Published in final edited form as:

Technol Cancer Res Treat. 2010 June ; 9(3): 299–306.

On correlated sources of uncertainty in four dimensional computed tomography data sets

Eric D. Ehler, Ph.D.^a and Wolfgang A. Tomé, Ph.D.^{a,b,*}

^a Department of Medical Physics, University of Wisconsin, Madison WI

^b Department of Human Oncology, University of Wisconsin, Madison WI

Abstract

The purpose of this work is to estimate the degree of uncertainty inherent to a given four dimensional computed tomography (4D-CT) imaging modality and to test for interaction of the investigated factors (i.e. object displacement, velocity, and the period of motion) when determining the object motion coordinates, motion envelope, and the conformality in which it can be defined within a time based data series. A motion phantom consisting of four glass spheres imbedded in low density foam on a one dimensional moving platform was used to investigate the interaction of uncertainty factors in motion trajectory that could be used in comparison of trajectory definition, motion envelope definition and conformality in an optimal 4D-CT imaging environment. The motion platform allowed for a highly defined motion trajectory that could be as the ground truth in the comparison with observed motion in 4D-CT data sets. 4D-CT data sets were acquired for 9 different motion patterns. Multifactor analysis of variance (ANOVA) was performed where the factors considered were the phantom maximum velocity, object volume, and the image intensity used to delineate the high density objects. No statistical significance was found for three factor interaction for definition of the motion trajectory, motion envelope, or Dice Similarity Coefficient (DSC) conformality. Two factor interactions were found to be statistically significant for the DSC for the interactions of 1) object volume and the HU threshold used for delineation and 2) the object velocity and object volume. Moreover, a statistically significant single factor direct proportionality was observed between the maximum velocity and the mean tracking error. In this work multiple factors impacting on the uncertainty in 4D data sets have been considered and some statistically significant two-factor interactions have been identified. Therefore, the detailed evaluation of errors and uncertainties in 4D imaging modalities is recommended in order to assess the clinical implications of interaction among the various uncertainty factors.

Keywords

4D-CT; intrafraction motion; quality control; 4D imaging and tracking; motion envelope; maximum intensity projection (MIP)

Introduction

Four dimensional computed tomography (4D-CT) data sets are commonly used in treatment planning to provide images of anatomic motion that occurs within the delivery of a treatment fraction. It has been proposed that 4D-CT data sets be further used for dosimetric calculation (1–3), inverse planning optimization (4) and for breathing synchronized delivery (5).

*Corresponding Author: Professor Wolfgang A. Tomé, Tel: (608) 263-8510, tome@humonc.wisc.edu.

Conflict of Interest Statement: None of the authors has a conflict of interest regarding the work reported on.

Furthermore, 4D image guidance modalities such as cone beam CT (CBCT) (6,7) and dynamic MRI have been proposed as image guided radiation therapy modalities since it has been shown that in some cases the tumor motion trajectory can be variable throughout the course of treatment (8,9). An advantage of these 4D image guidance modalities is that they may allow for reduced set-up margins (10,11). In order to implement these modalities for these additional applications their inherent accuracy, or inaccuracy, must be determined. However, for each application the definition of accuracy can be different depending on whether one wishes to simply define the motion magnitude, the motion kernel, or to define the structure and shape of the anatomy throughout the breathing motion.

Variation in shape and volume of the gross tumor volume throughout the phases of 4D-CT data sets have been described in segmentation studies (12,13). Furthermore, 4D-CT image artifacts and associated errors have been considered by previous investigators (14–17). These previous studies have, cumulatively, considered 4D-CT scans of phantoms with various motion magnitudes and periods, various motion kernels, and phantom delineation at various CT density levels. The image sets used in these studies were often compared to their respective known motion paths, CT densities, and object shapes in order to determine the accuracy of the 4D images.

In consideration of the findings in prior work on the subject, this study focuses on the statistical interaction of following three factors: *the maximum object velocity*, *the object volume*, and *the density threshold* used for delineation. In other words, prior studies have only considered the main effects (single factors) individually; this study not only considers main effects but also considers the possible interaction of multiple factors. The dependence of these factors is considered with respect to ability to determine: 1) the motion kernel, 2) the motion envelope (ME), and 3) the shape and volume of scanned objects. Furthermore statistical analysis was performed for the single factors in order to determine which had a greater influence on the ability to define the motion kernel, ME and the shape and volume of the scanned objects. It should be noted that this study only reports results for one method of 4D-CT acquisition. Similar studies for other 4D-CT acquisition methods and other 4D imaging modalities, such as 4D-CBCT and dynamic MRI may differ from those observed here. They can, however, still be analyzed employing the methodology outlined in this study.

Methods and Materials

A phantom was designed that consisted of 4 glass spheres embedded in polyurethane foam. The measured diameters of the four glass spheres were 48.25 ± 0.14 mm, 40.60 ± 0.15 mm, 34.84 ± 0.35 mm, 22.21 ± 0.09 mm for Spheres 1–4, respectively. The diameters were measured at 16 distinct points on each glass sphere and the uncertainty values result from the 95% confidence interval of the 16 measured diameters. The density of the glass spheres determined in a static CT scan was determined to be approximately 1700 Hounsfield Units (HU) and the polyurethane foam had an approximate density of –950 HU.

Nine motion trajectories were studied with maximum velocities ranging from approximately $1.5 - 4.0 \text{ cm s}^{-1}$. Lung tumors have been observed to have velocities up to 5.4 cm s^{-1} (18). A motion platform was utilized (Respiratory Gating Platform, Standard Imaging Inc., Middleton, WI). The device provides sinusoidal motion in one spatial dimension with various motion amplitudes and periods that can be selected by the user. In this study, the motion was limited to being orthogonal to the plane of the bore which is typically along the inferior-superior axis of the patient reference frame. The nominal motion period and displacement settings that were used are shown in Table I with the associated maximum velocity of the sinusoidal motion. From Table I it can be seen that, of the nine investigated motion paths, four share a common

maximum velocity of 3.14 cm s^{-1} (scans 1, 6, 8, and 9), three share a common displacement of 1.5 cm (scans 1, 2, and 5), and four share a motion period of 4.0 s (scans 4–7).

4D-CT scans were acquired on a 4-slice GE Discovery CT/PET scanner (General Electric Healthcare, Waukesha, WI) with an infrared marker motion tracking system (RPM, Varian Medical Systems Inc., Palo Alto, CA). Clinical scanning parameters were used for the image acquisition in order to replicate realistic scenarios. Scans were performed in cine mode with a 2.5 cm slice thickness and a 0.4 s gantry rotation. It should be noted that it is unclear if other 4D-CT systems (i.e. helical acquisition, etc.) or other 4D imaging modalities will result in significantly different results. For each scanning position, the cine mode acquisition time was performed for at least 0.5 s greater than the motion period. The scans were retrospectively phase sorted into 10 breathing phases (4D Advantage, General Electric Healthcare, Waukesha, WI).

The motion tracking system utilizes infrared markers and an infrared camera. The markers were attached to a block on a string and pulley system that was attached to the motion platform, which the phantom was also placed on. This was intended to provide a direct link between the motion of the phantom and the observed motion by the infrared camera and marker system. However, a phase lag was observed employing an in-house motion tracking analysis system. The system compared the motion of markers placed on the string pulley system and those placed on the gating platform. A mirror was then placed in order to reflect the markers on the gating platform such that the markers and their motion appeared parallel to the motion of the markers on the block, which was attached to the gating platform via the string-pulley system. Comparison of the gating platform motion with the motion of the markers attached to the string pulley system was performed for each motion technique used (cf. Table I). The magnitude of the phase lag was determined to range from 1.49 ± 0.61 degrees to 3.42 ± 0.42 degrees. The magnitude of the phase lag had an inverse relationship with the phantom displacement. The uncertainty values represent the 95% confidence interval in the phase lag.

The Pinnacle³ treatment planning system (TPS) (Philips Radiation Oncology Systems, Fitchburg, WI) was used to contour the glass spheres. Automatic contouring tools within the TPS were used to segment the images of the glass spheres at HU threshold values of -500, 0, 500, 1000, and 1500 in each phase of the 4D-CT data set. From the contours, the centroid position of the glass sphere contour was determined. Centroid positions were determined for each glass sphere at each automatic contouring threshold level. The mean disagreement in spatial coordinates throughout the 4D data set between the observed centroid position and the associated known position was determined for each glass sphere as well as each of the five automatic contouring levels. A three-factor analysis of variance (ANOVA) was performed for the mean error in centroid position with respect to the following three factors: the glass sphere volume, the threshold used to define the spheres, and the maximum velocity of the spheres of the associated 4D-CT data set. Along with the three factors described previously, analysis was also performed for two-factor and three-factor interactions. The ANOVA was performed employing the statistical analysis toolkit within MATLAB (The Mathworks Inc., Natick, MA).

The extent of the motion of each glass sphere, called the motion envelope (ME), was compared to the known motion extent of the gating platform for the five displacements considered. MEs were created in the Pinnacle³ treatment planning system for each glass sphere and for each of the automatic contouring threshold values. The known ME was determined from the measured glass sphere diameters along with measured displacements of the gating platform. The accuracy in defining the ME was analyzed by comparison of the observed ME magnitude with the known ME magnitude. Observed ME magnitude was defined as the distance from the superior most segmented slice and the inferior most segmented slice for a given sphere in the full 4D-CT data set. The difference in known and observed ME magnitudes was compared by a three-factor

ANOVA test with respect to the glass sphere volume, the HU threshold used to create the contours, and the maximum velocity of the glass spheres.

The conformity of the glass spheres contours were compared to a volumetric model representative of the spheres determined from the measured diameters. The resolution of the representative models was 0.1 mm in all three spatial dimensions. The modeled glass sphere diameters were equal to the measured mean sphere diameter. The purpose of this was to compare each glass sphere contour to the true object shape. Observed motion data from the in-house tracking system was used to determine the position of the glass spheres at the midpoint of the 4D-CT phase interval for each 4D-CT phase dataset. False positive (FP) and false negative (FN) volumes were determined from the comparison of glass sphere contours and the representative spheres at the determined positions. From this data, the Dice Similarity Coefficient (DSC) (19) was determined for each automatic contour threshold in each phase of the 4D-CT data set. The DSC can be used to determine the extent of spatial overlap between 2 regions of interest. The DSC gives a higher weighting to parts where two regions of interest agree than where they do not and ranges between 0 (no overlap) and 1 (perfect overlap). Let R_1 and R_2 be the two regions of interest to be compared for overlap then the DSC is given by the following expression:

$$DSC = \frac{2|R_1 \cap R_2|}{|R_1 \cap R_2| + |R_1 \cup R_2|} \quad [1]$$

where $|\bullet|$ denotes the volume of a region of interest. An average DSC was determined as the mean DSC value over all phases for a given glass sphere. Three-factor ANOVA, along with consideration of interactions, was performed with respect to the glass sphere volume, the HU threshold used to create the contours, and the maximum velocity of the glass spheres.

For factors determined to have a statistical significance by ANOVA, further analysis was performed with multiple comparison tests. Multiple comparison tests were performed with Scheffé's method (20) because it allows for unequal sample size and is the most conservative method. Scheffé's method is also robust regarding violations of the assumptions underlying ANOVA (i.e. homogeneity of variance and the normality of the underlying population). In order to assure the validity of utilizing parametric three-way ANOVA, the assumptions of ANOVA were tested via skew and kurtosis (normality) and by Levene's test (homogeneity of variance). The results are shown in Table II.

It can be seen that the ratio of maximum to minimum variance was less than 3.00 in all cases which suggests that the parametric ANOVA was not appreciably degraded (21–22). It should be further noted that the sample sizes were equal among groups. F-ratios are also included in the analysis, which is also referred to as the F Test statistic. The F-ratio is the ratio of the variance of the group means to the mean of the within group variances. The P -value is a function of the F-ratio and the number of degrees of freedom. F-ratios were reported as well as P -values in order to give additional insight into the level of significance for cases where the P -value is only reported as <0.0001 . Thus, for factors with similar degrees of freedom, some additional comparison is possible. For further explanation of the F-ratio, it is suggested to consult a statistics text.

Results

Results from the three factor ANOVA analyses are shown in Table III. Along with the reported P -values, the F-ratios are included in order to better compare factors with small P -values. It is important to note that F-ratios from a single factor and from interactions of multiple factors

cannot be directly compared due to the difference in the degrees of freedom used to determine the *P*-value. However, generalized comparison between factors with the similar degrees of freedom can be made.

3.1. Motion trajectory

The observed motion of the glass spheres was compared to the known motion (cf. Figure 1). In Figure 1, the determined centroids of the glass sphere contours are shown at each automatic contour threshold value and they are compared to the known position of the glass sphere at the associated phantom motion phase. The known motion is shown with and without the phase lag correction (black and grey lines respectively). It can be seen that the correction for the phase lag results in better agreement between the known and observed motion coordinates. In Figure 1a–d the motion trajectories are shown for the cases with a) 1 cm and b) 3 cm displacement; it shows the data for c) Sphere 1 and d) Sphere 4 which are the largest and smallest of the glass spheres considered.

A multiple comparison test of the glass sphere velocities (cf. Figure 2) shows that there are multiple data set pairs that are statistically significant and a relationship can be observed where the increase in maximum velocity results in an increase in the mean tracking error. In Figure 2, the tracking error is the median tracking error for a 4D data set.

Figure 3 shows the distribution in the spatial difference between the observed and known sphere motion throughout all 4D-CT image sets. It can be seen that tracking errors are distributed with a mean of approximately 0 cm and with most points falling between ± 0.5 mm, which is a fairly high degree of precision. It should be noted that this study was limited to motion in one dimension, hence the tracking error will likely increase with motion in all three spatial dimensions. However, tracking error in the superior – inferior direction would be expected to be the greatest due to CT slice thickness resulting in limited resolution in that direction.

3.2. ME magnitude

In terms of the comparison of the known ME and the observed ME magnitudes, the HU threshold used for the automatic contouring had the greatest significance (cf. Table III). Again a multiple comparison test was performed (cf. Figure 4); it can be seen that HU thresholds of 0 and greater result in underestimation of the ME magnitude while at the lowest threshold the observed ME magnitude tends to be an overestimate. The definition of the ME magnitude at 0 HU, which is slightly less than the half maximum HU intensity over the background, resulted in a mean error of close to 0 mm. It can be observed that there is a strong linear relationship with the greatest HU thresholds resulting in the greatest underestimation of the ME magnitude. The dependence on the HU threshold is likely due to partial projection effects (17) as well as partial volume effects.

3.3. Conformity

In addition to testing the ability to define a moving object's trajectory and the ME magnitude, the constancy of the observed object volume was also considered. Figure 5a–b show contoured volumes throughout all phases of a given 4D-CT for two motion trajectories with maximum velocities of 1.57 and 3.93 cm s⁻¹ for Sphere 1. It can be seen that the contours based on the 1500 and -500 HU thresholds result in the greatest variation in observed volume. Furthermore, the volume is smaller for the higher HU threshold values. It can also be seen that the change in observed volume is greater for the case with the greatest maximum velocity and that the observed change in volumes resembles the sinusoidal velocity pattern of the phantom motion. Thus, the motion artifacts are greatest in the 20% and 70% phases where the sphere velocity is greatest.

The DSC is determined in part by the FP and FN. Figure 6 shows distributions of a) the relative FP and b) FN volumes as well as c) the DSC for Sphere 4 for the case of a maximum velocity of 3.93 cm s^{-1} . The relative FP and FN volumes are obtained by dividing the FP and FN volumes by the total known volumes of the sphere under consideration. From Figure 6 it can be seen that the relative FN volume is greatest for the contours defined at HU thresholds near the true density of the glass spheres while the relative FP volume is greatest for the lowest HU thresholds. It can be seen in Figure 6 that the DSC does typically degrade in the phases (20% and 70%) where the associated velocity of the spheres is greatest (cf. Table I) because of increased motion artifacts.

The best conformity, as defined by the DSC, is found for contours based near half of the true density intensity over the background (cf. Figure 6 and Figure 7), which is similar to findings in previous studies (17). It can further be seen from Table III and Figure 7 that a statistically better conformity was found for the 500 and 0 HU threshold based contours and for larger sphere volumes (cf. Figure 7b) and lower maximum phantom velocities (cf. Figures 7a and 7c). The interaction between the factors Volume $\times$ HU Threshold and the Velocity $\times$ Volume were found to be statistically significant (cf. Table III and Figure 7). The interaction of Volume $\times$ Threshold was the strongest interaction observed over all the data considered because of the lesser P-value and greater F-value. This is supported by the fact that the F-ratio associated with Volume $\times$ Threshold ($F = 30.28$) is much greater than for any other two-factor interactions considered, including those for the tracking error and ME magnitude error data. It can be seen in Figure 7 that the effect of the HU intensity used to define the spheres was affected by the size of the glass sphere, i.e. the effect of varying the HU threshold was greater for smaller spheres.

Discussion

While this experiment is a first principles design compared to real clinical cases, from this data it appears it is possible that when comparing the conformality between the segmented moving target in 4D-CT and the actual moving target, complex relationships exist between the velocity and volume of the moving target as well as the HU threshold used to define the structure. Significance in the interaction of Velocity $\times$ Volume suggests that restricted breathing techniques in motion management may be more or less beneficial depending on the volume of the moving structure, with smaller volume targets benefiting more from motion restriction (cf. Figure 7a). Similarly, the significance in the interaction of Volume $\times$ HU Threshold suggests that the optimal Window and Level combination in a CT scan used for segmentation may also be a function of the volume of the target of interest. For the factor combinations determined not to have a significant interaction, this data suggests that when there is no complex relationship however the single factor dependencies still apply. It is stressed that further investigation is needed in order to determine if the preceding is manifested clinically and, if so, to what degree.

Characterization of 4D-CT artifacts with respect to parameters such as the velocity and volume of the target structures as well as the HU level used to define them is necessary to determine the feasibility of implementing further uses of 4D-CT data in the treatment planning process. For the 4D-CT system investigated, interaction of multiple factors have been found to be statistically significant for some two-factor combinations, in terms of DSC (cf. Table III). However, further investigation is required to determine how correlation of these factors is manifested clinically. In order to determine the clinical implications, anthropomorphic phantoms with defined intrafraction motion are likely required along with dosimetric analysis for the cases where 4D-CT data is proposed to be used in dosimetric calculations. Such proposed uses are to calculate dose on each breathing phase (1–3) and to optimize output fluence on each 4D-CT phase (4).

In real patient imaging cases it should be considered that patient anatomy is typically irregularly shaped and/or has inhomogeneous density. This may have an effect on the accuracy of a 4D imaging modality. Furthermore, 4D data sets are heavily dependent on breathing regularity (23–25) and different image binning algorithms are being considered to reduce the effects of breathing irregularities during image acquisition (26). These factors were not considered in this work in order to focus on the inaccuracies in a best case scenario in tracking well defined objects in a high contrast environment. Future work will include more realistic 4D imaging environments. An interesting study (27) has described the use of porcine lungs in a breathing chest phantom imaged with multiple modalities; however it may be beneficial to study a case where the true motion is user defined with a high precision in order to compare the motion in the image set with a known user defined motion. Furthermore, consideration of multiple factors for other 4D imaging modalities (i.e. CBCT, dynamic MRI, etc.) is also required for both simple and realistic motion phantom configurations. Since other 4D modalities tend to have larger uncertainties than 4D-CT (27) and some future motion management techniques propose to utilize multiple 4D imaging modalities, further intermodality comparisons are required.

This study defines errors in motion observed in 4D-CT, which agree with the work of others (i.e. the single factor ANOVA analysis in Table III). Moreover, this study has also considered interactions of multiple factors and has found significance in some two factor combinations for DSC. More work is required to determine how these interactions are manifested clinically.

Conclusion

One two and three factor ANOVA was performed for 4D-CT image sets of various sized high density spherical objects surrounded by a low density medium at various motion periods and trajectory lengths. For DSC conformality, some two factor interactions (Volume $\times$ HU Threshold & Velocity $\times$ Volume) were determined to be significant. Significance among single factors was found for definition of the motion trajectory, the motion envelope, and the DSC conformality. Similar analysis of other 4D imaging modalities is suggested as well as the intercomparison of these modalities. Finally, evaluation of errors and uncertainties in anthropomorphic phantoms is recommended in order to assess the clinical implications of interaction among uncertainty factors.

Acknowledgments

This work was supported in part by a research grant from Philips Radiation Oncology Systems and NIH R01-CA109656. The Authors would like to thank Standard Imaging Inc. for donating the use of the Respiratory Gating Platform.

References

1. Glide-Hurst CK, Hugo GD, Liang J, Yan D. A simplified method of four-dimensional dose accumulation using the mean patient density representation. *Med Phys* 2008;35:5269–77. [PubMed: 19175086]
2. Rietzel E, Chen GTY, Choi NC, Willet CG. Four-dimensional image-based treatment planning: target volume segmentation and dose calculation in the presence of respiratory motion. *Int J Radiat Oncol Biol Phys* 2005;61:1535–50. [PubMed: 15817360]
3. Starkschall G, Britton K, McAleer MF, Jeter MD, Kaus MR, Bzdusek K, Mohan R, Cox JD. Potential dosimetric benefits of four-dimensional radiation treatment planning. *Int J Radiat Oncol Biol Phys* 2009;73:1560–5. [PubMed: 19231098]
4. Zhang T, Jeraj R, Keller H, Lu W, Olivera GH, McNutt TR, Mackie TR, Paliwal B. Treatment plan optimization incorporating respiratory motion. *Med Phys* 2004;31:1576–86. [PubMed: 15259662]
5. Tewatia D, Zhang T, Tomé W, Paliwal B, Mehta M. Clinical implementation of target tracking by breathing synchronized delivery. *Med Phys* 2006;33:4330–6. [PubMed: 17153412]

6. Sonke JJ, Rossi M, Wolthaus J, van Herk M, Damen E, Belderbos J. Frameless Stereotactic Body Radiotherapy for Lung Cancer Using Four-Dimensional Cone Beam CT Guidance. *Int J Radiat Oncol Biol Phys* 2009;74:567–574. [PubMed: 19046825]
7. Poulsen PR, Cho B, Keall PJ. A method to estimate mean position, motion magnitude, motion correlation, and trajectory of a tumor from cone-beam CT projections for image-guided radiotherapy. *Int J Radiat Oncol Biol Phys* 2008;72:1587–96. [PubMed: 19028282]
8. Britton KR, Starkschall G, Tucker SL, Pan T, Nelson C, Chang JY, Cox JD, Mohan R, Komaki R. Assessment of gross tumor volume regression and motion changes during radiotherapy for non-small-cell lung cancer as measured by four-dimensional computed tomography. *Int J Radiat Oncol Biol Phys* 2007;68:1036–46. [PubMed: 17379442]
9. Bosmans G, van Baardwijk A, Dekker A, Ollers M, Boersma L, Minken A, Lambin P, De Ruyscher D. Intra-patient variability of tumor volume and tumor motion during conventionally fractionated radiotherapy for locally advanced non-small-cell lung cancer: a prospective clinical study. *Int J Radiat Oncol Biol Phys* 2006;66:748–53. [PubMed: 17011450]
10. Guckenberger M, Krieger T, Richter A, Baier K, Wilbert J, Sweeney RA, Flentje M. Potential of image-guidance, gating and real-time tracking to improve accuracy in pulmonary stereotactic body radiotherapy. *Radiother Oncol* 2009;91:288–295. [PubMed: 18835650]
11. Mutaf YD, Brinkmann DH. Optimization of internal margin to account for dosimetric effects of respiratory motion. *Int J Radiat Oncol Biol Phys* 2008;70:1561–70. [PubMed: 18374230]
12. Ehler ED, Bzdusek K, Tomé WA. A method to automate the segmentation of the GTV and ITV for lung tumor. *Medical Dosimetry* 2009;34:145–53. [PubMed: 19410144]
13. Weiss E, Wijesooriya K, Dill SV, Keall PJ. Tumor and normal tissue motion in the thorax during respiration: Analysis of volumetric and positional variations using 4D CT. *Int J Radiat Oncol Biol Phys* 2007;67:296–307. [PubMed: 17189078]
14. Simon L, Giraud P, Servois V, Rosenwald JC. Initial evaluation of a four-dimensional computed tomography system, using a programmable motor. *J Appl Clin Med Phys* 2006;7:50–65. [PubMed: 17533356]
15. Starkschall G, Desai N, Balter P, Prado K, Luo D, Cody D, Pan T. Quantitative assessment of four-dimensional computed tomography image acquisition quality. *J Appl Clin Med Phys* 2008;8:1–20. 2007.
16. Mutaf YD, Brinkmann DH. An investigation of temporal resolution parameters in cine-mode four-dimensional computed tomography acquisition. *J Appl Clin Med Phys* 2008;9:172–80.
17. Rietzel E, Pan T, Chen GTY. Four-dimensional computed tomography: image formation and clinical protocol. *Med Phys* 2005;32:874–89. [PubMed: 15895570]
18. Shirato H, Suzuki K, Sharp GC, Fujita K, Onimaru R, Fujino M, Kato N, Osaka Y, Kinoshita R, Taguchi H, Onodera S, Miyasaka K. Speed and amplitude of lung tumor motion precisely detected in four-dimensional setup and in real-time tumor tracking radiotherapy. *J Radiat Oncol Biol Phys* 2006;54:1229–1236.
19. Dice LR. Measures of the amount of ecologic association between species. *Ecology* 1945;26:297–302.
20. Hochberg, Y.; Tamhane, AC. *Multiple Comparison Procedures*. Wiley; 1987.
21. Box GEP. Some theorems on quadratic forms applied in the study of analysis of variance problems, I. Effect of inequality of variance in the one-way classification. *Ann Math Stat* 1954;25:290–302.
22. Moore, DS. *The basic practice of statistics*. Freeman and Co.; NY: 1995.
23. Cai J, Read PW, Baisden JM, Lerner JM, Benedict SH, Sheng K. Estimation of error in maximal intensity projection-based internal target volume of lung tumors: a simulation and comparison study using dynamic magnetic resonance imaging. *Int J Radiat Oncol Biol Phys* 2007;69:895–902. [PubMed: 17889270]
24. Cai J, Read PW, Sheng K. The effect of respiratory motion variability and tumor size on the accuracy of average intensity projection from four-dimensional computed tomography: an investigation based on dynamic MRI. *Med Phys* 2008;35:4974–81. [PubMed: 19070231]
25. Park K, Huang L, Gagne H, Papiez L. Do maximum intensity projection images truly capture tumor motion? *Int J Radiat Oncol Biol Phys* 2009;73:618–25. [PubMed: 19147026]

26. Langner UW, Keall PJ. Accuracy in the localization of thoracic and abdominal tumors using respiratory displacement, velocity, and phase. *Med Phys* 2009;36:386–93. [PubMed: 19291977]
27. Biederer J, Dinkel J, Remmert G, Jetter S, Nill S, Moser T, Bendl R, Thierfelder C, Fabel M, Oelfke U, Bock M, Plathow C, Bolte H, Welzel T, Hoffmann B, Hartmann G, Schlegel W, Debus J, Heller M, Kauczor HU. 4D-imaging of the lung: reproducibility of lesion size and displacement on helical CT, MRI, and cone beam CT in a ventilated *ex vivo* system. *Int J Radiat Oncol Biol Phys* 2009;73:919–26. [PubMed: 19215826]

Figure 1.

a–d. Comparison of observed motion determined at each automatic contouring threshold. Data is shown for a) 4 s period and 1 cm displacement and b) 6 s period and 3 cm displacement for the largest and smallest glass spheres, c) Sphere 1 and d) Sphere 4 respectively.

480

Multiple Comparison - Sphere Velocity

Figure 2.

Multiple comparison plot of the tracking error with respect to the velocity. It is important to note that the x-axis of a multiple comparison plot is categorical and not numerical. Hence, the x-axis is not to scale but can be used to observe general trends.

Figure 3.

Histogram of the disagreement between observed and known glass sphere centroid positions throughout all 4D –CT datasets. Differences between observed and known are normally distributed around 0 with greatest differences of –0.30 cm (inferior) and 0.35 cm (superior).

490

Figure 4. Mean and associated 99% confidence interval of the underestimation of the ME magnitude with respect to the HU threshold.

Figure 5.

a–b. Observed volumes for Sphere 1 for the cases with maximum velocities of a) 1.57 and b) 3.93 cm s⁻¹. The maximum velocity for both cases occurs at about the 25% and 75% breathing phases.

Figure 6.
a–c. a) Relative FN and b) FP distributions as well as c) the associated DSC for Sphere 4 for the 4 s period and 2.5 cm motion case. The best conformity defined by the DSC in this case was the contours based on 500 and 0 HU.

Figure 7.

a–c. Interaction plots for the conformity with respect to the volume and velocity of the glass spheres as well as the HU threshold used to generate the contours. DSC values are shown in the y axis with velocity and HU threshold shown in the x axes. Glass spheres S1–S4 are the largest – smallest of the glass spheres, respectively.

Table I

Motion parameters for the associated 4D-CT scans

Scan	Period (s)	Displacement (cm)	Max. Velocity (cm s ⁻¹)
1	3.0	1.5	314
2	3.5	1.5	2.69
3	3.5	2.0	3.59
4	4.0	1.0	1.57
5	4.0	1.5	2.36
6	4.0	2.0	3.14
7	4.0	2.5	3.93
8	5.0	2.5	3.14
9	6.0	3.0	3.14

Table II

Analysis of normality and homogeneity of variance described by the ratio of the maximum to minimum variance.

	Normality		Equality of Variance	
	Test	Value	Group	Ratio of σ
Tracking Error	Skew	0.172	Velocity	2.49
	Kurtosis	2.699	Size	1.6
			HU	1.1
Motion Envelope	Skew	0.153	Velocity	1.37
	Kurtosis	2.436	Size	1.23
			HU	1.11
Dice Coefficient	Skew	-1.745	Velocity	2.87
	Kurtosis	3.959	Size	1.41
			HU	2.88

Table III
P-values and F-ratios for three-factor ANOVA analysis with two-factor interactions (significance level set at $P = 0.01$)

	<i>Tracking</i>		<i>ME Magnitude</i>		<i>DSC</i>	
	<i>P - Value</i>	<i>F - ratio</i>	<i>P - Value</i>	<i>F - ratio</i>	<i>P - Value</i>	<i>F - ratio</i>
Velocity	< 0.0001	39.92	0.0243	2.80	< 0.0001	452.61
Volume	0.0289	3.22	0.8235	0.30	< 0.0001	63.20
Threshold	0.2452	1.40	< 0.0001	49.39	< 0.0001	272.37
Velocity × Volume	0.0118	2.30	0.7868	0.69	0.0009	3.35
Velocity × Threshold	1.0000	0.17	0.7222	0.78	0.1291	1.50
Volume × Threshold	0.9999	0.12	0.5869	0.86	< 0.0001	30.28
Velocity × Volume × Threshold	1.0000	0.19	0.6117	0.93	0.9778	0.57