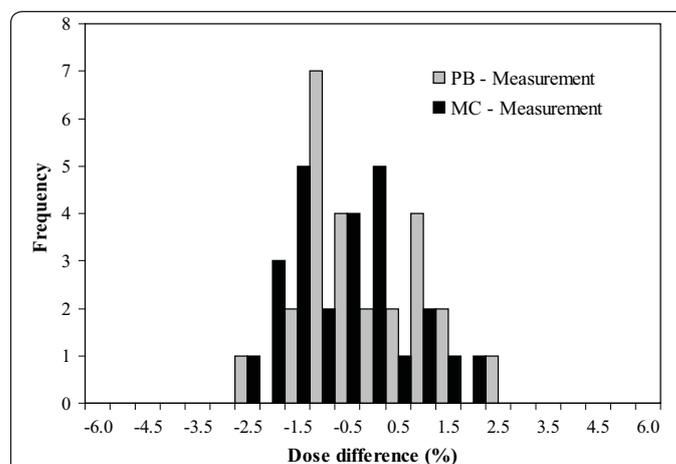


Figure 1: The percentage dose difference between PB, MC and the corresponding ion chamber measurements in a water-equivalent slab phantom.

The mean dose differences between the PB and MC algorithms for the GTV, PTV, and OARs of all 25 patients are shown in Table 2. Statistically different results between the two algorithms were obtained for the GTV, PTV, esophagus and spinal cord mean dose. The greatest difference was for the PTV (4.3%,  $p < 0.01$ ) and the GTV (3.6%,  $p < 0.01$ ). The esophagus and spinal cord differences were both within 1% ( $p = 0.01$ ).

**DVH comparison:** The difference in  $D_{95}$  and  $D_{98}$  values vary from patient to patient with an average of 4.4% (min: 0.3, max: 10.6) and 5.0% (min: 0.3, max: 15.2), respectively. The percent dose difference of  $D_{95}$  and  $D_{98}$  for PTV with average of 7.1% (min: 0.3, max: 19.6) and 7.8% (min: 0.0 max: 22.6), respectively. The average difference of the GTV dose volume points ( $D_{95}$ ,  $D_{98}$ ) is less than that of the PTV. Based on individual patient's comparison, the difference in the  $D_{98}$  varies from the difference in the  $D_{95}$  for some patients. Overall, the variations in both  $V_{20}$  and  $V_{30}$  for the ipsilateral lung are in the range of  $\pm 3\%$ . The mean average of the variations is 0.2% (min: -2.2, max: 2.2) for  $V_{20}$  and 0.6% (min: -1.4, max: 3.2) for  $V_{30}$ . On average, no significant dose difference between PB and MC is found for the percentage dose difference for spinal cord maximum dose ( $D_{max}$ ), heart  $D_{33}$ , and esophagus  $D_{33}$ . However, in the case of the esophagus, individual comparisons show differences for two patients of +6.2% and -5.5%.

The DVH differences between the PB and MC algorithms are shown in Table 3. A statistically significant difference between the two algorithms is found for GTV  $D_{95}$  (4.4%,  $p < 0.01$ ) and  $D_{98}$  (5.3%,  $p < 0.01$ ) as well as the PTV  $D_{95}$  (7.1%,  $p < 0.01$ ) and  $D_{98}$  (7.8%,  $p < 0.01$ ). The ipsilateral lung  $V_{30}$  (0.6%,  $p < 0.01$ ) is the only OAR for which a statistically significant difference between the two algorithms is found.

Structure name	avg $\pm \sigma$ (min, max) %	$p$ (PB vs. MC)
GTV	3.6 $\pm$ 2.3 (0.0, 8.4)	<0.01
PTV	4.3 $\pm$ 2.8 (0.0, 10.5)	<0.01
Esophagus	0.6 $\pm$ 1.0 (-2.6, 2.1)	0.01
Spinal cord	0.2 $\pm$ 0.6 (-1.2, 1.1)	0.04
Heart	0.2 $\pm$ 0.5 (-1.6, 0.8)	0.12

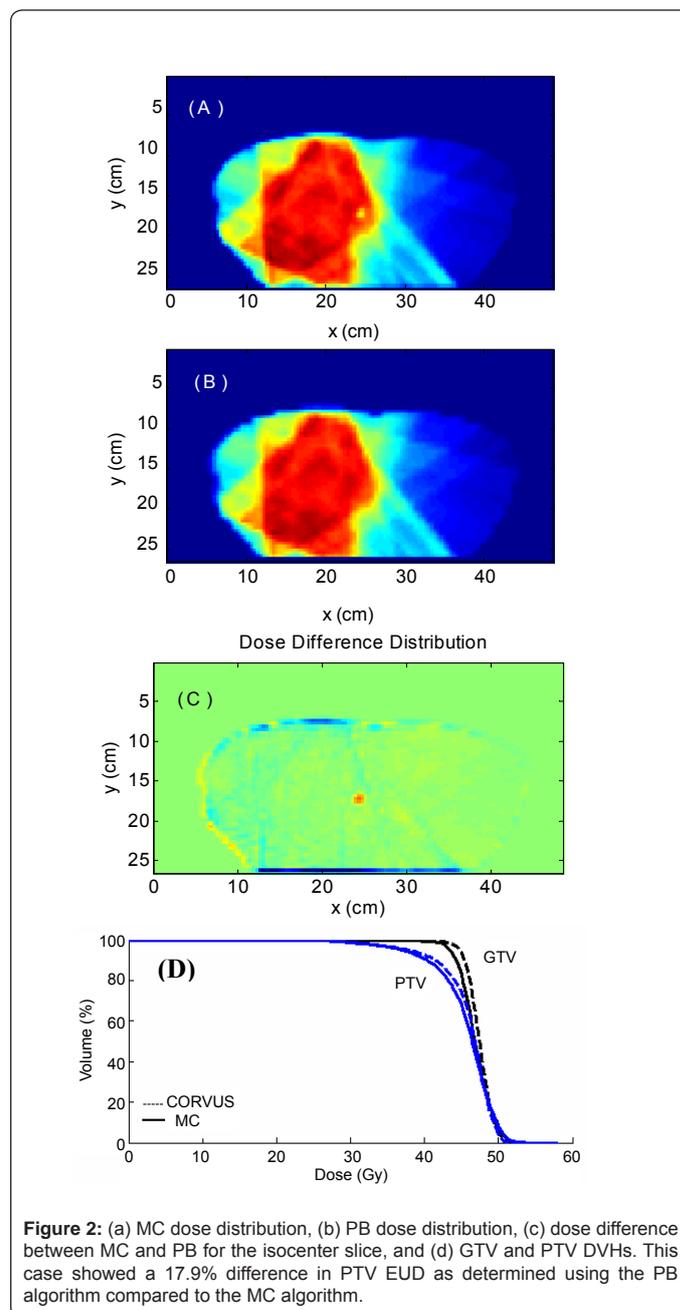
**Table 2:** Mean dose difference comparison between the pencil beam (PB) and Monte Carlo (MC) algorithms (i.e., PB – MC) for the GTV, PTV, and OARs for all 25 patients. Note that the lung mean dose is the same as the lung EUD ( $a = 1$ ). The results for the lungs are reported in Table 4.

Structure name	avg $\pm \sigma$ (min, max) %	$p$ (PB vs. MC)
GTV $D_{95}$	4.4 $\pm$ 2.6 (0.3, 10.6)	<0.01
GTV $D_{98}$	5.3 $\pm$ 4.1 (0.3, 15.2)	<0.01
PTV $D_{95}$	7.1 $\pm$ 5.1 (0.3, 19.6)	<0.01
PTV $D_{98}$	7.8 $\pm$ 5.8 (0.0, 22.6)	<0.01
Ipsilateral Lung $V_{20}$	0.2 $\pm$ 1.1 (-2.2, 2.2)	0.43
Ipsilateral Lung $V_{30}$	0.6 $\pm$ 1.2 (-1.4, 3.2)	<0.01
Spinal Cord $D_{max}$	0.5 $\pm$ 2.0 (-4.1, 4.8)	0.36
Esophagus $D_{33}$	0.5 $\pm$ 2.2 (-5.6, 6.2)	0.24
Heart $D_{33}$	0.1 $\pm$ 0.9 (-2.3, 2.2)	0.24

**Table 3:** Dose-volume comparison between the pencil beam (PB) and Monte Carlo (MC) algorithms (i.e., PB – MC) for the GTV, PTV, and OARs for all 25 patients.

Structure name	avg $\pm \sigma$ (min, max) %	$p$ (PB vs. MC)
GTV	4.1 $\pm$ 2.4 (-0.2, 8.7)	<0.01
PTV	4.8 $\pm$ 6.7 (-17.9, 14.6)	<0.01
Esophagus	1.8 $\pm$ 1.9 (-2.8, 4.9)	<0.01
Spinal cord	0.1 $\pm$ 1.7 (-5.1, 3.6)	0.82
Heart	0.2 $\pm$ 1.0 (-1.9, 2.7)	0.38
Ipsilateral Lung	0.7 $\pm$ 1.0 (-1.4, 2.8)	<0.01
Contralateral Lung	0.6 $\pm$ 0.3 (-0.4, 1.1)	<0.01

**Table 4:** EUD comparison between the pencil beam (PB) and Monte Carlo (MC) algorithms (i.e., PB – MC) for GTV, PTV, and OARs for all 25 patients.



**Figure 2:** (a) MC dose distribution, (b) PB dose distribution, (c) dose difference between MC and PB for the isocenter slice, and (d) GTV and PTV DVHs. This case showed a 17.9% difference in PTV EUD as determined using the PB algorithm compared to the MC algorithm.

**EUD comparison:** For the GTV, the EUD average difference between PB and MC is 4.1% (min: -0.2, max: 8.7). The EUD average difference for the PTV is 4.8% (min: -17.9, max: 14.6). For all patients, the EUD average difference for both the GTV and PTV are close in values, but the EUD difference data for PTV show higher deviation  $\sigma = 6.7\%$  compared to 2.4% for the GTV. Both the GTV and PTV EUD, however, are different depending on whether the PB or MC algorithm is used ( $p < 0.01$ ). For the OARs, the EUD average difference for all the patients for each structure is within 0 to 2%. The esophagus shows relatively higher EUD difference for some patients compared to other OARs such as the spinal cord. The EUD differences of the PTV are most likely greater than zero except one patient which shows about 17.9% reduction in the PB PTV EUD. For this patient, the DVH of the GTV and PTV as well as the PB dose distribution, MC dose distribution, and the dose difference at the isocenter slice are shown

in Figure 2. The EUD and the mean dose difference for OARs in this case and the mean dose difference are within 3%. The calculated dose from PB and MC and the measured dose from the ion chamber in the water phantom are all within a difference of 2%.

The EUD difference for the GTV, PTV, and OARs are shown in Table 4. Statistically significant differences between the PB and MC results are seen for the GTV (4.1%,  $p < 0.01$ ), PTV (4.8%,  $p < 0.01$ ) and OARs of esophagus (1.8%,  $p < 0.01$ ), ipsilateral lung (0.7%,  $p < 0.01$ ), and contralateral lung (0.6%,  $p < 0.01$ ).

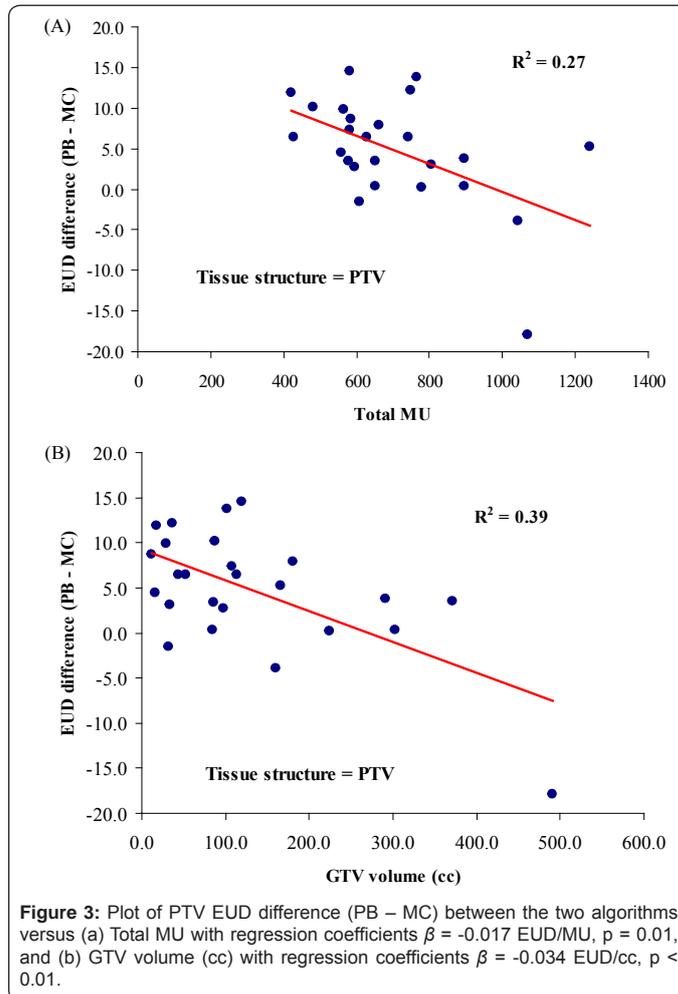
**Parameter correlations**

Only the treatment and plan evaluation parameter correlations with  $p < 0.05$  are presented. Linear regression parameters associated with the difference in the mean dose between PB and MC algorithms are shown in Table 5. Covariates that have predictive ability for the difference in EUD between the PB and MC algorithms for the PTV are shown in Table 6. With linear modeling of the GTV, PTV, and OAR plan parameters' dependence on total MU, number of fields, GTV volume, and PTV volume, the best fits were for the PTV EUD versus GTV volume ( $R^2 = 0.39$ ) and esophagus mean dose versus total MU ( $R^2 = 0.33$ ). Other statistically significant associations under univariate analysis were PTV EUD versus total MU ( $R^2 = 0.27$ ), PTV EUD versus number of field ( $R^2 = 0.21$ ), and PTV versus PTV volume ( $R^2 = 0.31$ ). Plots of the PTV EUD with respect to GTV volume and total MU are shown in Figure 3.

**Discussion**

In this study we have used a large cohort of patients (25) to investigate the effect of dose calculation accuracy on plan evaluation parameters for IMRT lung treatments. Differences have been found in comparing the two algorithms. Several factors are illustrated regarding the variation in the dose difference between PB and MC.

PB-based algorithms calculate the dose-to-water while the MC results reported herein are dose-to-medium, which composed the patient anatomy based on the CT number conversions (tissue, air, bone). As reported by Siebers et al. (2000) for 6 MV beam (Siebers et al., 2000), the dose-to-medium dose-to-water correction is about 1, 10, 13% for tissue, bone, and air, respectively. In this work, we have considered the dose values calculated by MC as the ground truth.



**Figure 3:** Plot of PTV EUD difference (PB – MC) between the two algorithms versus (a) Total MU with regression coefficients  $\beta = -0.017$  EUD/MU,  $p = 0.01$ , and (b) GTV volume (cc) with regression coefficients  $\beta = -0.034$  EUD/cc,  $p < 0.01$ .

Therefore, the dose differences between dose-to-water and dose-to-tissue are taken as real differences between the two algorithms. Others have found that the dose-to-material dose-to-water conversion can affect the final plan parameters from 0 – 8% for head and neck and prostate cases (Dogan et al., 2006). The largest error (8%) in that study was for the femoral heads in prostate cases that included nodal irradiation. In heterogeneous geometries such as lung, the dose-to-medium dose-to-water conversion will not account for the internal scattered dose that affects the dose distribution differences between PB and MC algorithms.

Similarly, another source of dose difference between PB and MC is the perturbation to the fluence of the secondary electrons by the tissue composition. It is well known that the ability of the path-length correction within the PB algorithm to account for this is limited. The level of the perturbation most likely depends on the tumor location and size relative to bone structures or air cavities. For example, a tumor located beyond or in front of a bony structure or an air cavity relative to the incident beam should have a different dose depending on whether PB or MC is used. Such an issue is shown in the two cases of Figure 4a (case number 1) and Figure 4b (case number 7) where both patients have an air cavity inside the PTV. The significant dose prediction differences between PB and MC due to the secondary electron fluence are evident in the isodose curves as shown in the figures (the MC isodose curves being much more constricted in the high-dose region).

Tissue structure	Variable	$\beta$	SE	$p$
GTV	GTV volume (cc)	-0.008	0.004	0.04
GTV	PTV volume (cc)	-0.004	0.002	0.03
PTV	GTV volume (cc)	-0.012	0.004	0.01
PTV	PTV volume (cc)	-0.005	0.002	0.01
PTV	Total MU	-0.006	0.003	0.04
PTV	Number of fields	-0.407	0.308	0.20
Esophagus	Total MU	-0.003	0.001	<0.01
Esophagus	Number of fields	-0.181	0.109	0.11

Abbreviations: SE  $\equiv$  Standard Error;  $\beta$   $\equiv$  change in mean dose difference per unit change of the independent variable.

**Table 5:** Univariate analysis of factors associated with the difference in mean dose between the pencil beam (PB) and Monte Carlo (MC) algorithms for tissue structures GTV, PTV, and esophagus.

Variable	$\beta$	SE	$p$
GTV volume (cc)	-0.03	0.009	<0.01
PTV volume (cc)	-0.013	0.004	<0.01
Total MU	-0.017	0.006	<0.01
Number of fields	-1.656	0.678	0.02

Abbreviations: SE  $\equiv$  standard error;  $\beta$   $\equiv$  change in EUD difference per unit change of the independent variable.

**Table 6:** Univariate analysis of factors associated with the difference in EUD between the pencil beam (PB) and Monte Carlo (MC) algorithms for the PTV.

Variations in the homogeneity of the target tissue composition may also contribute to dose differences between PB and MC algorithms. For example, the GTV usually represents the tumor which has a relatively uniform density of  $\sim 1 \text{ gm/cm}^3$  while the PTV for lung has less tissue homogeneity, usually consisting of a lung tissue shell (density =  $0.23 \text{ gm/cm}^3$ ) surrounding the GTV. The EUD is similarly affected where EUD differences of 13% and 14% are found for case number 1 and 7, respectively. These effects are minimized in homogeneous phantoms as demonstrated by the results in Figure 1 where the agreement between PB, MC, and ion chamber measurements are within 2%.

Previous studies usually compared the performance of a TPS or MC with a TPS based on dose-volume points from the DVH or the mean dose difference. The DVH points usually used are based on their ability to specify target coverage or predict OAR complications. These dose-volume points are unreliable due to their dependence on the DVH slope at the location of the comparison points. Therefore, a clinical decision or comparison based on these points could be misleading. The mean difference evenly penalized positive and negative outliers. Therefore, if there is any cold or hot spot in the target or the structure, it could be canceled out and not appear in the calculated mean. For example, even though a cold spot in MC calculated dose is shown in the patient of Figure 2, the mean dose difference is 1%. This patient has a large tumor GTV of 490cc and a PTV of 1189cc. This case has 12 gantry angles which is the highest among the other 25 plans tested in this study. The PTV for this patient contains tissues from lung, esophagus and bronchi. The mean dose

difference is 1%. The dose calculated by MC shows a cold spot at the PTV which is shown in Figure 2a. The beam number and arrangement and the tissue complication mainly affects the significant reduction in the EUD value.

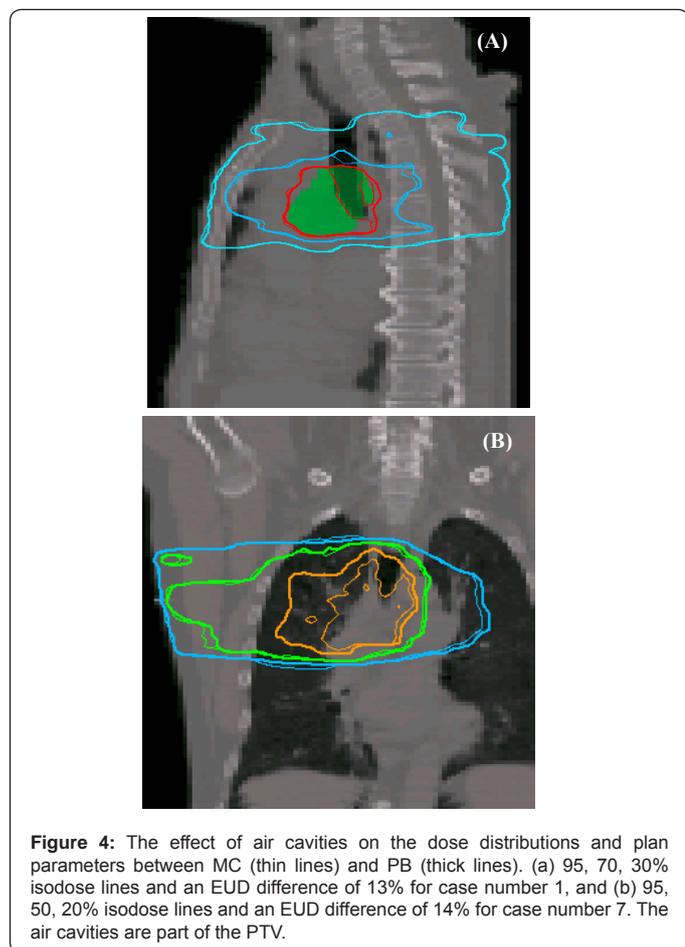
It is a natural assumption that the patient anatomy, tumor size and location, number of fields and field angles, and number of segments may play an important role in the dose distribution differences between PB and MC. As previously discussed, the exact reason for the dosimetric differences on a per case basis can be difficult if not impossible to identify. Even if one could identify the physical reason for the differences, it would not be clinically useful. The statistical differences in plan parameters between the PB and MC algorithms for the cohort of patients studied may be more clinically relevant. To this end, the targets show the greatest difference in plan evaluation parameters for the two algorithms. Even though the OAR plan evaluation parameters have shown a statistically significant difference between the two algorithms, the differences are only within about 2%. These differences may or may not be clinically significant but they nevertheless represent a systematic dose difference that can be addressed by accurate dose calculations.

The EUD shows more variability on a case-by-case basis for the targets than the OARs compared to the other plan parameters. For example, the PTV mean dose differences have a minimum of 0.0% and maximum of 10.5%. Correspondingly, the PTV EUD dose differences have a minimum of -17.9% to a maximum of 14.6%. This variability difference between the EUD and other plan parameters reflects the ability of the EUD to appropriately (i.e., in terms of radiobiological response) account for hot and cold spots in the targets.

The linear regression analysis is helpful to identify which, if any, treatment plan parameters are predictive of potential differences between a PB calculation model and a MC calculation model. One would expect that as the number of beams increases that the differences between the two algorithms would be mitigated. This is what was found in our analysis. There is a statistically significant (although weak) negative correlation with PTV EUD and total MU, number of fields, GTV and PTV volume. As the total plan MU, number of fields, GTV and PTV volume increase, then there is a decrease in the difference between the PTV EUD calculated with the PB algorithm and the MC algorithm. The only OAR that shows a similar dependence was for the esophagus mean dose difference with respect to the total MU. It is unclear why this is the only OAR that shows dependence the treatment planning parameters studied. In any case, this study shows that the degree to which PB and MC differences can be predicted using readily available treatment planning parameters (e.g., MU, GTV volume, etc.) has been shown to be weakly linked at best. Furthermore, the difference between PB and MC on plan evaluation parameters can vary widely from case to case as previously shown in the results of Figure 2. It must, therefore, be recommended that developers should still continue to work to implement MC methods into routine clinical use.

## Conclusion

The dose distributions for 25 IMRT lung patients were evaluated in this study using MC and PB dose calculation methods. Dose calculated using the PB algorithm on average overestimated the GTV, PTV, and OARs plan evaluation parameters compared to MC calculations. The average difference for the PTV is 4.3% while it is 7.8% for the  $D_{95}$  of the PTV. Even though the dose comparison in terms of average difference did not indicate high significant variation



**Figure 4:** The effect of air cavities on the dose distributions and plan parameters between MC (thin lines) and PB (thick lines). (a) 95, 70, 30% isodose lines and an EUD difference of 13% for case number 1, and (b) 95, 50, 20% isodose lines and an EUD difference of 14% for case number 7. The air cavities are part of the PTV.

for targets or OARs, the dose difference based on individual patient comparisons could vary significantly for some patients. There was a statistically significant difference between plan evaluation parameters calculated using the PB algorithm compared to the MC algorithm. The plan evaluation parameters of mean dose and EUD have a weak but statistically significant dependence on the number of fields, total MU, GTV volume and PTV volume for some structures. One main conclusion of this work is that the most accurate dose calculation methods (i.e. Monte Carlo) should still be implemented in the clinical routine. One can not tell *a priori* which cases will show the largest errors in plan evaluation parameters. Based on the statistical analysis, we know there are appreciable differences between the PB and MC calculations for the targets. Accurate MC calculations can remove those remaining systematic errors from treatment plans compared to PB calculations. Lastly, MC algorithms will need to be benchmarked or standardized upon clinical implementation so different levels of accuracy are avoided.

### Acknowledgements

The authors would like to acknowledge Mr. Andrew Kaplan for his help with computer cluster used in this work.

### References

- Ahnesjo A, Saxner M, Trepp A (1992) A pencil beam model for photon dose calculation. *Med Phys* 19: 263-73. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Aljarrah K, Sharp GC, Neicu T, Jiang S (2006) Determination of the initial beam parameters in Monte Carlo linac simulation. *Med Phys* 33: 850-58. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Blazy L, Baltes D, Bordy JM, Cutarella D, Delaunay F, et al. (2006) Comparison of PENELOPE Monte Carlo dose calculations with Fricke dosimeter and ionization chamber measurements in heterogeneous phantoms (18 MeV electron and 12 MV photon beams). *Phys Med Biol* 51: 5951-65. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Boudreau C, Heath E, Seuntjens J, Ballivy O, Parker W (2005) IMRT head and neck treatment planning with a commercially available Monte Carlo based planning system. *Phys Med Biol* 50: 879-90. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Boyer A, Mok E (1985) A photon dose distribution model employing convolution calculations. *Med Phys* 12: 169-77. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Butson MJ, Elferink R, Cheung T, Yu PK, Stokes M, et al. (2000) Verification of lung dose in an anthropomorphic phantom calculated by the collapsed cone convolution method. *Phys Med Biol* 45: N143-9. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Carrasco P, Jorret N, Duch MA, Weber L, Ginjaume M, et al. (2004) Comparison of dose calculation algorithms in phantoms with lung equivalent heterogeneities under conditions of lateral electronic disequilibrium. *Med Phys* 31: 2899-911. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Charland PM, Chetty IJ, Yokoyama S, Fraass BA (2003) Dosimetric comparison of extended dose range film with ionization measurements in water and lung equivalent heterogeneous media exposed to megavoltage photons. *J Appl Clin Med Phys* 4: 25-39. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Choi B, Deasy JO (2002) The generalized equivalent uniform dose function as a basis for intensity-modulated treatment planning. *Phys Med Biol* 47: 3579-89. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Chui CS, LoSasso T, Spirou S (1994) Dose calculation for photon beams with intensity modulation generated by dynamic jaw or multileaf collimations. *Med Phys* 21: 1237-44. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Davidson SE, Ibbott GS, Prado KL, Dong L, Liao Z, et al. (2007) Accuracy of two heterogeneity dose calculation algorithms for IMRT in treatment plans designed using an anthropomorphic thorax phantom. *Med Phys* 34: 1850-7. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Dogan N, Siebers JV, Keall PJ (2006) Clinical comparison of head and neck and prostate IMRT plans using absorbed dose to medium and absorbed dose to water. *Phys Med Biol* 51: 4967-80. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Hounsell AR (1998) Monitor chamber backscatter for intensity modulated radiation therapy using multileaf collimators. *Phys Med Biol* 43: 445-54. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Jelen U, Sohn M, Alber M (2005) A finite size pencil beam for IMRT dose optimization. *Phys Med Biol* 50: 1747-66. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Jiang SB, Boyer AL, Ma CM (2001) Modeling the extrafocal radiation and monitor chamber backscatter for photon beam dose calculation. *Med Phys* 28: 55-66. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Keall PJ, Siebers JV, Jeraj R, Mohan R (2000) The effect of dose calculation uncertainty on the evaluation of radiotherapy plans. *Med Phys* 27: 478-84. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Krieger T, Sauer OA (2005) Monte Carlo- versus pencil-beam-/collapsed-cone-dose calculation in a heterogeneous multi-layer phantom. *Phys Med Biol* 50: 859-68. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Laub WU, Bakai A, Nusslin F (2001) Intensity modulated irradiation of a thorax phantom: comparisons between measurements, Monte Carlo calculations and pencil beam calculations. *Phys Med Biol* 46: 1695-706. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Ma CM (2004) MCSIM: A Monte Carlo dose calculation tool for radiation therapy. *Proc of the 14th International Conference on the Use of Computer in Radiation Therapy (ICCR)*, Eds: BY Yi, SD Ahn, EK Choi and SW Ha (Jeong Publishing, Seoul, Korea). » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Ma CM, Jiang SB, Pawlicki T, Chen Y, Li JS, et al. (2003) A quality assurance phantom for IMRT dose verification. *Phys Med Biol* 48: 561-72. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Ma CM, Mok E, Kapur A, Pawlicki T, Findley D, et al. (1999) Clinical implementation of a Monte Carlo treatment planning system. *Med Phys* 26: 2133-43. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Ma CM, Pawlicki T, Jiang SB, Li JS, Deng J, et al. (2000) Monte Carlo verification of IMRT dose distributions from a commercial treatment planning optimization system. *Phys Med Biol* 45: 2483-95. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Ma CM, Price RA Jr, Li JS, Chen L, Wang L, et al. (2004) Monitor unit calculation for Monte Carlo treatment planning. *Phys Med Biol* 49: 1671-87. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- McDermott PN, He T, DeYoung A (2003) Dose calculation accuracy of lung planning with a commercial IMRT treatment planning system. *J Appl Clin Med Phys* 4: 341-51. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Metcalfe PE, Wong TP, Hoban PW (1993) Radiotherapy X-ray beam inhomogeneity corrections: the problem of lateral electronic disequilibrium in lung. *Australas Phys Eng Sci Med* 16: 155-67. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Mohan R, Chui C, Lidofsky L (1986) Differential pencil beam dose computation model for photons. *Med Phys* 13: 64-73. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Niemierko A (1997) Reporting and analyzing dose distributions: a concept of equivalent uniform dose. *Med Phys* 24: 103-10. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Oldham M, Webb S (1997) Intensity-modulated radiotherapy by means of static tomotherapy: a planning and verification study. *Med Phys* 24: 827-36. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Paelinck L, Reynaert N, Thierens H, De Neve W, De Wagter C (2005) Experimental verification of lung dose with radiochromic film: comparison with Monte Carlo simulations and commercially available treatment planning systems. *Phys Med Biol* 50: 2055-69. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Rogers DW (2006) Fifty years of Monte Carlo simulations for medical physics. *Phys Med Biol* 51: R287-301. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Rogers DW, Faddegon BA, Ding GX, Ma CM, We J, et al. (1995) BEAM: a Monte Carlo code to simulate radiotherapy treatment units. *Med Phys* 22: 503-24. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Siebers JV, Keall PJ, Nahum AE, Mohan R (2000) Converting absorbed dose to medium to absorbed dose to water for Monte Carlo based photon beam dose calculations. *Phys Med Biol* 45: 983-95. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
- Vanderstraeten B, Reynaert N, Paelinck L, Madani I, De Wagter C, et al. (2006) Accuracy of patient dose calculation for lung IMRT: A comparison of Monte Carlo, convolution/superposition, and pencil beam computations. *Med Phys* 33: 3149-58. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)

34. Wang L, Chui CS, Lovelock M (1998) A patient-specific Monte Carlo dose-calculation method for photon beams. *Med Phys* 25: 867-78. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
35. Wang L, Yorke E, Chui CS (2002) Monte Carlo evaluation of 6 MV intensity modulated radiotherapy plans for head and neck and lung treatments. *Med Phys* 29: 2705-17. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
36. Wang X, Spirou S, LoSasso T, Stein J, Chui CS, et al. (1996) Dosimetric verification of intensity-modulated fields. *Med Phys* 23: 317-27. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
37. Wu Q, Djajaputra D, Wu Y, Zhou J, Liu HH, et al. (2003) Intensity-modulated radiotherapy optimization with gEUD-guided dose-volume objectives. *Phys Med Biol* 48: 279-91. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
38. Yang J, Li J, Chen L, Price R, McNeeley S, et al. (2005) Dosimetric verification of IMRT treatment planning using Monte Carlo simulations for prostate cancer. *Phys Med Biol* 50: 869-78. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
39. Zhou SM, Das S, Wang Z, Marks LB (2004) Relationship between the generalized equivalent uniform dose formulation and the Poisson statistics-based tumor control probability model. *Med Phys* 31: 2606-9. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)

