Dosimetric impact of geometric errors due to respiratory motion prediction on dynamic multileaf collimator-based four-dimensional radiation delivery

S. Vedam
Department of Radiation Oncology, Virginia Commonwealth University, Richmond, Virginia 23298

A. Docef
Department of Electrical Engineering, Virginia Commonwealth University, Richmond Virginia 23298

M. Fix, M. Murphy, and P. Keall
Department of Radiation Oncology, Virginia Commonwealth University, Richmond, Virginia 23298

(Received 28 December 2004; revised 11 March 2005; accepted for publication 28 March 2005; published 17 May 2005)

The synchronization of dynamic multileaf collimator (DMLC) response with respiratory motion is critical to ensure the accuracy of DMLC-based four dimensional (4D) radiation delivery. In practice, however, a finite time delay (response time) between the acquisition of tumor position and multileaf collimator response necessitates predictive models of respiratory tumor motion to synchronize radiation delivery. Predicting a complex process such as respiratory motion introduces geometric errors, which have been reported in several publications. However, the dosimetric effect of such errors on 4D radiation delivery has not yet been investigated. Thus, our aim in this work was to quantify the dosimetric effects of geometric error due to prediction under several different conditions. Conformal and intensity modulated radiation therapy (IMRT) plans for a lung patient were generated for anterior-posterior/posterior-anterior (AP/PA) beam arrangements at 6 and 18 MV energies to provide planned dose distributions. Respiratory motion data was obtained from 60 diaphragm-motion fluoroscopy recordings from five patients. A linear adaptive filter was employed to predict the tumor position. The geometric error of prediction was defined as the absolute difference between predicted and actual positions at each diaphragm position. Distributions of geometric error of prediction were obtained for all of the respiratory motion data. Planned dose distributions were then convolved with distributions for the geometric error of prediction to obtain convolved dose distributions. The dosimetric effect of such geometric errors was determined as a function of several variables: response time (0-0.6 s), beam energy (6/18 MV), treatment delivery (3D/4D), treatment type (conformal/IMRT), beam direction (AP/PA), and breathing training type (free breathing/audio instruction/visual feedback). Dose difference and distance-to-agreement analysis was employed to quantify results. Based on our data, the dosimetric impact of prediction (a) increased with response time, (b) was larger for 3D radiation therapy as compared with 4D radiation therapy, (c) was relatively insensitive to change in beam energy and beam direction, (d) was greater for IMRT distributions as compared with conformal distributions, (e) was smaller than the dosimetric impact of latency, and (f) was greatest for respiration motion with audio instructions, followed by visual feedback and free breathing. Geometric errors of prediction that occur during 4D radiation delivery introduce dosimetric errors that are dependent on several factors, such as response time, treatment-delivery type, and beam energy. Even for relatively small response times of 0.6 s into the future, dosimetric errors due to prediction could approach delivery errors when respiratory motion is not accounted for at all. To reduce the dosimetric impact, better predictive models and/or shorter response times are required. © 2005 American Association of Physicists in Medicine. [DOI: 10.1118/1.1915017]

I. INTRODUCTION

Management of intrafraction respiratory motion in radiotherapy, specifically in the treatment of thoracic and abdominal tumors, has received increasing attention recently. Among the several approaches suggested in the literature, DMLC-based 4D radiotherapy has emerged as one of the possible solutions. Recent literature has also reported on the development of techniques to reliably acquire 4D CT datasets. With advances in multislice CT technology, such 4D CT image acquisition capability will be further enhanced, thereby making it possible to develop 4D treatment plans for thoracic and abdominal lesions.

Studies have also reported on the implementation of 4D radiation delivery for both gantry-mounted and robotic-arm-mounted linear accelerator architecture. The success of such an approach, especially when dealing with intrafraction respiratory motion, is dependent on the efficiency of real-time feedback of tumor/surrogate position during respiration to the radiation-delivery device. The response...
time of this feedback loop is a measure of the time that elapses between the acquisition of the anatomic position of the tumor/surrogate and the delivery device response. Tumor motion during such a response time causes a lag in the response of the delivery device, which if not accounted for introduces a latency error, thus reducing the accuracy of respiratory-synchronized 4D radiation delivery. Predicting the respiratory motion of a tumor/surrogate compensates for the lag in response and resynchronizes radiation delivery.

In increasing order of complexity, predictive filter algorithms span the range from the simpler stationary linear filter models to the highly complex adaptive nonlinear filter models. A few examples of such models have been studied recently for the purpose of image-guided and adaptive radiotherapy. Murphy et al. applied adaptive linear filters, and nonlinear Kalman filters and neural networks, to predict non-stationary respiratory motion. Sharp et al. have analyzed the performance of stationary linear filter models. In earlier work from this institution, we have compared the performance of stationary and adaptive linear models of prediction.

The prediction process, irrespective of the actual predictive filter algorithms mentioned above, is not always perfect. Consequently, tumor/surrogate positions predicted by the prediction algorithm do not always match the corresponding actual anticipated positions. Such a discrepancy in position due to an imperfect prediction leads to geometric errors. The magnitude of such geometric errors, due to imperfect prediction, referred to as prediction error herewith, increases with the response time. In a previous study at our institution, we have reported geometric errors due to the prediction of \( \pm 0.2 \) cm (1 standard deviation) for response times as low as 0.4 s.

Thus, while the geometric errors that occur due to the imperfect prediction of respiratory motion have been studied, the dosimetric impact of such geometric errors during 4D radiation therapy has not been investigated.

In this study, we evaluate the dosimetric outcomes of a treatment plan delivered according to three different scenarios: (1) shifting the MLC aperture synchronously with tumor motion according to a predicted compensation for latency, (2) shifting the MLC aperture synchronously with tumor motion but without any allowance for latency, and (3) without any adaptation to respiratory tumor/surrogate motion.

Our primary interest was in the dosimetric impact of the residual geometric error due to respiratory motion prediction that is used to compensate for the geometric error due to latency (as described in scenario (1) above) and compare this impact to the other two scenarios, as a function of several variables such as response time (0-0.6 s), beam energy (6/18 MV), treatment type (conformal/IMRT), beam direction (AP/PA), and breathing training type (free breathing/auditory instruction/visual feedback).

II. METHODS

Three types of data had to be assembled to quantify the dosimetric impact of residual geometric errors in each of the scenarios mentioned above. These were (a) respiratory motion data, (b) residual geometric error distributions due to each of the scenarios mentioned above, and (c) dose distributions from patient treatment plans. In the sections below, a description of the study design is presented.

A. Respiratory motion data

Respiratory motion information was obtained by tracking the position of the apex of the diaphragm from over 60 fluoroscopy recordings of diaphragm motion, as described by Vedam et al., from five patients, according to three different breathing training types (free breathing, audio instruction, and visual feedback). The diaphragm motion data collected above represented one-dimensional (1D) displacement of the diaphragm along the superior–inferior (SI) direction. A single session of respiratory motion lasted about 30 s at a data-sampling rate of 10 Hz.

B. Obtaining geometric error distributions

Although the primary aim of this work is to quantify the dosimetric outcome of a treatment plan delivered by shifting the MLC aperture synchronously with tumor motion according to a predicted compensation for latency (described as scenario (1)), it is important to compare this dosimetric impact to that caused by shifting the MLC aperture synchronously with tumor motion but without any allowance for latency and (scenario (2)), without any adaptation to respiratory tumor/surrogate motion (scenario (3)).

The residual geometric errors due each of the above scenarios will be defined for the rest of this work as prediction error, latency error, and respiratory motion error, respectively. Thus, in our study, error distributions that represented prediction error, latency error, and respiratory motion error were determined according to the methods described in the sections below. In general, such 1D geometric error distributions are identified by the symbol \( G_{error} \).

Figure 1 provides a graphical illustration of the same concept. As mentioned before, the error distributions computed for prediction error, latency error, and respiratory motion error are one-dimensional (1D) in nature and are computed from 1D respiratory motion data obtained through tracking the motion of a tumor/surrogate along the superior–inferior direction.

1. Prediction error distribution

Each session of respiratory motion data was provided as input to a linear adaptive filter model to obtain the predicted tumor position at several future instants in time (response time ranging from 0 to 600 ms in 200 ms increments). Prediction error at any given instant of time \( t_n \) for a response time \( \Delta \) into the future was then defined as the absolute difference between the predicted and the actual positions at that given instant, as seen in the equation below,

\[
\epsilon_{pred}(t_n, \Delta) = \text{pred}(t_{n+\Delta}) - \text{act}(t_{n+\Delta}),
\]

where \( \epsilon(t_{n+\Delta}) \) is the prediction error, \( x_{pred}(t_{n+\Delta}) \) is the predicted position for time instant \( t_{n+\Delta} \) and \( x_{act}(t_{n+\Delta}) \) is the actual position.
position at the same instant in time. The distribution of such error data obtained over an entire session thus represented the prediction error distribution and was then used for further analysis.

2. Latency error distribution

In addition to determining the dosimetric impact of prediction errors, it is important to determine the dosimetric impact of geometric errors that occur when the delivery system is simply allowed to follow the respiratory motion of the tumor with all of the associated mechanical and computational delays that represent the system response time. The term “latency” is used to refer to this type of error. In similar notation as in the previous section, latency error at any given instant of time \( t_n \) for a response time \( \Delta \) was then defined as the absolute difference between the current actual position at time \( t_n \) and the position of the tumor with a time lag due to response time of \( \Delta \), i.e., at time instant \( t_n-\Delta \), as seen in the equation below,

\[
\epsilon_{\text{latency}}(t_n,\Delta) = x_{\text{act}}(t_n) - x_{\text{act}}(t_n-\Delta).
\]

The corresponding error distribution, derived from the respiratory motion distribution, is referred to as the latency error distribution.

3. Respiratory motion error distribution

The distribution of the diaphragm’s positions during respiration was determined directly from the diaphragm position displacements from original respiratory motion data according to the equation below,

\[
\epsilon_{\text{resp. motion}}(t_n) = x_{\text{act}}(t_n).
\]

The corresponding distribution represents the respiratory motion error distribution.

4. Testing geometric error distributions for normality

Though the original geometric error distributions were used in further analysis, it is useful to study the nature of the error distributions. Thus, we tested for the Gaussian nature of the input error distributions by using the Kolmogorov–Smirnov (K–S) test that is available in the procedure PROC UNIVARIATE through the SAS (SAS Institute, Cary, NC) interface. The K–S test is essentially a goodness-of-fit test used to decide if a test sample (geometric error distribution in our case) comes from a population with a specific distribution (Gaussian in our case). The means and the standard errors in the corresponding mean values of each of the error distributions were also examined for any systematic errors.

C. Dose distributions

Conformal (CRT) and intensity modulated radiation therapy (IMRT) treatment plans for a lung patient, using the Pinnacle (Philips Medical Systems Milpitas, CA) treatment-planning system, were generated for AP/PA beam arrangements at 6 and 18 MV energies. The aim was to provide 66 Gy to at least 98% of the planning target volume (PTV). Gross tumor volume (GTV) was initially expanded to the clinical target volume (CTV) through the addition of a uniform margin of 0.8 cm to account for possible microscopic...
tumor extension. To this, an additional margin of 0.8 cm (2σ) was also added to obtain the PTV. Two different PTV-block margins of 0.6 cm and 0.8 cm were added for the 6 MV and 18 MV CRT beams, respectively, to account for beam penumbra. A 0.2 × 0.2 × 0.2 cm³ dose-grid resolution (the minimum allowed by the treatment-planning system) was used to compute dose within the tumor/surrogate. Dose distributions were thus obtained for two beam-energy values (6, 18 MV), two types of treatment (CRT and IMRT), and two beam directions (AP and PA), resulting in a total of eight planned dose distributions. IMRT plans were designed with the Pinnacle treatment-planning system for segmental IMRT delivery. Such dose distributions were referred to as planned dose distributions, and they represented the dose planned for the patient according to the given prescription. Figure 2 illustrates the method employed for obtaining a planned dose distribution. For the purposes of the analysis described further in this work, the planned dose distributions are assumed to be invariant with respect to changes in the underlying anatomy.

D. Determination of dosimetric impact: Dose convolution

The dosimetric impact of geometric errors due to prediction/latency/respiratory motion was determined by employing a dose-convolution approach. In essence, according to this approach, a planned dose distribution, when convolved with a given geometric error distribution, yields a resultant convolved dose distribution that includes the effects of the geometric error . The same concept is illustrated in Fig. 3. Accordingly, each of the eight planned dose distributions was convolved (MATLAB 7.0, © Mathworks Inc., Natick MA) with each type of geometric error distribution as follows: (a) prediction error, (b) latency error, and (c) respiratory motion error, in order to obtain the resultant convolved dose distributions that represented the effects of such geometric errors. Expressed in equation form,

\[
D_{\text{conv}} = D_{\text{planned}} \otimes G_{\text{error}},
\]

where \( G_{\text{error}} \) represents the geometric error distribution due to prediction, latency, or respiratory motion error for respiratory motion that occurs in a plane perpendicular to the central axis of the beam, \( D_{\text{conv}} \) represents the overall 3D result of the sequential convolution of the 1D \( G_{\text{error}} \) distribution with each 1D profile of the 3D planned distribution \( D_{\text{planned}} \) that is parallel to the direction of respiratory motion.

1. Validation of convolution algorithm with an ideal dose distribution

Validation of the convolution algorithm was carried out by convolving various geometric error distributions with an “ideal” planned step-function type of dose distribution...
(where maximum dose is delivered to all regions of the PTV, while zero dose is delivered to the surrounding normal structures). Such a step-function type of dose distribution represents the case for which the greatest effect of convolution is observed at the field edges.

E. Quantification of dosimetric impact: Dose difference and distance-to-agreement

While the dose distributions that included the effects of geometric errors provided data on the variations in the planned dose distribution due to geometric errors, quantification of such variations/differences between the planned and resultant convolved dose distributions was still essential in order to pass judgment on the dosimetric impact of the geometric errors described in Sec. II B. Two parameters, percent dose difference ($DD$) and distance to agreement ($DTA$), were employed for this purpose.

1. Percent dose difference

For each voxel, the $DD$ was calculated from the ratio of the difference between the convolved dose and the planned dose at that voxel to the maximum dose value in the planned dose distribution. In general, the lower the value of the percent dose difference ($DD$), the greater was the similarity between the two dose voxels under comparison.

2. Distance to agreement

For each voxel in the convolved dose distribution, the $DTA$ was defined as the minimal distance between the current convolved dose voxel and the location of a voxel in the planned distribution that has a dose value equal to the convolved dose. If an exact match for the current convolved dose voxel cannot be obtained, then an interpolation procedure is employed to obtain the corresponding $DTA$ value. In general, the lower the value of the $DTA$, the greater was the similarity between the two dose distributions under comparison, implying less of an impact due to the corresponding geometric error distribution.

3. DD and DTA threshold criteria

Percent dose difference or $DTA$ by itself does not provide a complete picture of the dosimetric impact of geometric errors. The dosimetric impact was therefore quantified by determining the percentage of voxels in the convolved dose distribution that differed from the planned dose distribution by greater than a threshold $DD$ value and a threshold $DTA$ value. Two different thresholds at (i) $>2\%$ $DD$ and 2 mm $DTA$ and (ii) $>1\%$ $DD$ and 1 mm $DTA$ were employed in this analysis. In general, the greater the percentage of voxels that exceeded each of the above threshold criteria, the greater...
was the disparity between the two dose distributions in question and therefore the greater the dosimetric impact of the corresponding geometric error distribution.

Using the above threshold criteria, comparisons of the dosimetric impact of geometric error distributions were made, according to several parameters such as response time, beam energy, treatment type, treatment-delivery mode, beam direction and breathing training type.

F. Dosimetric impact on a treatment plan

A clinical perspective of the dosimetric impact of prediction errors was obtained by importing the convolved dose distribution due to prediction and respiratory motion errors, which represented the greatest dosimetric impact from the analysis in Sec. II E, back into the Pinnacle treatment-planning system for comparison with the planned dose distribution. Conclusions were based on dose-volume histograms (DVHs) generated for the PTV and the lung.

III. RESULTS

A. Normality results for error distributions

It can be seen from Table I that the mean and the standard error of the mean for all three types of geometric error distributions (prediction, latency, and respiratory motion) are essentially zero, implying that there is no systematic error component being propagated throughout the analysis. The range of D values from the K–S test (0.03–0.08) also indicates that the geometric error distributions can be considered to be normally distributed. It can also be seen that the magnitude of respiratory motion with audio instructions (0.7 cm, 1 standard deviation) is greater than that for visual feedback (0.44 cm) and free breathing (0.33 cm).

B. Dose convolution

Figure 4 shows the dose profiles obtained from the convolution of an “ideal” planned step-function type of dose distribution (where the maximum dose is delivered to all regions of the PTV, while zero dose is delivered to the surrounding normal structures) with prediction error distributions (for response time=200–600 ms) and respiratory motion error distributions. The step-function type of dose distribution represents the case for which the greatest effect of convolution is observed at the field edges. In general, the effect of convolution is that of “blurring” of the sharp beam edges found in the original dose distribution. Thus, the dose drop-off at the beam edge is shallower after the convolution process. The degree to which this blurring occurs is dependent on the response time of the corresponding prediction/latency error distribution. The respiratory motion error distribution produces the greatest amount of blurring due to the greater magnitude of motion involved.

C. Dose difference and DTA analysis

Dosimetric impact, as defined in Sec. II E 3 above, is determined by the percentage of voxels in the convolved dose distribution that exceeds the corresponding threshold crite-
Thus, the greater the percentage of voxels exceeding the threshold criteria, the greater is the dosimetric impact.

1. General results

Figure 5 illustrates the dosimetric impact of geometric errors due to prediction, latency, and respiratory motion, as quantified by the percentage of points in the convolved dose distribution that differs from the corresponding planned dose distribution by greater than 1% of the maximum dose and a DTA of greater than 1 mm across several parameters such as beam energy (6 MV/18 MV), treatment type (CRT/IMRT) and breathing training type (free breathing, audio instruction, or visual feedback) for all AP beams. Figure 6 represents similar data for all PA beams. Figures 7 and 8, in turn, represent corresponding results, as described above, for the 2% DD and the 2 mm DTA threshold criterion. Based on the graphics shown in Fig. 5 to Fig. 8, some general conclusions can be drawn about the dosimetric impact. The dosimetric impact of prediction latency errors, in general, increases with an increase in the response time. Latency error, at a given response time, has a greater dosimetric impact than the prediction error for the same response time. This implies that prediction does in fact improve delivery accuracy and is necessary for accurate and effective 4D radiation delivery, especially for longer response times.

2. Effect of beam energy (6 vs 18 MV)

From a comparison of the results for the percentage of voxels exceeding both >2% DD and 2 mm DTA and >1% DD and 1 mm DTA, no consistent trends were seen for dosimetric impact when comparing 6 and 18 MV beam energies. Though it was initially expected that the 18 MV results would be less sensitive to the inclusion of geometric error, due to the broader penumbra, the similarity of the results with energy can be attributed to the definition of PDD, which is expressed relative to the maximum dose in the dose distribution. As the falloff with depth becomes shallower at higher energies, the number of points near the maximum dose becomes higher.
Fig. 5. Dosimetric impact of geometric errors due to prediction, latency, and respiratory motion, as quantified by the percentage of points in the convolved dose distribution differing from the corresponding planned dose distribution by >1% of maximum dose and have a distance to agreement (DTA) >1 mm. Each graph represents results at two different beam energies (6 MV and 18 MV) for a particular treatment type (CRT/IMRT) and a breathing training type (free breathing/audio instruction/visual feedback), for all AP beams. AP: anterior–posterior; CRT: conformal radiation therapy; IMRT: intensity modulated radiation therapy; Pred: prediction; Resp.: respiratory.
Fig. 6. Dosimetric impact of geometric errors due to prediction, latency, and respiratory motion, as quantified by the percentage of points in the convolved dose distribution differing from the corresponding planned dose distribution by >1% of maximum dose and have a distance to agreement (DTA) >1 mm. Each graph represents results at two different beam energies (6 MV and 18 MV) for a particular treatment type (CRT/IMRT) and a breathing training type (free breathing/audio instruction/visual feedback), for all posterior–anterior (PA) beams. CRT: conformal radiation therapy; IMRT: intensity modulated radiation therapy; Pred.: prediction; Resp.: respiratory.
Fig. 7. Dosimetric impact of geometric errors due to prediction, latency, and respiratory motion, as quantified by the percentage of points in the convolved dose distribution differing from the corresponding planned dose distribution by >2% of maximum dose and have a distance to agreement (DTA) >2 mm. Each graph represents results at two different beam energies (6 MV and 18 MV) for a particular treatment type (CRT/IMRT) and a breathing training type (free breathing/audio instruction/visual feedback), for all anterior–posterior (AP) beams. CRT: conformal radiation therapy; IMRT: intensity modulated radiation therapy; Pred.: prediction; Resp.: respiratory.
Fig. 8. Dosimetric impact of geometric errors due to prediction, latency, and respiratory motion, as quantified by the percentage of points in the convolved dose distribution differing from the corresponding planned dose distribution by >2% of the maximum dose and have a distance to agreement (DTA) >2 mm. Each graph represents results at two different beam energies (6 MV and 18 MV) for a particular treatment type (CRT/IMRT) and a breathing training type (free breathing/audio instruction/visual feedback), for all posterior–anterior (PA) beams. CRT: conformal radiation therapy; IMRT: intensity modulated radiation therapy; Pred.: prediction; Resp.: respiratory.
3. Effect of treatment type (CRT versus IMRT)

In general, CRT-based dose distributions resulted in lower numbers of voxels exceeding the threshold criteria as compared with IMRT distributions, irrespective of other parameters such as beam energy, beam direction, and breathing training type. This is indicative of a greater dosimetric impact due to error distributions for IMRT as compared with CRT. IMRT dose distributions, by nature, contain greater frequency content in their distribution spectrum as compared with CRT dose distributions, and, therefore, convolution methods affect IMRT distributions more than CRT distributions.

4. Effect of beam direction (AP versus PA)

Both AP and PA beams exhibited similar dosimetric impact regardless of treatment type and breathing training type.

5. Effect of breathing training type (free breathing versus audio instruction versus visual feedback)

In general, a significantly higher percentage of voxels exceeded the threshold criteria for error distributions obtained with audio instructions as compared with either free breathing or visual feedback. This was due to the increased amplitudes in the respiratory patterns that were seen when patients were instructed to follow audio instructions. Similar results for increased respiratory motion amplitude with audio instructions have been reported previously. Visual feedback, however, did result in a slightly higher percentage of voxels exceeding the threshold criteria, as compared with free breathing. Thus, the dosimetric impact due to error distributions obtained from respiratory motion with audio instruction was greater than error distributions with visual feedback, which was subsequently greater than error distributions obtained with free breathing respiratory motion. Improved combined audio and visual training methods should increase the respiration motion reproducibility and, hence, decrease geometric error and subsequently dosimetric error.

D. Dosimetric impact on a treatment plan

To obtain a clinical perspective about the dosimetric impact due to prediction and respiratory motion errors, PTV DVHs from distributions that exhibited the greatest dosimetric impact due to prediction and respiratory motion error were compared with the corresponding DVH from the planned dose distribution (see Fig. 9). It is interesting to note that the PTV DVH for the worst-case prediction convolved dose distribution deviates from the corresponding DVH for the planned dose distribution by a small amount. It can also be seen that the differences in the PTV volume coverage between the prediction convolved dose distribution (with response time=600 ms) and the planned dose distribution are fewer than the differences in the corresponding PTV volume coverage between the respiratory motion convolved dose distribution and the planned dose distribution. Also, a visual examination of the corresponding lung DVHs did not reveal any significant differences which was consistent with the mean lung dose values of 6.92, 6.91, and 6.87 Gy for the planned, prediction error-convolved (response time=600 ms), and respiratory motion error distributions, respectively. Thus, while respiratory motion error does not seem to significantly affect the dose distributions delivered to a patient, the dosimetric impact due to implementing 4D radiation delivery with a short response time is lesser.

IV. DISCUSSION

We have presented a dose convolution-based approach to investigate the dosimetric impact of prediction errors during 4D radiation delivery. Results from this study provide us with a basic understanding of how the various intricacies such as the respiratory motion prediction of a complex treatment strategy, such as 4D radiation therapy, affect the dose
prescribed to a tumor/surrogate volume. It has to be emphasized that prediction of a complex process such as respiratory motion is not completely error-proof. Depending on the type of prediction algorithm employed, errors arise due to imperfect prediction. It is, however, evident from the results in this study and from other studies\(^{20,21}\) that prediction does compensate for the geometric errors that occur due to system latency, which is again a function of the response time.

The response time, as defined earlier, is the sum of all the delays that occur between the acquisition of the current tumor position and the response of the delivery system to a change in tumor position. This response time depends on several factors such as the characteristics of the respiratory motion acquisition system, the linac manufacturer and the computational complexity, and efficiency of the compensation mechanisms that determine a change in position and predict the position ahead of time. The values of the response time studied here span the range of estimated response times for various respiratory motion signal and delivery system configurations and are not just limited to the configuration presented in this work.

Also, with a rotational gantry type linac and MLC-based 4D radiation delivery, the collimator rotational position is fixed during 4D radiation delivery, so that the MLC leaves move in a direction that is parallel to the long axis of tumor motion. With dynamic IMRT, one has to include constraints such as the speed of gantry rotation and maximum leaf velocity and acceleration, while with segmental IMRT, the speed of gantry rotation is not particularly important.

The convolution of the dose distributions with respiratory motion error distributions also provides a perspective of the dosimetric impact of such 4D delivery intricacies in relation to the overall dosimetric impact of patient motion. The convolution approach employed in this work makes the assumption about the invariance of the planned distributions with respect to the changes in the underlying anatomy. This assumption may hold true in only some situations.

Mechalakos et al.,\(^{31}\) in their study on dosimetric errors caused by intrafraction respiratory motion, have stated that the effects of normal breathing on tumor/surrogate PTV coverage are small on average, which is consistent with results obtained from this work (see Sec. III D). Furthermore, geometric prediction errors, at least for response times as long as 600 ms, have been reported to be less than the corresponding respiratory motion errors.\(^{21}\) Therefore, the dosimetric impact due to prediction is understandably minor when compared with the corresponding impact of respiratory motion, as is the case when 4D radiation delivery is not implemented.

Respiratory motion by itself is only one of several factors, e.g., systematic errors, caused by beam-to-bony-anatomy alignment (setup error), interfraction motion, etc., that introduce treatment-delivery errors. It has to be emphasized that the geometric error with which we are primarily concerned in this study basically represents a measure of the discrepancy between the estimated/predicted position of the surrogate/tumor and its actual position. Thus, other patient-specific errors that could alter the dose distributions, such as setup errors and interfraction motion, have not been included as part of this work, in an effort to make the study more focused. In addition, this study has also provided a quantified dosimetric comparison of the impact of beam energy, treatment type, breathing training type, and beam direction. As expected, geometric errors lead to a greater dosimetric impact for IMRT dose distributions over CRT dose distributions. While beam energy and beam direction had no consistent trend for dosimetric impact, the breathing training type used did exhibit differences in dosimetric impact, with audio instruction reporting the greatest impact.

The fact that latency error contributed to a greater dosimetric impact as compared with prediction error, especially at longer response times (>200 ms), is an important conclusion that confirms the necessity for prediction algorithms during 4D radiation delivery.

Other sources of error not considered in this study include errors in tumor/surrogate-volume delineation during 4D CT imaging that propagate through the remaining planning and delivery stages. This study also did not examine the effects of the interplay between the multileaf collimator motion and respiratory motion during 4D IMRT. Bortfeld et al.\(^{22}\) have reported on the averaging effect of respiratory organ motion on IMRT dose distributions, especially over multiple treatment fractions, leading to minimal perturbation of the original dose distribution due to the interplay effect over a whole course of treatment. Chui et al.\(^{33}\) in a similar study on lung patients, reported on the broadened and degraded coverage as a result of the interplay effect. However, they concluded that such effects for normal breathing patterns were small in magnitude. Nevertheless, any dosimetric errors due to the interplay between leaf motion and respiratory motion, especially during 4D IMRT delivery, require further attention.

Despite the above-mentioned limitations, this work does provide us with quantification of the various sources of errors that are associated with a complex treatment technique such as 4D radiotherapy. It is evident that the prediction of respiratory motion is a necessary component of DMLC-based 4D radiation delivery, given the current technical limitations of system response. It is also evident that geometric errors and, therefore, dosimetric errors due to prediction are reduced with shorter response times. Thus, the dosimetric impact due to prediction can be reduced by either employing better prediction algorithms and/or by improving the technical limitations of the system response (i.e., using a shorter response time).

ACKNOWLEDGMENTS

This research was supported by NCI Grant No. RO1 CA 93626. The authors acknowledge the useful discussions and insightful comments of Dr. Jeffrey Siebers. The authors thank Ms. Devon Murphy for carefully reviewing and significantly improving the clarity of this manuscript. The authors also acknowledge Rohini George for her valuable input on the statistical analysis of the geometric error distributions included in this paper.


