Geometric uncertainty analysis of MLC tracking for lung SABR

To cite this article before publication: Vincent Caillet *et al* 2020 *Phys. Med. Biol.* in press https://doi.org/10.1088/1361-6560/abb0c6

Geometric uncertainty analysis of MLC tracking for lung SABR

Vincent Caillet[1,2], Benjamin Zwan[1,5], Adam Briggs[1], Nicholas Hardcastle[3,4], Kathryn

Szymura[1], Alexander Prodreka[1], Ricky O'Brien[2], Ben E Harris[7,8], Peter Greer[5,6], Carol Haddad[1],

Dasantha Jayamanne[1], Thomas Eade[1], Jeremy Booth[1, 4], Paul Keall[2]

[1] Northern Sydney Cancer Centre, Royal North Shore Hospital, St Leonards NSW

Australia

[2] ACRF Image X Institute, the University of Sydney, Australia

[3] Physical Sciences, Peter MacCallum Cancer Centre, Melbourne, Australia

[4] Institute of Medical Physics, University of Sydney, School of Physics, Sydney

[5] School of Mathematical and Physical Sciences, University of Newcastle, Newcastle,

Australia.

[6] Department of Radiation Oncology, Calvary Mater Hospital, Newcastle, Australia

[7] Department of Respiratory Medicine, Royal North Shore Hospital, Sydney, New South

Wales, Australia.

[8] Faculty of Health Sciences, University of Sydney, Australia

Abstract

Purpose

The purpose of this work was to report on the geometric uncertainty for patients treated with MLC tracking for lung SABR to verify the accuracy of the system.

Methods

Seventeen patients were treated as part of the MLC tracking for lung SABR clinical trial using electromagnetic beacons implanted around the tumor acting as a surrogate for target motion. Sources of uncertainties evaluated in the study included the surrogate-target positional uncertainty, the beam-surrogate tracking uncertainty, the surrogate localization uncertainty, and the target delineation uncertainty. Probability density functions (PDFs) for each source of uncertainty were constructed for the cohort and each patient. The total PDFs was computed using a convolution approach. The 95% confidence interval (CI) was used to quantify these uncertainties.

Results

For the cohort, the surrogate-target positional uncertainty 95% CIs were ±2.5 mm (-2.0/3.0 mm) in LR, ±3.0 mm (-1.6/4.5 mm) in SI and ±2.0 mm (-1.8/2.1 mm) in AP. The beam-surrogate tracking uncertainty 95% CIs were ±2.1 mm (-2.1/2.1 mm) in LR, ±2.8 mm (-2.8/2.7 mm) in SI and ±2.1 mm (-2.1/2.0 mm) in AP directions. The surrogate localization uncertainty minimally impacted the total PDF with a width of ±0.6 mm. The target delineation uncertainty distribution 95% CIs were ±5.4 mm. For the total PDF, the 95% CIs were ±5.9 mm (-5.8/6.0 mm) in LR, ±6.7 mm (-5.8/7.5 mm) in SI and ±6.0 mm (-5.5/6.5 mm) in AP.

Conclusion

This work reports the geometric uncertainty of MLC tracking for lung SABR by accounting for the main sources of uncertainties that occurred during treatment. The overall geometric uncertainty is within ±6.0 mm in LR and AP directions and ±6.7 mm in SI. The dominant uncertainty was the target delineation uncertainty. This geometric analysis helps put into context the range of uncertainties that may be expected during MLC tracking for lung SABR. (ClinicalTrials.gov registration number: NCT02514512).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Introduction

The goal of radiation therapy is to provide adequate target coverage while limiting dose to adjacent critical organs. Tumors however are subject to motion from physiological processes such as respiration, which necessitates motion management strategies to ensure the goal of treatment is achieved [1, 2].

Dedicated radiation therapy linear accelerators that allow to track the tumor motion in real time have been used to treat patients. Among those, the CyberKnife system (Accuray Inc. Sunnyvale, USA), the Accuray Radixact (Accuray, Sunnyvale, CA), the Vero system (BrainLab AG, Feldkirchen, Germany) and Multi-Leaf Collimator (MLC) tracking enable the treatment delivery to adapt to the target position and patient's internal dynamics in real time while the treatment is delivered.

The detailed geometric information recorded during MLC tracking treatment creates the opportunity to quantify the various sources of geometric uncertainties for the patient cohort, and each patient**.** The framework proposed by Stroom *et al.* [3] and Sawkey *et al.* [4] was adopted to quantify the geometric uncertainty, whereby individual and independent sources of uncertainties were computed as Probability Density Functions (PDFs) and then convolved to obtain the total PDF. The key sources of uncertainties pertinent in MLC tracking for lung SABR include the surrogate-target positional uncertainty, the beam-surrogate tracking uncertainty, the surrogate localization uncertainty, and the target delineation uncertainty.

The purpose of this work is to report on the geometric uncertainties of MLC tracking for lung SABR patients to verify the accuracy of the system and inform margin requirements for future applications of MLC tracking in thoracic targets.

Methods

1. Patients, beacon insertion, planning and treatment

Seventeen patients diagnosed with stage I non-small cell lung cancer or lung metastases were recruited as part of the phase I/II clinical trial (LIGHTSABR, NCT02514512) and treated using SABR with MLC tracking between November 2015 and November 2018.

Three electromagnetic transponders (Varian Medical Systems, Palo Alto, US) were bronchoscopically implanted for each patient in the airways surrounding the tumor using C-arm fluoroscopic image guidance.

Each patient received a 4D-CT simulation one week following the bronchoscopic implantation. For the 4D-CT simulation and the subsequent treatment, patients were lying supine (N=11), prone (N=5) or lateral decubitus (N=1). The prone and lateral positions were used to meet the limitations imposed by the Calypso system (Varian Medical Systems, Palo Alto, US) that electromagnetic beacons must be within 19 cm of the detector panel. Phase binning for 4D-CT imaging was performed for supine patients using the Real-time Position Management (RPM) infrared camera (Varian Medical Systems, Palo Alto, US) and the pneumatic belt (bellows, Philips Medical Systems, Cleveland, US) for prone and lateral decubitus patients. Ten phases were reconstructed with a slice thickness of 1.5 mm. Six patients out of seventeen were simulated with visual biofeedback using in-house software [5].

MLC tracking treatment planning was performed on the end-exhale phase to provide reliable tumor delineation [6], unless other phases were considered of higher quality. The gross tumor volume (GTV) was drawn by the treating oncologist. The clinical tumor volume (CTV) was assumed to equal the GTV. Isotropic margins of 5 mm were added to the CTV to define the Planning Target Volume (PTV).

Patients were planned with volumetric modulated arc therapy (VMAT) using prescribed doses of 4×12 Gy or 5×10 Gy depending on the tumor location. The planning protocol prescribed 100% of the CTV to get more than 100% of the prescribed dose and at least 98% of the PTV to get 100% of the prescribed dose. There was no criteria specified for the maximum dose in the protocol. From a review of the treatment plans, the maximum dose to the GTV did not exceed 130% of the prescribed dose. The collimator angle was aligned along the most dominant tumor

motion direction observed at 4D-CT, either in the superior-inferior (SI) direction (N=16) or in the left-right (LR)/anterior-posterior (AP) direction (N=1).

As shown in Figure 1, treatments were delivered with a Varian Trilogy linear accelerator equipped with the Millennium MLC. The patient was placed on the table such that the Calypso-measured tumor position matched with the planned position. Cone beam computed tomography was then acquired to verify the tumor position at the end-of-exhale relative to the surrounding organs. The in-house MLC tracking software was initiated to take control of the leaves and the MLC tracking treatment was delivered similarly to a standard treatment. Further details about the software can be found in Keall *et al.* [7]. A prediction algorithm [8] was used to account for the measured 230 ms system latency. The average beam-on time for two-arc VMAT was 5.03 min (± 0.5 min) using 600 MU/min with an average of 1415 MU (± 210 MU) per arc. Output logs from the MLC tracking software recorded the positions of the MLC, gantry angle and surrogate position. These were exported for analysis using MATLAB 2019 (MathWorks, USA).

2. Methods for quantifying the geometric uncertainty

Using the output logs, the geometric uncertainty was computed for both the cohort and each patient. The geometric uncertainty refers to any geometric deviations at a given time that contributed to a misalignment between the centroid of the MLC aperture and the centroid of the target.

As shown in Figure 1, following an MLC tracking treatment, four sources of uncertainty were quantified to construct four individual PDFs. Those four PDFs were then combined with a convolution method described in Stroom *et al.* [3] and Sawkey *et al.* [4] to obtain the total probability of geometric deviation. The method for building each PDF is described in the following sections. Using this total PDF, the geometric uncertainty for each patient was reported as being the 95% CI (2.5th and 97.5th percentile) [4, 9]. The 2.5th and the 97.5th percentile values were read directly from the PDF.

Each uncertainty was evaluated in the frame of reference of a static point located at the centroid of the GTV contoured in the planned 4D-CT phase. The LR, SI and AP directions correspond to a patient in the supine position. For prone and lateral patients, the uncertainties were transformed to match the same frame of reference.
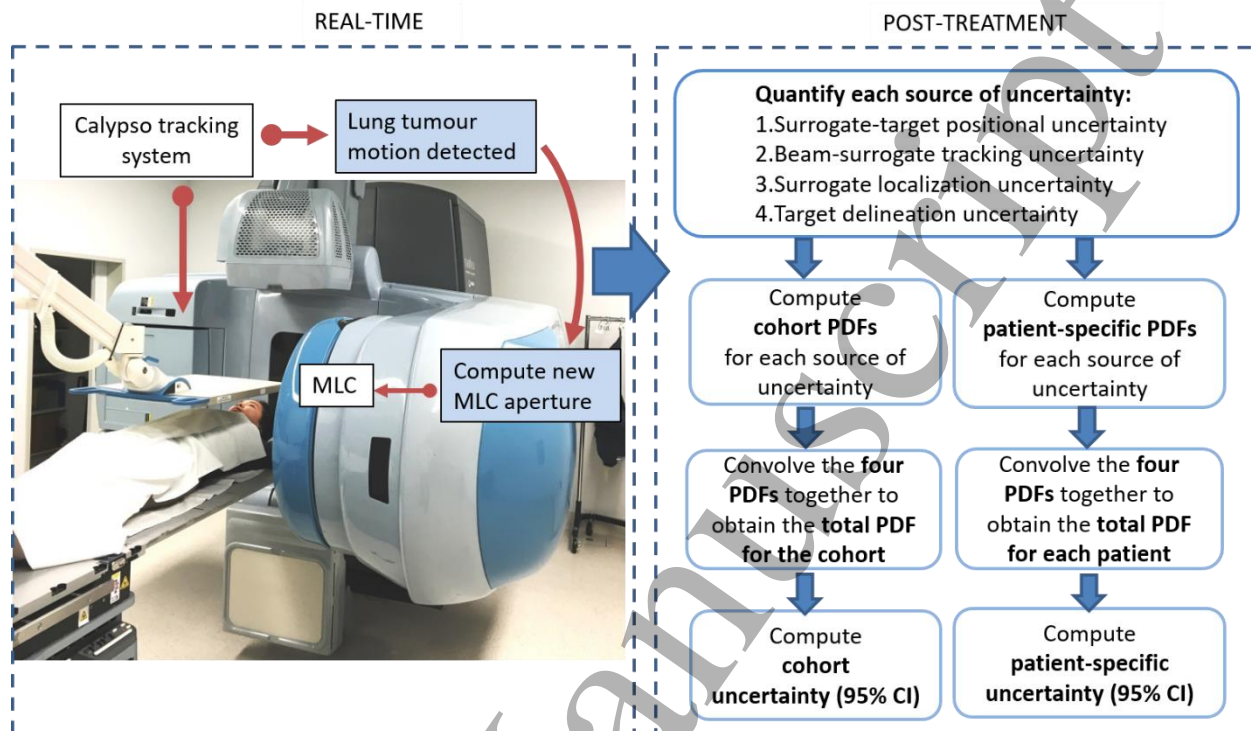
Figure 1. Schematic of the MLC tracking beam adaptation method using the Calypso tracking system. The output logs generated allowed a post-treatment analysis to quantify each source of uncertainty.

Surrogate-target positional uncertainty

The use of a surrogate to infer the target position introduces a geometric offset that needed to be accounted for. On each phase (10 phases per patient) of the 4D-CT, both the transponders and the target (GTV) were manually contoured and the position of the centroids in each phase computed. The differential motion relative to the end-of-exhale phase between the surrogate and the target constituted the surrogate-target position uncertainty.

The PDF for the surrogate-target positional uncertainties was built based on the aggregated uncertainties obtained from all 4D-CT while each patient-specific PDF was built using their individual 4D-CT.

Beam-surrogate tracking uncertainty

In the context of MLC tracking, geometric deviations due to the system's latency, finite leaf width and leaf speed are known to introduce an offset between the position of the beam and the position of the surrogate. This differential position corresponds to the beam-surrogate tracking uncertainty.

The output logs were directly read from the linac in real-time using the MLC tracking software. Agnew *et al.* [10] assessed the accuracy of the logs in terms of MLC positions and found that the errors obtained from the linac were expected to range between 0.12 mm and 0.28 m.

To compute the beam-surrogate tracking uncertainty, the actual beam position relative to the planned beam position was obtained using the output logs containing both the actual and the planned leaf positions. The leaf positions were used to draw two binary images (set to zero outside the treatment field and one within the aperture) within the beam's-eye-view corresponding respectively to the actual aperture and the planned aperture. Using an image registration algorithm based on cross-correlation [11], the two-dimensional offset between the actual and planned aperture was computed for each data entry (25 Hz) to obtain the actual aperture shift relative to its planned position.

The surrogate position (centroid of the 2-3 beacons) relative to its planned position at the end-of-exhale was obtained directly from the output logs at the same frequency (25 Hz).

Transformation of the beam-surrogate tracking uncertainties from the two-dimensional beam's-eye-view into the three-dimensional reference coordinate system was calculated assuming that the dosimetric uncertainty directly along the therapeutic path was negligible as for photon beams the dose generally varies less with depth than laterally [12].

The beam-surrogate tracking uncertainty was computed for each data entry at 25 Hz. The total PDF was built using the uncertainties from the seventeen patients concatenated into one single matrix while the patient-specific PDF were built using their individual uncertainties.

Surrogate localization uncertainty

The surrogate localization uncertainty relates to the accuracy and precision of the electromagnetic transponders' reported positions using the Calypso tracking system. This uncertainty was obtained based on published data that reported a standard deviation in the range of 0.2 to 0.6 mm for stationary [13] and moving targets [14], respectively. For this analysis, the surrogate localization

uncertainty PDF was modelled as a step function, with a mean of zero and width of 0.6 mm in each dimension.

Target delineation uncertainty

Delineation of the tumor within the treatment planning system remains an important source of geometrical uncertainty. The delineation uncertainty was taken from measurements by Peulen *et al.* [15] and Mercieca *et al.* [16] who reported on the target delineation uncertainty for a cohort of lung cancer patients. GTVs obtained from 4D-CTs were contoured by several lung radiation oncologists onto the MIP (Peulen *et al.*, Mercieca *et al.*), mid-ventilation phase (Mercieca *et al.*) and the mid-position phase (Mercieca *et al.*).

For Peulen *et al.*, the collated data (in Figure 2.a. of their manuscript) was chosen as it represents the distribution of standard deviation for the entire group of lung radiation oncologists. In Mercieca et *al.*, the Mid-V data (Figure 2, top left) was chosen as this most closely aligns with the current study of contouring on a single phase. We used these distributions to derive a PDF that was implemented in our model.

3. Quantifying the total uncertainty

For both the cohort and the patient-specific model, the total PDF was computed as the convolution of all PDFs. Kolmogorov-Smirnoff (KS) normality tests were ran for each cohort PDF. **Table 1** summarizes the key points for each PDF.

Table 1. List of uncertainties with their respective PDF functional form, uncertainty type and data source

	Surrogate-target positional uncertainty	Beam-surrogate tracking uncertainty	Surrogate localization uncertainty	Target delineation uncertainty
PDF functional form				
Cohort	PDF built from entire cohort data	PDF built from entire cohort data	Step function from the literature	PDF built from two publications
Patient-specific	PDF built from patient-specific 4D-CT	PDF built from patient-specific data	Step function from the literature	PDF built from two publications

Uncertainty type	3D*	Beam's eye view	3D	3D
Data Source	4D-CT manual segmentation	Position of surrogate and leaf positions in beam's-eye-view transformed to 3D	Literature [13, 14]	Literature [15, 16]

*3D refers to the static coordinate system in the planned 4D-CT phase.

4. Relationship between the beam-surrogate tracking uncertainties and the average surrogate peak-to-trough distance

Since the primary objective of MLC tracking is to compensate for the surrogate peak-to-trough distance, the relationship between the surrogate peak-to-trough distance and the beam-surrogate tracking uncertainty was investigated for each fraction (N=70). The Pearson correlation was used to compute the degree of linearity between the two variables.

Results

Seventeen patients were treated with MLC tracking for lung SABR, totaling 70 fractions of treatment with dual-arc VMAT.
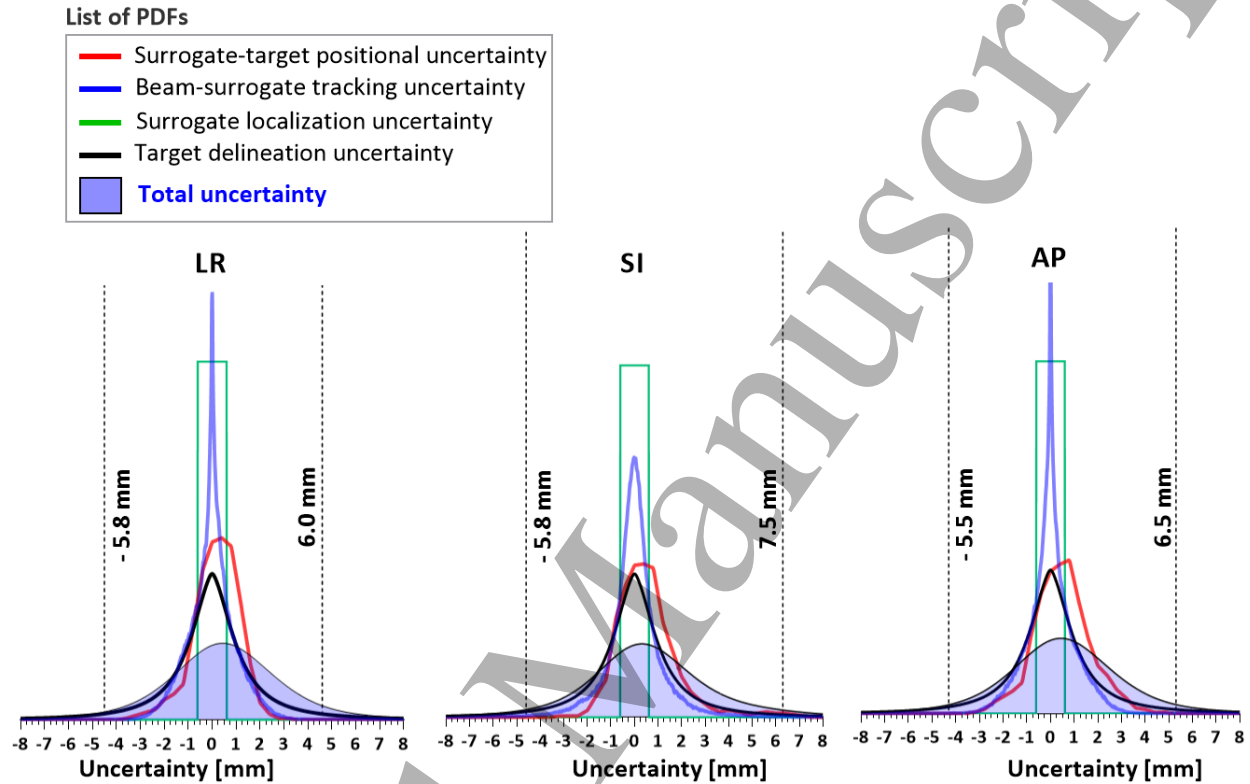
Figure 2. PDFs for the cohort in LR, SI and AP directions. PDFs for individual source of uncertainties are shown as curves and the total PDF as the shaded distribution. The uncertainties are represented as the 95% CIs (vertical lines). KS test showed that all PDF were significantly different than a normal distribution (p<0.01).

Uncertainties for the cohort

As seen in **Figure 2** with values summarized in **Table 2**, the total PDF for the cohort (shaded distributions) shows that the uncertainties, the 95% CI (2.5th and 97.5th percentile, lie within ±5.9 mm (-5.8/6.0 mm) in LR, ±6.7 mm (-5.8/7.5 mm) in SI and ±6.0 mm (-5.5/6.5 mm) in AP directions.

The average surrogate peak-to-trough distance (± standard deviation) for the entire cohort is 3.2 ± 1.7 mm in LR, 8.6 ± 5.4 mm in SI and 4.8 ± 2.6 mm in AP directions.

The surrogate-target positional uncertainty computed from the 4D-CT shows that the uncertainties (2.5th/97.5th) are approximately ±2.5 mm (-2.0/3.0 mm) in LR, ±3.0 mm (-1.6/4.5 mm) in SI and ±2.0 mm (-1.8/2.1 mm) in AP directions. The asymmetric distribution in the SI direction showing the near-maximum uncertainty of 4.5 mm is likely due to one patient (**Error! Reference source not found.**, Patient 10) with considerable visual 4D-CT artefacts.

The beam-surrogate uncertainty attributed to the MLC tracking technology contributed less or equal than 2.8 mm in all directions, with uncertainties of ±2.1 mm (-2.1/2.1 mm) in LR, ±2.8 mm (-2.8/2.7 mm) in SI and ±2.1 mm (-2.1/2.0 mm) in AP directions.

The surrogate localization uncertainty minimally impacted the total PDF with width of ± 0.6 mm.

The PDF from Peulen *et al.* provided 95% confidence intervals (CIs) of ±4.2 mm and Mercieca *et al.* ±6.5 mm. The target delineation uncertainty, being the average PDF of Peulen *et al.* and Mercieca *et al.*, had a CI of ±5.4 mm. The target delineation uncertainty dominated the overall shape of the total PDF.

Table 2. Summary of the uncertainties for the entire cohort

	Surrogate-target uncertainty [mm]			Beam-surrogate uncertainty [mm]			Surrogate localization uncertainty [mm]	Target delineation uncertainty [mm]	**Total uncertainty [mm]**		
	LR†	SI	AP	LR	SI	AP	LR/SI/AP	LR/SI/AP	LR	SI	AP
95% CI	±2.5	±3.0	±2.0	±2.1	±2.8	±2.0	±0.6	±5.4	±5.9	±6.6	±6.0
2.5th	-2.0	-1.6	-1.8	-2.1	-2.8	-2.1	-0.6	-5.4	-5.8	-5.8	-5.5
97.5th	3.0	4.5	2.1	2.1	2.7	2.0	0.6	5.4	6.0	7.5	6.5

†LR (+Left, -Right), SI (+Superior, -Inferior), AP (+Anterior, - Posterior)

Uncertainties for each patient

Error! Reference source not found. summarizes the magnitude of each source of uncertainty per patient. The first three columns show the average surrogate centroid peak-to-trough distance. Patients 2, 6, 10, 11 and 12 exhibited large (>10 mm) surrogate peak-to-trough distance in the SI direction, the largest being patient 6 with 18.3 mm. The dominant direction of motion was SI for all patients, excluding patients 5, 13 and 15 who had AP dominant motion. Largest AP motion was patient 15 with 10.4 mm. No patient exhibited LR dominant motion.

The surrogate-target positional uncertainty demonstrated that the average (± standard deviation) patient's CIs were 1.3mm ± 0.7mm in LR, 1.9mm ± 1.0mm in SI and 1.7mm ± 0.8mm in AP directions. Large uncertainties from patient 10 (±7.3 mm in SI) are believed to be due to 4D-CT under-sampling artefacts. For LR and AP, the largest uncertainty was traced to patient 2 (±3.2 mm in LR) and patient 6 (±4.0 mm in AP), both patients with large surrogate peak-to-trough distances.

For the beam-surrogate uncertainty, noticeable large uncertainties were observed for patient 2 (±5.6 mm in SI) that were mostly attributed to large and erratic surrogate motion that, due to the system's latency of 230 ms, impaired the accurate prediction of the surrogate position ahead of time. During their first fraction, patient 11 also exhibited erratic surrogate motion with large beam-surrogate tracking uncertainties.

Table 3. Summary of average surrogate peak-to-trough distance and uncertainties for each patient (95% CI). Bold text is used to highlight the patient with the largest uncertainty in each column.

Patient Id	Average surrogate peak-to-trough distance [mm]			Surrogate-target positional uncertainty [mm]			Beam-surrogate tracking uncertainty [mm]			Surrogate localization uncertainty [mm]	Target delineation uncertainty [mm]	**Total uncertainty**		
	LR	SI	AP	LR	SI	AP	LR	SI	AP	All directions	All directions	LR	SI	AP
1	2.6	6.4	6.5	±0.8	±2.4	±1.0	±1.7	±2.3	±2.3	±0.6	±5.4	±6.0	±6.4	±6.0
2	5.6	15.2	5.9	**±3.2**	±1.7	±3.1	**±2.3**	**±5.6**	±2.4	±0.6	±5.4	**±6.1**	**±9.2**	±6.0
3	2	4.6	1.2	±0.9	±1.5	±2.7	±1.0	±1.8	±0.8	±0.6	±5.4	±5.8	±6.3	±5.7
4	2.2	4.1	3.8	±2.5	±1.3	±2.1	±1.5	±1.6	±2.0	±0.6	±5.4	±5.9	±6.2	±5.9
5	**6.8**	5.5	7.5	±1.5	±1.1	±1.9	±1.6	±3.1	±2.3	±0.6	±5.4	±5.9	±6.9	±6.1
6	1.9	**18.3**	5.3	±1.6	±2.5	**±4.0**	±2.0	±3.6	±1.9	±0.6	±5.4	±6.1	±7.7	±6.0
7	2.3	4	3.1	±1.5	±1.5	±1.2	±1.4	±1.2	±1.6	±0.6	±5.4	±5.9	±6.1	±5.9
8	1.3	3.5	4.3	±1.0	±0.9	±1.1	±1.7	±1.3	±2.2	±0.6	±5.4	±5.9	±6.2	±6.0
9	4.5	12.9	2.3	±1.2	±1.2	±2.5	±2.1	±3.2	±1.5	±0.6	±5.4	±6.0	±6.7	±5.8
10[†]	5.3	11.5	5.2	±0.7	**±7.3**	±3.5	±2.0	±1.3	±2.0	±0.6	±5.4	±6.0	±7.5	±5.9
11	2.8	14.3	4.1	±0.6	±1.9	±1.2	±1.7	±3.5	±2.0	±0.6	±5.4	±6.1	±6.2	±5.9
12	0.9	12.1	3.8	±1.4	±0.7	±2.0	±1.1	±1.3	±1.7	±0.6	±5.4	±5.8	±6.2	±5.8
13	2.6	1.9	6.1	±0.5	±4.1	±2.5	±2.0	±0.7	±2.0	±0.6	±5.4	±6.0	±6.1	±5.9
14	4.1	6.3	3.7	±0.8	±0.5	±0.8	±2.1	±1.4	±2.2	±0.6	±5.4	±6.1	±6.2	±6.0
15	2.6	8.1	**10.4**	±1.5	±2.2	±1.0	±2.0	±2.2	**±2.5**	±0.6	±5.4	±6.0	±6.4	**±6.1**
16	1.1	2.7	3.4	±1.0	±1.5	±1.0	±1.2	±2.1	±1.6	±0.6	±5.4	±6.0	±6.4	±6.0
17	2.7	3.1	1.9	±1.1	±1.7	±1.2	±1.1	±2.2	±1.5	±0.6	±5.4	±6.0	±6.4	±6.0

[†] Patient 10's 4D-CT showed significant artefacts that may explain large surrogate-target positional uncertainty

Relationship between the beam-surrogate tracking uncertainties and the average surrogate peak-to-trough distance
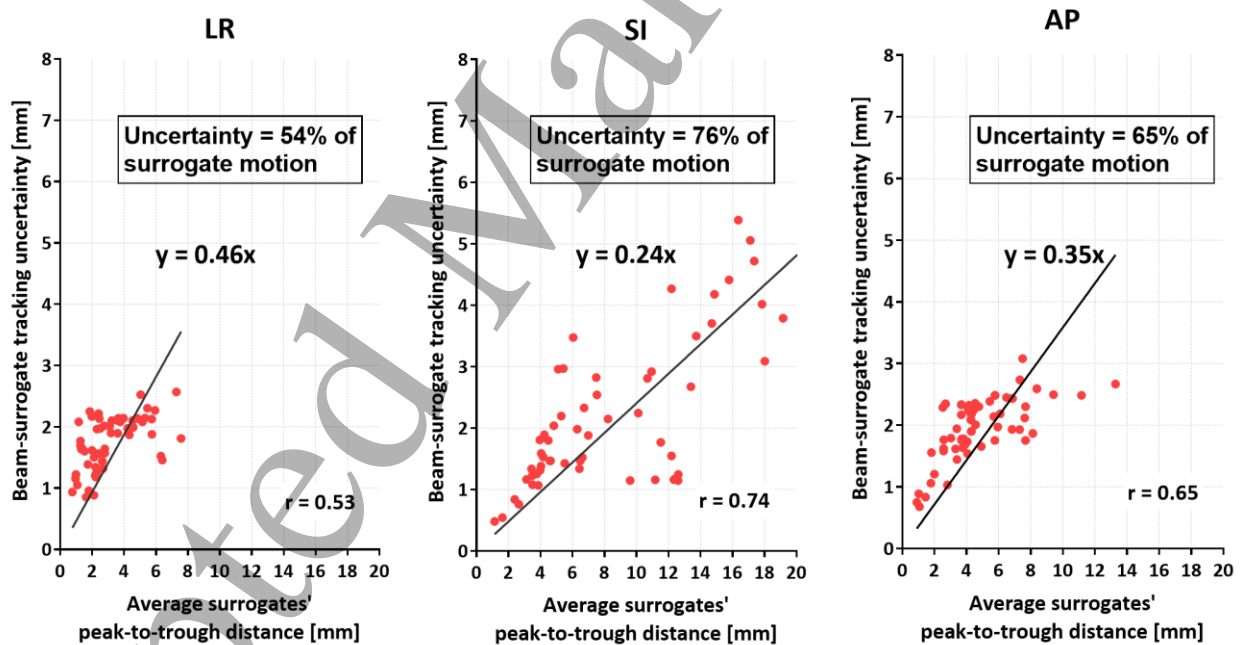
Beam-surrogate tracking uncertainties during treatment were plotted against the average surrogate peak-to-trough distance, as shown in **Figure 3**. Correlation using the Pearson coefficient *r* was found to be the lowest LR (r=0.53) and AP (r=0.65) directions and the highest in SI (r=0.74) directions.

In the LR and AP directions, the line of best fit shows a slope of 0.46 and 0.35 respectively, meaning that the beam-surrogate tracking uncertainty represents 54% and 65% of the average surrogate peak-to-trough motion. It is worth noting that most of the error for the surrogate-target

positional uncertainty in AP are plateauing at ~2.5 mm which corresponds to half a leaf width. This implies that the major contributor to the beam-surrogate tracking uncertainty is probably the leaf width, and this also explains the low Pearson correlation factor. This pattern is not as noticeable in the LR direction probably because most patient's average surrogate peak-to-trough distances were less than 2.5 mm.

In the SI direction, MLC tracking reduced by 76% the average surrogates' peak-to-trough distance. In the SI direction for patients with SI dominant motion, the magnitude of uncertainties is due to a combination of system's latency and leaf speed uncertainties, detailed further in the following section.

Previous work published by Steiner *et al.*[17] in our group compared the delivered and planned average surrogates' peak-to-trough and found that motion during 4D-CT consistently underestimates the true tumor motion during treatment.

Figure 3. Beam-surrogate tracking uncertainty (95% CI) plotted against the average surrogate peak-to-trough distance for 17 patients (70 fractions). Line of best fit and the correlation factor *r* are shown in bold.

Discussion

The purpose of this work was to compute the total geometric uncertainty for a cohort of patients treated with MLC tracking for lung SABR. Using the data collected from the clinical trial, the convolution approach developed by Stroom *et al.* [3] and Sawkey *et al.* [4] enabled the quantification of a set of uncertainties for the cohort and each patient. The overall geometric uncertainty is within ±6.0 mm in LR and AP directions and ±6.7 mm in SI. Individual analysis of uncertainties shows that the surrogate-target positional uncertainty was less than ±3.0 mm and the beam-surrogate tracking uncertainty was accountable for uncertainties equal to or less than ±2.8 mm. The largest uncertainty stemmed from the target delineation uncertainty (±5.4 mm) and the smallest from surrogate target localization uncertainty (±0.6 mm).

By treating the surrogate-target uncertainty and the beam-surrogate uncertainty without splitting the random and systematic error means that the random errors are potentially over-estimated. This means that the overall final margin is computed conservatively. However, we would like to emphasize that this study was designed to provide geometric uncertainty information for the GTV and does not deal with dosimetric margins. A finding from this study, which has been discussed in other studies [4, 18, 19], is that the errors are not normally distributed as seen in Figure 2 with the reported KS p-value. In these cases, to compute the dosimetric margins either an approximation needs to be made to assume normality, or the convolution approach used here would need to be used to incorporate other sources of uncertainty in order to create a margin, such as microscopic spread and beam penumbra required for GTV to CTV and CTV to PTV margins. With a convolution approach, the conventional concepts of a group mean, systematic error and random error do not apply.

Various studies have reported on the performance of other radiation therapy tracking devices comparable to the MLC tracking technology for lung treatment. Floriano *et al.* [20] combined the uncertainties using summation in quadrature and found that 95% of the CTV is geometrically covered with ±5.0 mm. It is worth noting that substantial differences can be found in this work compared with ours. Floriano *et al.* accounted for the deformation of the tumor in their analysis (margins ranging from 1.5 to 2.5 mm) but did not include the target delineation uncertainty that is responsible for the largest uncertainty in our model.

Similarly for the MHI Vero 4DRT, Depuydt *et al*. [21] calculated the 95% CI for the equivalent of the beam-surrogate tracking in the beam's-eye-view to be 3.9 mm on average. These uncertainties were larger than those from this study with calculated 95% CIs to be ±2.7 mm in the direction parallel to the MLC leaves and ±2.1 mm perpendicular to the leaves.

The uncertainties evaluated in this study were from an in-house developed version of MLC tracking that was an adaptation of existing technology and not a commercially designed system. Therefore, some of the geometric uncertainties measured here should be considered as upper bounds of those expected from a dedicated MLC tracking system.

The surrogate-target positional uncertainty could be reduced by tracking the target without implanted markers. Markerless tracking for lung treatment has been tested on fluoroscopic images on patients by Yang *et al.* [22] with the CyberKnife Xsight lung tracking software segmentation with reported segmentation errors of 0.38 ±0.54 mm, noticeably smaller than the surrogate-target positional uncertainty described in this paper (~2.5 mm). Mueller *et al.* [23] tested online real-time markerless tracking software on a Varian TrueBeam linear accelerator with a moving phantom and reported errors of 0.4-3.2 mm (LR), 0.7-1.6 mm (SI) and 0.8-1.5 mm (AP). A common issue described by both Yang *et al.* and Mueller *et al.* is that for patients with a tumor located adjacent to other organs, the lack of tumor contrast with fluoroscopic images increases the segmentation uncertainty and, in some cases, may restrict the patient selection for treatment. However, with the emergence of integrated MRI-linear accelerators [24] capable of MLC tracking as shown in a proof-of-concept in Glitzner *et al.* [25], high-contrast images that could be used to track and adapt to the target position in real-time.

Improvement in the beam-surrogate uncertainty is achievable by reducing the latency of the system or using thinner leaves. Reduction in the system's latency would facilitate predicting the surrogate position ahead of time and in return reduce the beam-surrogate uncertainties. Faster leaf-fitting algorithms have been tested by Caillet *et al.* [26] and it was found that the speed of the calculation of the new leaf pattern has no effect on the overall system's latency. Reducing the system's latency could be achieved, similarly to the CyberKnife and Vero, by coupling the Calypso motion detection system with a faster input (e.g. thoracic belt, vest) and building a correlation model to help reduce the overall system's latency. Despite using fluoroscopic images for beacon segmentations sporadically (Yang *et al.* reported on fluoroscopic images obtained every 40 seconds), CyberKnife studies [27] report latencies of 115 ms. The impact of system latencies may

be reduced with improved motion prediction algorithms or biofeedback technology. The latter has been shown to reduce irregular motion with the ability to stabilize the patient's internal motion [28]. The impact of latency may also be reduced with compression belt and CPAP devices [29, 30] that dampen the magnitude of the target motion.

Previous work published by Steiner *et al.*[17] relied on ten patients from this cohort to compare the delivered and planned average surrogates' peak-to-trough and found that motion during 4D-CT consistently underestimates the true tumor motion during treatment.

Thinner leaves have been explored by Pommer et *al.* [31] showing that for prostate motion (i.e. small motion) the most limiting factor was the leaf width with consistently better dose distributions for treatments with thinner leaves. Pommer *et al.* conclusion corroborate the data shown in **Figure 3** where the AP graphs show a cluster of errors plateauing at approximately 2.5 mm, corresponding to half of a leaf width. As an alternative to implementing thinner leaves, couch tracking [32] could also help reduce residual uncertainties due to leaf width. Ehrbar *et al.* [33] showed that couch tracking alleviates the impact of the leaf width on the beam-surrogate tracking uncertainty.

The target delineation represents the largest uncertainty in this geometrical model. Studies have shown that this variability between observers can be reduced with adherence to contouring guidelines to harmonize the application of target delineation [34]. The use of multimodality imaging, such as PET/CT or PET/MRI [31], has been shown to improve delineation accuracy. The emergence of automatic or semi-automatic target segmentation [31-33] is also showing promise to reduce the inter-observer variability while also potentially saving time for the clinician [34].

Conclusion

This work reports on the geometric uncertainty of MLC tracking for lung SABR to be on average within ±6.0 mm in LR and AP directions and ±6.7 mm in SI and demonstrates the accuracy of the overall system by accounting for the main sources of uncertainties that occurred during treatment. This geometric analysis helps putting into context the range of uncertainties that may be expected during MLC tracking for lung SABR.

Acknowledgement

Conflicts of interest

PJK is an inventor on the awarded US patents 7,469,035 and 8,971,489 that are related to MLC tracking. Patent #7,469,035 is unlicensed, patient #8,971,489 has been licensed by the University of Sydney to Leo Cancer Care. Paul J Keall and Jeremy Booth are investigators on one completed and two ongoing MLC tracking clinical trials that have been partially supported by Varian Medical Systems. Paul J Keall is an inventor on one licensed patent and one unlicensed patent related to MLC tracking. Paul J Keall acknowledges funding from an Australian Government NHMRC Senior Principal Research. Nicholas Hardcastle receives funding from Varian Medical Systems for unrelated research.

Ethical Statement

Data presented in this work was collected as part of an ethics-approved clinical trial (ClinicalTrials.gov registration number: NCT02514512) and approved by the Northern Sydney Local Health District's Human Research Ethics Committee (reference number HREC/15/HAWKE/55). All patients provided informed consent for participation in the study and publication of the results. All investigations were conducted in accordance with the principles embodied in the Declaration of Helsinki and in accordance with the Australian Therapeutic Goods Administration (TGA).

References

1. Molitoris, J.K., et al., *Advances in the use of motion management and image guidance in radiation therapy treatment for lung cancer.* Journal of thoracic disease, 2018. **10**(Suppl 21): p. S2437.

2. Brandner, E.D., et al., *Motion management strategies and technical issues associated with stereotactic body radiotherapy of thoracic and upper abdominal tumors: a review from NRG oncology.* Medical physics, 2017. **44**(6): p. 2595-2612.

3. Stroom, J.C., et al., *Inclusion of geometrical uncertainties in radiotherapy treatment planning by means of coverage probability.* International Journal of Radiation Oncology• Biology• Physics, 1999. **43**(4): p. 905-919.

4. Sawkey, D., M. Svatos, and C. Zankowski, *Evaluation of motion management strategies based on required margins.* Physics in Medicine & Biology, 2012. **57**(20): p. 6347.

5. Venkat, R.B., et al., *Development and preliminary evaluation of a prototype audiovisual biofeedback device incorporating a patient-specific guiding waveform.* Physics in Medicine & Biology, 2008. **53**(11): p. N197.

6. Thomas, H.M., et al., *Impact of tumour motion compensation and delineation methods on FDG PET‑based dose painting plan quality for NSCLC radiation therapy.* Journal of medical imaging and radiation oncology, 2018. **62**(1): p. 81-90.

7. Keall, P.J., et al., *The first clinical implementation of electromagnetic transponder‑guided MLC tracking.* Medical physics, 2014. **41**(2).

8. Ruan, D. and P. Keall, *Online prediction of respiratory motion: multidimensional processing with low-dimensional feature learning.* Physics in Medicine & Biology, 2010. **55**(11): p. 3011.

9. Cerviño, L.I., et al., *The diaphragm as an anatomic surrogate for lung tumor motion.* Physics in Medicine & Biology, 2009. **54**(11): p. 3529.

10. Agnew, C., et al., *Implementation of phantom-less IMRT delivery verification using Varian DynaLog files and R/V output.* Physics in Medicine & Biology, 2012. **57**(21): p. 6761.

11. Guizar-Sicairos, M., S.T. Thurman, and J.R. Fienup, *Efficient subpixel image registration algorithms.* Optics letters, 2008. **33**(2): p. 156-158.

12. Kry, S.F., et al., *AAPM TG 158: Measurement and calculation of doses outside the treated volume from external‑beam radiation therapy.* Medical physics, 2017. **44**(10): p. e391-e429.

13. Balter, J.M., et al., *Accuracy of a wireless localization system for radiotherapy.* International Journal of Radiation Oncology* Biology* Physics, 2005. **61**(3): p. 933-937.

14. Murphy, M.J., et al., *The effect of transponder motion on the accuracy of the Calypso Electromagnetic localization system.* International Journal of Radiation Oncology* Biology* Physics, 2008. **72**(1): p. 295-299.

15. Peulen, H., et al., *Target delineation variability and corresponding margins of peripheral early stage NSCLC treated with stereotactic body radiotherapy.* Radiotherapy and Oncology, 2015. **114**(3): p. 361-366.

16. Mercieca, S., et al., *Interobserver variability in the delineation of the primary lung cancer and lymph nodes on different four-dimensional computed tomography reconstructions.* Radiotherapy and Oncology, 2017.

17.   Steiner, E., et al., *Both four-dimensional computed tomography and four-dimensional cone beam computed tomography under-predict lung target motion during radiotherapy.* Radiotherapy and Oncology, 2019. **135**: p. 65-73.

18.   van Herk, M., et al., *Biologic and physical fractionation effects of random geometric errors.* International Journal of Radiation Oncology* Biology* Physics, 2003. **57**(5): p. 1460-1471.

19.   Zhang, Q., et al., *Three dimensional expansion of margins for single-fraction treatments: Stereotactic radiosurgery brain cases.* arXiv preprint arXiv:1205.1747, 2012.

20.   Floriano, A., et al., *Retrospective evaluation of CTV to PTV margins using CyberKnife in patients with thoracic tumors.* Journal of applied clinical medical physics, 2014. **15**(6): p. 59-72.

21.   Depuydt, T., et al., *Treating patients with real-time tumor tracking using the Vero gimbaled linac system: implementation and first review.* Radiotherapy and Oncology, 2014. **112**(3): p. 343-351.

22.   Yang, Z.-Y., et al., *Target margin design for real-time lung tumor tracking stereotactic body radiation therapy using CyberKnife Xsight Lung Tracking System.* Scientific Reports, 2017. **7**(1): p. 10826.

23.   Mueller, M., et al., *The first prospective implementation of markerless lung target tracking in an experimental quality assurance procedure on a standard linear accelerator.* Physics in Medicine & Biology, 2019.

24.   Raaymakers, B., et al., *First patients treated with a 1.5 T MRI-Linac: clinical proof of concept of a high-precision, high-field MRI guided radiotherapy treatment.* Physics in Medicine & Biology, 2017. **62**(23): p. L41.

25.   Glitzner, M., et al., *MLC-tracking performance on the Elekta unity MRI-linac.* Physics in Medicine & Biology, 2019.

26.   Caillet, V., et al., *In silico and experimental evaluation of two leaf‐fitting algorithms for MLC tracking based on exposure error and plan complexity.* Medical physics, 2019. **46**(4): p. 1814-1820.

27.   Pepin, E.W., et al., *Correlation and prediction uncertainties in the cyberknife synchrony respiratory tracking system.* Medical physics, 2011. **38**(7): p. 4036-4044.

28.   Lee, D., et al., *Audiovisual biofeedback improves the correlation between internal/external surrogate motion and lung tumor motion.* Medical physics, 2018. **45**(3): p. 1009-1017.

29.   Goldstein, J.D., et al., *Continuous positive airway pressure for motion management in stereotactic body radiation therapy to the lung: A controlled pilot study.* International Journal of Radiation Oncology* Biology* Physics, 2015. **93**(2): p. 391-399.

30.   T Eade, J.B., P Keall. *Lung Cancer Radiotherapy Using Realtime Dynamic Multileaf Collimator (MLC) Adaptation And Radiofrequency Tracking (LIGHTSABR).* 2015; Available from: https://clinicaltrials.gov/ct2/show/record/NCT02514512.

31.   Pommer, T., et al., *The impact of leaf width and plan complexity on DMLC tracking of prostate intensity modulated arc therapy.* Medical physics, 2013. **40**(11): p. 111717.

32.   Jöhl, A., et al., *The ideal couch tracking system—Requirements and evaluation of current systems.* Journal of applied clinical medical physics, 2019.

33.   Ehrbar, S., et al., *Validation of dynamic treatment‐couch tracking for prostate SBRT.* Medical physics, 2017. **44**(6): p. 2466-2477.

34.   Oar, A., et al., *Australasian Gastrointestinal Trials Group (AGITG) and Trans-Tasman Radiation Oncology Group (TROG) Guidelines for Pancreatic Stereotactic Body Radiation Therapy (SBRT).* Practical radiation oncology, 2019.