International Journal of Radiation Oncology biology • physics

www.redjournal.org

Physics Contribution

Residual Setup Errors Towards the Heart After Image Guidance Linked With Poorer Survival in Lung Cancer Patients: Do We Need Stricter IGRT Protocols?

Corinne N. Johnson-Hart, MSc,*,† Gareth J. Price, PhD,*,† Corinne Faivre-Finn, MD, PhD,*,† Marianne C. Aznar, PhD,*,† and Marcel van Herk, PhD*,†

**Manchester Cancer Research Centre, Division of Molecular and Clinical Cancer Science, School of Medical Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester Academic Health Sciences Centre, UK; and †The Christie NHS Foundation Trust, Manchester Academic Health Sciences Centre, Manchester, UK*

Received Nov 28, 2017, and in revised form May 15, 2018. Accepted for publication May 18, 2018.

Summary

Image guided radiation therapy is widely used, but data providing evidence of its direct effect on patient outcome are scarce. We related residual patient set-up errors from image guided radiation therapy data to overall survival for non-small cell lung cancer. The direction of the residual errors was found to be significant, with patients with residual shifts toward the heart having significantly worse overall survival than those with shifts away from the heart. This result was

Purpose: Image guided radiation therapy (IGRT) is widely used, but data directly relating set-up errors to patient outcome are scarce. This study investigates the relationship between residual IGRT shifts and overall patient survival and uses the observed relations to identify structures sensitive to radiation dose.

Methods and Materials: Residual shift data for 780 patients with non-small cell lung cancer were summarized for each patient over the course of treatment by determining the mean shifts, standard deviations, and the vector shift in the direction of the heart. These variables were related to overall survival, and significant variables were used to produce Kaplan-Meier plots of survival. The effect of shift directionality was studied by splitting the cohort into left, right, anterior, posterior, superior, and inferior groups and by analyzing the vector shift in the direction of the heart. The observed relationship was independently validated in an esophageal cancer cohort (n = 177).

Results: The shift data showed strong associations with survival. Left and right cohorts showed opposite directional shift effects, suggesting shifts toward the mediastinum have a negative effect on survival. Projection of the vector shift in the direction of the heart showed that patients with a residual shift toward the heart have significantly worse overall survival ($P = .007$, hazard ratio 1.091). The same effect was observed in the esophageal cancer cohort ($P = .041$, hazard ratio 1.164).

Reprint requests to: Corinne N. Johnson-Hart, MSc, Dept 32, The Christie Hospital NHS Foundation Trust Wilmslow Rd, Manchester M20 4BX, UK. E-mail: corinne.johnson@physics.cr.man.ac.uk

Conflict of interest: none.

Supplementary material for this article can be found at www.redjournal.org.

Acknowledgments—This study was supported by the Manchester Cancer Research UK Major Centre Award.

Int J Radiation Oncol Biol Phys, Vol. 102, No. 2, pp. 434–442, 2018
0360-3016/$ - see front matter
https://doi.org/10.1016/j.ijrobp.2018.05.052

independently validated in an esophageal cancer cohort.

Conclusions: Residual shift metrics derived from IGRT data can categorize patients with non-small cell lung cancer and those with esophageal cancer into populations with significantly different survival times on the basis of the size of the residual shift in the direction of the heart, thus providing evidence of the importance of using strict IGRT protocols to spare organs at risk and highlighting the heart as a dose-sensitive organ.

Introduction

Lung cancer is one of the most commonly occurring cancers worldwide (1), with approximately 46,000 new cases diagnosed each year in the UK alone (2). Patients with lung cancer receive surgery, radiation therapy, chemotherapy, or a combination of therapies, with radiation therapy playing an important role for all tumor stages. Radiation therapy plans are numerically optimized on computed tomography images acquired with the patient in treatment position. At each treatment fraction, this position must be replicated as accurately as possible because deviations from the planned position will result in a different dose distribution being delivered than was planned, which could affect the probability of tumor control and normal tissue complications. Previous studies have confirmed that the quality of the delivered radiation therapy can greatly affect patient outcomes (3) and the negative effects that deviations can have on tumor control (4).

Image guided radiation therapy (IGRT) has been widely implemented over the last decades to aid patient positioning. Typically using integrated planar kV or MV radiographs or cone beam computed tomography (CBCT), patients are imaged on the treatment machine before irradiation, and their position and anatomy are compared to that at planning (5). Many studies have demonstrated the advantages of using IGRT for patient positioning (6-8), reporting superior geometric and dosimetric conformance to the plan. As such, IGRT is now routinely used for the correction of patient set-up errors (9, 10), has facilitated the use of hypofractionation by providing confidence in the patient set-up, and has been applied in adaptive radiation therapy strategies (11, 12).

Much of the evidence base for IGRT, as mentioned, relies on surrogate geometric and dosimetric outcomes, with the implicit assumption that better conformance to the planned treatment correlates with improved clinical outcome. To our knowledge, the only study directly relating clinical outcome to the use of IGRT is the work of Zelefsky et al (13) in prostate cancer. Their study showed that use of daily fiducial marker–based set-up error correction improved biochemical tumor control and reduced incidence of late urinary toxicity, as compared to a cohort treated without IGRT. As discussed by Bujold et al (14) and Jaffray et al (15), because of the assumed benefit of IGRT, quantifying its efficacy in prospective clinical trials is problematic, given the difficulty in arguing equipoise between the 2 arms. In this context, retrospective observational studies may provide insight.

After its immediate clinical use, IGRT data are generally archived in accordance with national data retention guidelines and are rarely analyzed further. In our study, we hypothesized that IGRT data could provide much deeper insight. Although some simulation studies have looked at the effect of residual set-up errors from different imaging protocols—for example, Han et al (8)—to the best of our knowledge, no studies have yet directly explored the relationship between set-up error data and patient outcomes. We analyzed a cohort of 780 patients with non-small cell lung cancer (NSCLC) who were treated with curative-intent radiation therapy and verified by CBCT-based IGRT. The aim of our study was to assess whether the magnitude of residual set-up errors after IGRT is associated with patient survival and, should any such relationship exist, whether the directionality of the residual errors can provide evidence of the underlying cause.

Methods and Materials

Anonymized routine clinical data for 780 patients with NSCLC and 167 patients with esophageal cancer treated with curative intent were collated with the approval of the UK Health Research Authority and the local Caldicott information governance committee (ethical approval reference no. 17/NW/0060).

Lung patients were treated with standard fractionation regimes of either 55 Gy in 20 fractions or 60 to 66 Gy in 30 to 33 fractions, with or without chemotherapy delivered sequentially or concurrently. Esophageal patients received either 50 Gy in 25 fractions, 55 Gy in 20 fractions, or 41.4 Gy in 23 fractions, with or without concurrent chemotherapy. The IGRT protocol was identical for both patient cohorts: CBCT images were acquired at the first 3 fractions and weekly thereafter and were registered to the planning scan using a rigid registration based on bony anatomy by placing a region of interest over the spine (XVI v. 4.0 or 5.0). The obtained translations and rotations from the image match were then applied to the center of the PTV to derive the appropriate couch shift. If any of the required shifts were greater than the 5 mm action threshold in any direction, an online correction was performed. A systematic correction to the patient set-up for subsequent fractions was made after the first 3 fractions if the average shift in the first 3 fractions exceeded the 5 mm

action level. If a systematic correction was applied, further images were acquired to verify its execution. Imaging was subsequently performed weekly to verify the patient position with respect to possible positional time trends, and online corrections and further monitoring were performed if required.

The image registration translations were collected for each patient along with routine clinical variables, including patient age, sex, performance status, comorbidities score, stage, and gross tumor volume (GTV) and the fractionation regime, as detailed in Table 1. The GTV variable consisted of both true GTVs and internal tumor volumes eroded using a kernel developed in a dataset with both volumes available, as described by Johnson et al (16). The natural logarithm transform of the GTV was taken to normalize the data. Variables with a large percentage of missing data (>15% missing) were excluded, and remaining missing data were imputed using a random forest method.

Each patient's shifts were processed as shown in Fig. 1a. First, the 5 mm action threshold was applied to the raw shift data to give a measure of the residual set-up error. Couch corrections were assumed to be performed perfectly because of the lack of data on the couch start and end positions. Any systematic corrections applied to the patient set-up were ignored because these changes will be reflected in shift values obtained in later treatment fractions. A nearest-neighbor weighting approach was then used to backfill fractions where no image was acquired to remove any bias resulting from overrepresentation of the first 3 fractions (which are always imaged for every patient) and to incorporate any longitudinal trends in the errors (10).

The weighted residual shift data then were summarized as a number of shift metrics (including the mean and standard deviation of the shifts in the lateral, longitudinal, and vertical directions) and as the vector shift to the heart, represented schematically in Fig. 1b, using the center of mass (CoM) as a representative point for the heart position. Because heart delineations are not available for all patients, we used a linear fit to synthesize the heart CoM from the lung CoM (because lung contours were available for all patients).

Briefly, using a cohort of 554 lung patients, all of whom had both a lung and a heart contour available, plots of the lung versus heart CoM position in each of the x, y, and z directions were created and linear fits determined. These linear fits then were used to convert lung CoM coordinates to heart CoM coordinates. The linear fit quality was assessed via the coefficient of determination (R^2) and standard deviation of the residuals. Further details are provided in the Material 1 (available online at www.redjournal.org). All data were analyzed in International Electrotechnical Commission coordinates, such that positive shifts in the x, y and z directions correspond to shifts to the patient's left, superior, and anterior, respectively.

Table 1 Patient cohort details

Variable	NSCLC cohort (n = 780)	Esophagus cohort (n = 177)
Age (y), mean (range)	70 (31-94)	69 (41-88)
Sex, n (%)		
Male	421 (53.97)	113 (63.84)
Female	359 (46.03)	64 (36.16)
ECOG-PS, n (%)		
0	111 (14.23)	31 (17.51)
1	364 (46.67)	96 (54.24)
2	198 (25.38)	36 (20.34)
3	47 (6.03)	7 (3.95)
4	2 (0.26)	0
Missing	58 (7.43)	7 (3.95)
Comorbidities, n (%)		
0	79 (10.13)	35 (19.77)
1	105 (13.46)	29 (16.38)
2	89 (11.41)	24 (13.56)
3	72 (9.23)	9 (5.09)
Missing	435 (55.77)	80 (45.20)
Stage, n (%)		
I	18 (2.31)	20 (11.30)
II	113 (14.49)	41 (23.16)
III	520 (66.67)	85 (48.02)
IV	38 (4.87)	8 (4.52)
Missing	91 (11.66)	23 (13.00)
Mean GTV (cm^3)	72	39
Missing, n (%)	104 (13.33)	10 (5.65)
Fractionation, n (%)		
60-66 Gy in 30-33	159 (20.38)	-
55 Gy in 20	621 (79.62)	63 (35.59)
50 Gy in 25	-	104 (58.76)
41.1 Gy in 23	-	10 (5.65)

Abbreviations: ECOG-PS = Eastern Cooperative Oncology Group performance status; GTV = gross tumor volume; NSCLC = non-small cell lung cancer.

Patient details for each variable in the NSCLC and esophagus cohorts, giving the number of cases (and corresponding percentage of the whole cohort) in each variable level. Where there are missing data in the variable, this condition is given as an additional level.

The main analysis took place in the NSCLC cohort. Univariate analysis of any relationship between clinical variables and the shift metrics was undertaken via Pearson correlation and analysis of variance. The resulting P values were adjusted using the Benjamini and Hochberg False Discovery Rate method (17) to correct for the effects of multiple comparisons. A separate comparison was performed for the comorbidities score, using the subset of the data with a comorbidity score available (n = 345), because of the large amount of missing data. Elastic net penalized Cox regression with equal ridge regression and least absolute shrinkage and selection operator penalty terms were used to select the variables most strongly related to patients' overall survival. The most significant residual shift variables were used to categorize patients into low- and high-shift groups, using an optimal cut-point determined by

a

Raw Shifts

X (mm)	Y (mm)	Z (mm)	Fraction
6.8	5.5	-3.1	1
2.6	1.5	-2.8	2
1.3	3.4	-3.3	3
-0.5	1.1	1.5	8
5.1	-0.8	-2.5	13
-1.9	-3.7	0.2	14
1.8	1.7	0.4	15
8.1	0.1	-0.6	20

Residual Shifts

X (mm)	Y (mm)	Z (mm)	Fraction
0	0	0	1
2.6	1.5	-2.8	2
1.3	3.4	-3.3	3
-0.5	1.1	1.5	8
0	0	0	13
-1.9	-3.7	0.2	14
1.8	1.7	0.4	15
8.1	0.1	-0.6	20

Weighted Residual Shifts

X (mm)	Y (mm)	Z (mm)	Fraction
0	0	0	1
2.6	1.5	-2.8	2
1.3	3.4	-3.3	3
1.3	3.4	-3.3	4
1.3	3.4	-3.3	5
-0.5	1.1	1.5	6
-0.5	1.1	1.5	7
-0.5	1.1	1.5	8
-0.5	1.1	1.5	9
-0.5	1.1	1.5	10
⋮	⋮	⋮	⋮

b

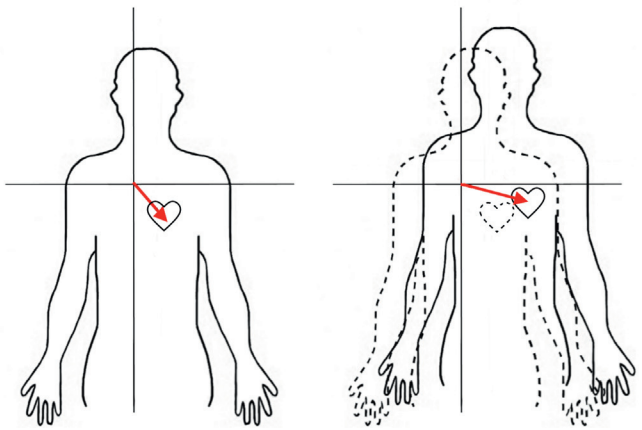

Fig. 1. (a) Schematic of how the residual image guided radiation therapy (IGRT) shifts were calculated by first retrospectively applying the shifts that were over the 5 mm action threshold (highlighted) and then using a nearest-neighbor weighting approach to backfill data in fractions where no imaging was used. (b) Schematic of how the vector shift to the heart was calculated. Using the center of mass (CoM) as a representative point for structure position, the vector length between the planning target volume CoM—taken to be the region of high dose represented here by crosshairs—and the heart CoM was determined. The residual shifts then were applied to the heart position, and this vector length was recalculated. The difference between these 2 values, represented by the 2 red arrows, determines whether a shift moves the heart toward or away from the high-dose region. (A color version of this figure is available at www.redjournal.org.)

maximizing the log-rank statistic between the 2 groups. Kaplan-Meier survival curves for each group were plotted, and multivariate analysis of both the categorized risk groups and the continuous residual shift magnitudes alongside well-known prognostic factors was performed using Cox regression.

Initial analysis looked at the whole cohort and then at subsets grouped by tumor position relative to the heart CoM to determine any directional effect. These subsets were limited to patients (n = 628) for whom both a GTV—internal tumor volume and heart—lung delineation were available. Additionally, the vector shift to the heart (toward or away) was evaluated. Results were validated in the esophagus cancer cohort using the same methodology. All analysis was performed in R version 3.3.2 (18-23).

Results

The volume of missing data for each variable is reported in Table 1. All variables had an acceptable level of missing data, apart from the comorbidities score, which was excluded from subsequent analysis.

NSCLC data

Univariate analysis found the shift variables to be independent of the clinical variables listed in Table 1. Additionally, no relation was observed between shift and comorbidities in the subset of patients for whom a comorbidity score was available (n = 345). Figure 2 shows

representative plots, along with their corresponding analysis of variance P values and Pearson correlation coefficient (where relevant).

Variable selection found the standard deviation of the lateral shift, the vector shift to the heart, age, Eastern Cooperative Oncology Group performance status, fractionation, and GTV volume to be significantly related to overall survival. When the shift direction toward or away from the heart was taken into account, the mean longitudinal, vertical, and lateral shifts were also selected in addition to those listed previously.

Table 2 details the split points used to categorize the risk and the multivariate analysis results for the different shift metrics selected as significant by regularized Cox regression. Results from the inferior and anterior cohort results are not provided because of the limited cohort sizes (n = 62 and n = 29, respectively).

It is clear from Table 2 that the right and left cohorts show opposite effects. In the left-sided tumor cohort, cases with mean residual shifts >0.1 mm had decreased overall survival. In the right-sided tumor cohort, cases with mean residual shifts >0.4 mm had improved overall survival. This result (illustrated in Figs. 3a and 3b) suggests there is a survival effect depending on whether the dose is shifted toward or away from the mediastinum. Similar effects are seen in the superior and posterior cohorts, with cases that have mean shifts toward the mediastinum having worse overall survival. The effect of the magnitude of the 3-dimensional residual shift, evaluated as the vector shift to the heart (range −4.34 to 4.66 mm, mean −0.09 mm), on overall survival is shown in Table 2 and Fig. 3c, where it can be seen that the patients with shifts toward the heart CoM have significantly worse overall survival.

Studying the vector shift to the heart as a continuous variable, we find a multivariate Cox corrected P value of .007 and hazard ratio of 1.091 per mm shift, demonstrating an increased risk of death with increasing shifts toward the heart. The same analysis using the overall magnitude of the mean shift (ie, the undirected residual set-up errors) yields a hazard ratio of 0.998 (P = .737), demonstrating that the magnitude alone is not correlated with patient survival.

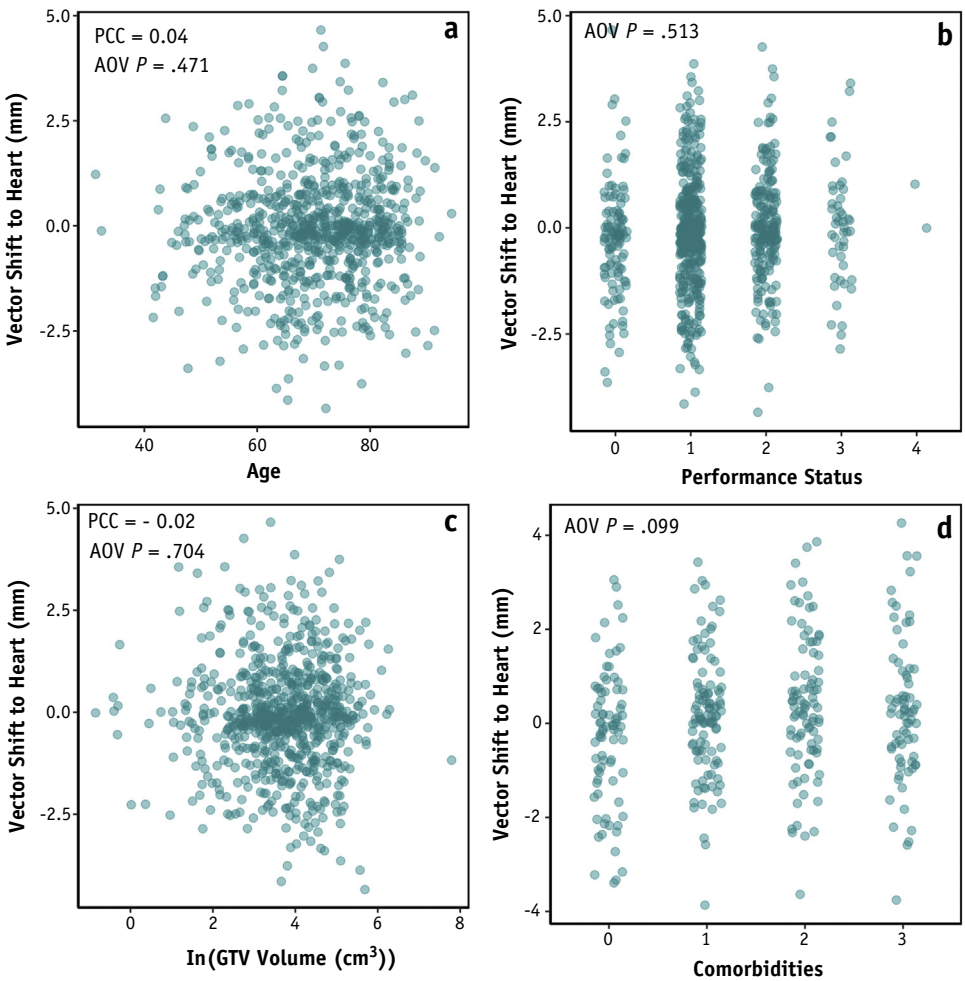

Fig. 2. Plots of the correlation with the vector shift to the heart for (a) patient age, (b) performance status, (c) gross tumor volume, and (d) comorbidities (subset of patients with comorbidity scores available, n = 345). *Abbreviations*: PCC = Pearson's correlation coefficient; AOV = analysis of variance.

Table 2 Multivariate Cox regression results

Cohort	N	Variable	*P* value	HR (low-shift group)
All	780	SD of lateral shift (>1.5 mm vs ≤1.5 mm)	<.001	1.405
		ECOG-PS	.032	1.121
		Age	.083	1.008
		Fractionation	<.001	0.954
		ln(GTV)	<.001	1.400
Left tumor cases	261	Mean lateral shift (>0.1 mm vs ≤0.1 mm)	.025	0.723
		ECOG-PS	.032	1.224
		Age	.430	1.007
		Fractionation	.044	0.966
		ln(GTV)	.002	1.263
Right tumor cases	367	Mean lateral shift (>0.4 mm vs ≤0.4 mm)	.007	1.401
		ECOG-PS	.094	1.132
		Age	.340	1.006
		Fractionation	<.001	0.943
		ln(GTV)	<.001	1.457
Superior tumor cases	566	Mean longitudinal shift (>−1.8 mm vs ≤−1.8 mm)	.011	0.664
		ECOG-PS	.020	1.151
		Age	.385	1.004
		Fractionation	<.001	0.949
		ln(GTV)	<.001	1.371
Posterior tumor cases	599	Mean vertical shift (>−1.2 mm vs ≤−1.2 mm)	.003	1.379
		ECOG-PS	.007	1.171
		Age	.220	1.006
		Fractionation	<.001	0.953
		ln(GTV)	<.001	1.386
All	780	Vector shift to heart (>0.0 mm vs ≤0.0 mm)	<.001	0.757
		ECOG-PS	.009	1.148
		Age	.214	1.006
		Fractionation	<.001	0.955
		ln(GTV)	<.001	1.405

Abbreviations: ECOG-PS = Eastern Cooperative Oncology Group performance status; HR = hazard ratio; ln(GTV) = natural logarithm of the gross tumor volume; SD = standard deviation.

Details of the split points used to assign the different residual shift metrics to risk categories, together with multivariate Cox regression HRS and p-values for each analysis. HRs associated with shift categories give the hazard of death for patients in the low residual shift group as compared with the high residual shift group (ie, HR of having low SD and low mean shifts).

Esophagus data

The vector shift to the heart in the esophagus cancer cohort was found to range between −3.85 and 4.16 mm (mean −0.13 mm). Using a split point of 0.0 mm on the vector shift to the heart (as for the lung cohort) resulted in 82 patients falling into the high-shift category and 95 falling into the low-shift category. The survival curves for the high- and low-shift groups are shown in Fig. 4.

In the NSCLC cases, patients who have a higher vector shift to the heart values have worse overall survival. This result is statistically significant when corrected for potential confounding factors (hazard ratio [low shift] = 0.660, P = .029). When the vector shift to the heart is used as a continuous variable, the hazard ratio is 1.164 per mm (P = .041).

Discussion

This study explored whether residual set-up errors directly correlate with overall patient survival after curative-intent radiation therapy. In a cohort of 780 patients with NSCLC, although the magnitude of the residual shifts over the course of treatment was not associated with poorer survival, the overall shift toward or away from the heart had a significant effect on overall patient survival (Fig. 3). The hazard ratio found when studying the vector shift to the heart as a continuous variable was 1.091 per mm, demonstrating that greater shifts toward the heart result in a greater risk (where positive shifts equate to the heart being moved toward the high-dose region, and negative shifts equate to the heart moving away). The same relation was observed when the shifts were normalized to the tumor's proximity to the heart (Material 1; available online

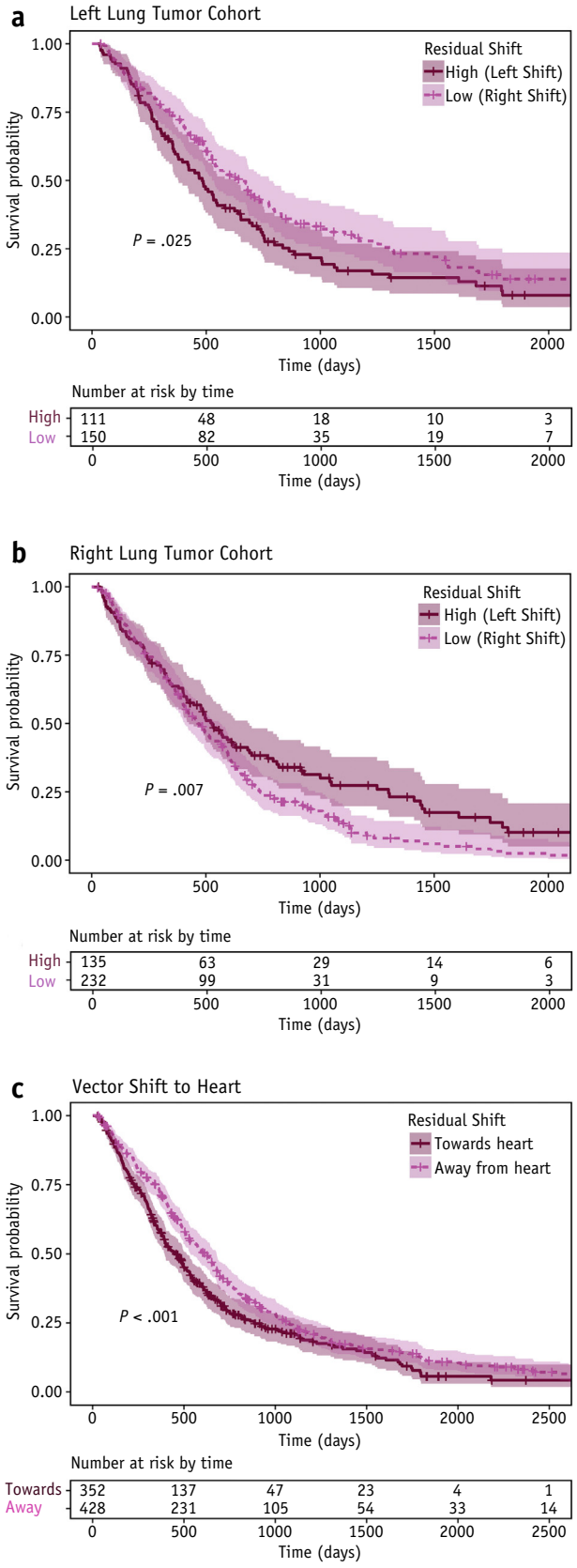

Fig. 3. Multivariate Cox regression survival curves for the (a) left (n = 261) and (b) right (n = 367) cohorts stratified into risk groups using the mean lateral shift split points detailed in Table 2 and (c) the whole cohort (n = 780) stratified on the vector shift toward or away from the heart (split point 0.0 mm). Patients with high shifts in the left cohort (those >0.1 mm, meaning that the majority of shifts will be to the left, equating to moving the heart closer to the high-dose region) have worse overall survival (P = .025), whereas patients with high shifts in the right cohort (those >0.4 mm, meaning that the majority of shifts will be to the left, moving the heart away from the high-dose region) have improved overall survival (P = .007). For the vector shift to the heart, a significant survival difference was observed (P < .001).

at www.redjournal.org) and in a cohort of patients with esophageal cancer (n = 177), using the same split point as in the lung cohort, suggesting that the effect of small residual errors is not limited to a particular cancer subtype. Univariate analysis showed that no patient variables were significantly related to any of the selected shift metrics, suggesting that the residual shift acts as an independent factor. All survival results were still significant when corrected for the effects of well-known prognostic factors.

As far as the authors are aware, this is the first study to quantify the effect of residual errors on patient outcomes. A previous study by Han et al (8) looked at quantifying esophageal cancer dose variations caused by residual set-up errors when different less-than-daily image guidance protocols were used. That study found that, even when the most frequent less-than-daily protocol was used, residual errors resulted in an underdosing of the clinical target volume and overdosing of the heart and lungs. However, the study did not quantify the effect of these dose differences on patient survival.

Recent studies have discussed the detrimental impact of heart dose on survival. The Radiation Therapy Oncology Group 0617 phase 3 study (24), which compared standard- and high-dose radiation therapy in patients with stage III NSCLC, reported worse survival in the high-dose arm. This result seemed counterintuitive because a higher dose was expected to improve local control. However, higher lung and heart doses also were reported in the high-dose arm, which has been suggested as a possible cause of this survival difference (25). A recent study performed by McWilliam et al used image-based data mining methods to identify anatomical regions where the radiation dose correlated with survival (26). The most significant difference was found in the base of the heart; when the mean dose to the heart was used to split the patients into high- and low-dose groups, a significant difference in survival was found.

Our study has several limitations. First, the number of CBCTs per patient varied. To correct for this limitation, we used a nearest-neighbor weighting approach. Although this method is relatively robust for compensating for the over-representation of the first 3 fractions, it will not perfectly reflect the variations seen in patient position on a day-to-day basis. The effects of such imperfections, however,

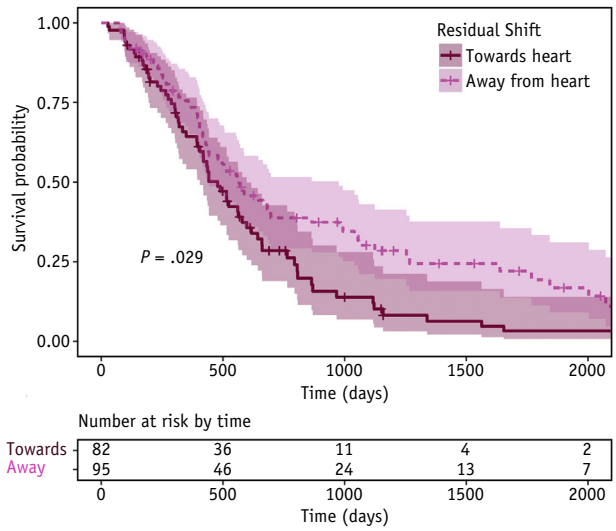

Fig. 4. Multivariate Cox regression survival curves for the high- and low-shift groups for the esophageal cancer cohort (n = 177), stratified on the 3-dimensional residual shift toward or away from the heart (split point 0.0 mm). Positive residual shifts represent the high-dose region being shifted toward the heart, whereas negative shifts represent the high-dose region being shifted away from the heart. A significant survival difference was seen between shifts toward and away from the heart ($P = .03$).

should be matched between the 2 groups we compare. Nonetheless, validation in a cohort with daily imaging is desirable to confirm this method. Additionally, the residual set-up errors are currently obtained by applying the imaging action threshold retrospectively to the recorded shift data, which assumes that the couch correction is always perfect. It would be preferable if patient couch positions were recorded at each fraction, but this was not the case. Over this large cohort of patients, we assume that the effect of imperfect corrections will be equivalent in both groups because the cohort is split at a 0 mm shift. This method of estimating the residual shifts is likely to underestimate the residuals to a certain extent (6). This underestimation is unbiased and, therefore, will not affect the primary finding (ie, that shifts toward and away from the heart affect survival). This bias will affect the results presented in Table 2, however, and the results must be interpreted carefully.

Second, our survival analysis has not been being corrected for comorbidities. However, in the cohort for which comorbidity scores were available, no significant relation was found between comorbidity score and any shift variables. Analysis should be performed again in a larger cohort to confirm that no relation exists and to allow the comorbidity score to be included in the multivariate correction procedure in case of any unexpected variable interactions.

Third, the heart CoM was chosen as a representative point for categorization. It could be that other reference points are more informative, such as the base of the heart, as discussed by McWilliam et al (26). This will be explored as part of a full dosimetric analysis.

Finally, in addition to the primary endpoint used in this study of overall survival, it would have been of interest to investigate cardiac toxicity as an endpoint. Unfortunately, this was not possible because the prospective recording of toxicities (including cardiac) in the routine setting is not robust enough. A particular issue is the recording of treatment-related cardiac deaths, which tend to be attributed to lung cancer by general practitioners without further investigations.

Other factors that could affect patient survival, such as the level of tumor shrinkage that occurs throughout the course of treatment, should also be investigated. However, given the independence of the residual shift from all of the clinical parameters we tested, we argue that it is fair to assume the distribution of tumor shrinkage and other effects to be comparable in both patient groups and hence that our results should not be unduly biased by any such effects.

As discussed in the introduction, there is little previous evidence linking IGRT directly with patient outcomes, largely because of the assumed benefit and therefore the difficulty in arguing equipoise for a clinical trial testing its benefit. This retrospective study provides some of the first direct evidence of the clinical impact of IGRT, showing that even small residual set-up errors affect survival in patients with lung cancer. Based on our results, we would suggest first that a 5 mm action threshold is too high; the patients who have the worst survival are those whose deviations are not large enough to be corrected. Our results suggest that this effect is due to toxicity and not tumor control because the magnitude of the residual errors alone does not correlate with survival. Second, we would advise that further investigations into the effect of dose to substructures of the heart be conducted and that stricter constraints be put in place. Currently, planning constraints for the heart are used (V30Gy <40% and V40Gy <30%), but these are compromised in favor of target coverage. Heart-sparing strategies, such as deep-inspiration breath hold, are being investigated for lung cancer (27), but their effectiveness seems to be highly patient-dependent. In the last year, the standard treatment imaging protocol at our institution has been updated, and CBCT images now are acquired at every fraction and the action level has been reduced to 2 mm. We intend to repeat this analysis in this new cohort once sufficient outcome data have been collected to confirm whether this change has removed the observed survival effect. The origin of the survival difference seems most likely linked to increasing or decreasing dose to the heart. We are now investigating the effect of residual set-up errors on the cumulative dose.

Conclusions

In this study, we have shown that residual IGRT shifts significantly correlate with survival. We found that patients who have a mean residual shift toward the heart have a worse prognosis compared with those who have a mean

shift away from the heart. This effect was observed in both NSCLC and esophageal cancer cohorts. These results provide solid evidence for the use of stricter IGRT protocols for thoracic radiation therapy because they show that even small residuals have a significant effect on survival; the results also provide further information on the dose response of organs at risk. We recommend that imaging action thresholds be reviewed, along with radiation therapy heart constraints, because increasing dose to the heart appears to have an early effect on survival.

References

1. Cancer Research UK. Cancer statistics - worldwide cancer. Available at: http://www.cancerresearchuk.org/health-professional/cancer-statistics/worldwide-cancer#heading-One. Accessed November 6, 2017.
2. Cancer Research UK. About cancer - lung cancer. Available at: http://www.cancerresearchuk.org/about-cancer/lung-cancer/about. Accessed March 23, 2017.
3. Peters LJ, O'Sullivan B, Giralt J, et al. Critical impact of radiotherapy protocol compliance and quality in the treatment of advanced head and neck cancer: Results from TROG 02.02. *J Clin Oncol* 2010;28:2996-3001.
4. De Crevoisier R, Tucker SL, Dong L, et al. Increased risk of biochemical and local failure in patients with distended rectum on the planning CT for prostate cancer radiotherapy. *Int J Radiat Oncol Biol Phys* 2005;62:965-973.
5. Dawson LA, Sharpe MB. Image-guided radiotherapy: Rationale, benefits, and limitations. *Lancet Oncol* 2006;7:848-958.
6. Bissonnette J-P, Purdie TG, Higgins JA, et al. Cone-beam computed tomographic image guidance for lung cancer radiation therapy. *Int J Radiat Oncol Biol Phys* 2009;73:927-934.
7. Li W, Moseley DJ, Bissonnette J-P, et al. Setup reproducibility for thoracic and upper gastrointestinal radiation therapy: Influence of immobilization method and on-line cone-beam ct guidance. *Med Dosim* 2010;35:287-296.
8. Han C, Schiffner DC, Schultheiss TE, et al. Residual setup errors and dose variations with less-than-daily image guided patient setup in external beam radiotherapy for esophageal cancer. *Radiother Oncol* 2012;102:309-314.
9. de Boer HCJ, Heijmen BJM. A protocol for the reduction of systematic patient setup errors with minimal portal imaging workload. *Int J Radiat Oncol Biol Phys* 2001;50:1350-1365.
10. de Boer HCJ, Heijmen BJM. eNAL: An extension of the nal setup correction protocol for effective use of weekly follow-up measurements. *Int J Radiat Oncol Biol Phys* 2007;67:1586-1595.
11. Nijkamp J, Pos FJ, Nuver TT, et al. Adaptive radiotherapy for prostate cancer using kilovoltage cone-beam computed tomography: First clinical results. *Int J Radiat Oncol Biol Phys* 2008;70:75-82.
12. Sonke J-J, Belderbos J. Adaptive radiotherapy for lung cancer. *Semin Radiat Oncol* 2010;20:94-106.
13. Zelefsky MJ, Kollmeier M, Cox B, et al. Improved clinical outcomes with high-dose image guided radiotherapy compared with non-IGRT for the treatment of clinically localized prostate cancer. *Int J Radiat Oncol Biol Phys* 2012;84:125-129.
14. Bujold A, Craig T, Jaffray D, et al. Image-guided radiotherapy: Has it influenced patient outcomes? *Semin Radiat Oncol* 2012;22:50-61.
15. Jaffray DA. Image-guided radiotherapy: From current concept to future perspectives. *Nat Rev Clin Oncol* 2012;9:688-699.
16. Johnson C, Price G, Khalifa J, et al. A method to combine target volume data from 3D and 4D planned thoracic radiotherapy patient cohorts for machine learning applications. *Radiother Oncol* 2018;126:355-361.
17. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc Ser B* 1995;57:289-300.
18. Ishwaran H, Kogalur UB. randomForestSRC: Random forests for survival, regression and classification (RF-SRC). Available at: https://cran.r-project.org/web/packages/randomForestSRC/index.html. Accessed May 17, 2016.
19. Ishwaran H, Kogalur U. Random survival forests for R. *RNews* 2007;7:25-31.
20. Ishwaran H, Kogalur UB, Blackstone EH, et al. Random survival forests. *Ann Appl Stat* 2008;2:841-860.
21. Kassambara A. Survminer: Drawing survival curves using "ggplot2." Available at: https://cran.r-project.org/package=survminer. Accessed March 13, 2017.
22. Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model. New York: Springer; 2000.
23. Therneau TM, Lumley T. A package for survival analysis in S; 2015; Available at: https://cran.r-project.org/package=survival. Accessed June 26, 2016.
24. Bradley JD, Paulus R, Komaki R, et al. Standard-dose versus high-dose conformal radiotherapy with concurrent and consolidation carboplatin plus paclitaxel with or without cetuximab for patients with stage IIIA or IIIB non-small-cell lung cancer (RTOG 0617): A randomised, two-by-two factorial p. *Lancet Oncol* 2015;187-199.
25. Faivre-Finn C. Dose escalation in lung cancer: Have we gone full circle? *Lancet Oncol* 2015;16:125-127.
26. Mcwilliam A, Kennedy J, Hodgson C, et al. Radiation dose to heart base linked with poorer survival in lung cancer patients. *Eur J Cancer* 2017;85:106-113.
27. Persson GF, Scherman Rydhög J, Josipovic M, et al. Deep inspiration breath-hold volumetric modulated arc radiation therapy decreases dose to mediastinal structures in locally advanced lung cancer. *Acta Oncol (Madr)* 2016;55:1053-1056.