Radiotherapy and Oncology xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Original article

Benefit of using motion compensated reconstructions for reducing inter-observer and intra-observer contouring variation for organs at risk in lung cancer patients

A. McWilliam [a,b,*], L. Lee [c], M. Harris [c], H. Sheikh [c], L. Pemberton [c], C. Faivre-Finn [a,b,1], M. van Herk [a,b,1]

[a] Division of Cancer Sciences, School of Medical Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic Health Science Centre; [b] Department of Radiotherapy Related Research, The Christie NHS Foundation Trust; and [c] Department of Clinical Oncology, The Christie NHS Foundation Trust, Manchester, UK

ARTICLE INFO

Article history:
Received 31 July 2017
Received in revised form 8 November 2017
Accepted 22 November 2017
Available online xxxx

Keywords:
Lung cancer
4D CT scan reconstruction
Inter-observer contour delineation
Intra-observer contour delineation

ABSTRACT

Background and purpose: In lung cancer patients, accuracy in contouring is hampered by image artefacts introduced by respiratory motion. With the widespread introduction of 4DCT there is additional uncertainty caused by the use of different reconstruction techniques which will influence contour definition. This work aims to assess both inter- and intra-observer contour variation on average and motion compensated (mid-position) reconstructions.

Material and methods: Eight early stage non-small cell lung cancer patients that received 4DCT were selected and these scans were reconstructed as average and motion compensated datasets. 5 observers contoured the organs at risk (trachea, oesophagus, proximal bronchial tree, heart and brachial plexus) for each patient and each reconstruction. Contours were compared against a STAPLE volume with distance to agreement metrics. Intra-observer variation was assessed by redelineation after 4 months.

Results: The inter-observer variation was significantly smaller using the motion compensated datasets for the trachea ($p = 0.006$) and proximal bronchial tree ($p = 0.004$). For intra-observer variation, a reduction in contour variation was seen across all organs at risk in using motion compensated reconstructions.

Conclusions: This work shows that there is benefit in using motion compensated reconstructions for reducing both inter-observer and intra-observer contouring variations for organs at risk in lung cancer patients.

Treatment outcomes for lung cancer survival are poor with less than 20% of patients surviving over 5 years [1,2]. Dose escalation studies have shown promise in improving outcomes, although recent studies suggest that the limits of this approach have been reached with standard dose fractionation. The RTOG 0617 trial showed worse outcome in the dose escalation arm [3] with the multivariate analysis indicating that dose to organs at risk (OARs) such as heart and lung was associated with poorer patient survival. However, there remains uncertainty in the reporting of dose statistics to OARs mainly due to variation in contouring. A secondary analysis of heart contours in the RTOG0617 showed large variation across observers, creating uncertainty in the dose delivered to OARs [4].

Contouring studies in tumour delineation for lung cancer patients have shown large variability, particularly in contouring lymph nodes with standard deviations of up to 1.5 cm [5]. The introduction of PET has significantly decreased inter-observer uncertainty [6], potentially allowing smaller target margins. However, uncertainty remains due to the respiratory motion in the lungs. 4-Dimensional computed tomography (4DCT) scans are now standard for radiotherapy planning for the majority of lung cancer patients. These allow the capture of the tumour motion and the potential to reduce or to personalise margins. The simplest approach is to contour a motion-adapted GTV that encompasses the extent of the tumour motion. An alternative is the mid-ventilation approach, where the phase closest to the mid-position of the respiratory cycle is selected for treatment planning. More recently, Wolthaus et al. introduced the mid-position concept where all anatomy is deformed to the true mid-position of the respiratory cycle, i.e. a motion compensated (MC) reconstruction is made [7]. This approach allows patient margins personalised to an individual's respiratory amplitude, which in most cases produces margins smaller than a motion-adapted GTV method [8]. This approach also results in better contrast in the scans as all phases are deformed to the same position and

* Corresponding author at: The Christie NHS Foundation Trust, Radiotherapy Related Research (Dept 58), Wilmslow Road, Manchester M20 4BK, UK.
E-mail address: alan.mcwilliam@manchester.ac.uk (A. McWilliam).
[1] Joint last authors.

https://doi.org/10.1016/j.radonc.2017.11.021
0167-8140/

averaged, reducing noise. Motion artefacts due to irregular respiration are also suppressed. The clinical use of this methodology has been shown to be acceptable with patient outcomes independent of motion amplitude [9].

The respiratory motion that results in uncertainty in delineation of the tumour volume will also cause uncertainty in the delineation of OARs. In current practice OARs are often contoured on a reconstructed average dataset (AVG), which is considered closest to the mean position, but where the respiratory motion causes image blur. With a move towards more complex conformal treatments, dose escalation and the wider adoption of stereotactic ablative body radiotherapy (SABR), accuracy in OAR delineation becomes more critical. There are also increasing numbers of structures to be outlined, with a number located close to the diaphragm where respiration will cause greater uncertainty, and in the mediastinum where signal-to-noise ratio is low and motion compensation can have additional benefit because effectively all dose in the 4DCT scan is used.

As for tumour delineation, implementation of MC workflows should allow for less observer variation in the contouring of OARs when compared to standard approach using an AVG reconstruction. To our knowledge, there are no current inter- or intra-observer delineation studies for OARs in lung cancer patients. This paper performs such a study for the first time with contours compared between AVG and MC reconstructions.

Materials and methods

Eight representative patients diagnosed with early or locally advanced non-small stage lung cancer and treated with SABR were randomly selected. Each patient received a 4DCT scan from which an AVG scan was reconstructed for use in the planning process. The 4DCT scan consisted of 10 phase bins and was acquired by a Philips Big Bore CT scanner (Philips Healthcare). The phase bins were used in the creation of the motion compensated scans utilising ADMIRE (Elekta AB, Stockholm, Sweden) and a Lua script running on a Conquest DICOM server. Additional image handling was done with tools from the Nifty deformable registration package (Nifty, UCLH).

The following steps were used to create the motion compensated scans.

- ADMIRE was used to deformable register each individual phase to a reference phase and export the DICOM deformation vector field (DVF) to the conquest DICOM server. The reference phase was chosen to be at exhale to minimise effects of motion induced image artefacts on the image registration.
- Each individual phase dataset and DVF was loaded into the Lua script. This calculated the mean DVF which was then subtracted from each individual DVF. These modified DVFs were used to deform their associated dataset to the mid-position.
- The four datasets that show the fastest motion (e.g. those during the inhale and exhale slopes) were discarded to reduce motion artefacts.
- The remaining deformed phase datasets were averaged to create the motion compensated dataset and exported.

For each patient the original AVG scan and the MC scan were loaded into Pinnacle vr9.8 (Philips Radiation oncology systems, Fitchburg, WI). Scans were blinded so the observers did not know which AVG and MC scan belonged to the same patient. Five clinical oncologists specialised in thoracic malignancies delineated the OARs on each scan; trachea, oesophagus, proximal bronchial tree (PBT), heart and brachial plexus. The oesophagus, heart and brachial plexus were contoured on the mediastinal level and window. The trachea and PBT were contoured using both the mediastinal and lung level and window. OARs were delineated as described in the United Kingdom SABR consortium guidelines [10]. Structure sets were exported as a DICOM RTSTRUCT object for analysis in ADMIRE.

Inter-patient analysis was performed by first creating a STAPLE (Simultaneous Truth and Performance Level Estimation) volume [11] from the five oncologist contours, for each OAR on each patient. All individual oncologist contours were compared against the STAPLE volume, calculating the unsigned mean and max distance to agreement (DTA). The mean DTA (mDTA) provided a good comparison across the whole volume while the max DTA will indicate the presence of outliers. This may highlight any differences seen from MC scans resulting in sharper boundaries, i.e. between the heart and liver. Results were combined for all observers on the AVG versus MC reconstructions and across all patients for each structure. A pairwise Student *t*-test was used to test for statistical significance, we considered each OAR individually and compared the distribution of the mDTA, averaged for all observers, on AVG and MC reconstructions for each patient. Secondly, contours were cropped so that each structure started and finished on the same CT slices across all observers for each patient. This analysis will enhance intra-slice differences, highlighting improvements between the two reconstruction techniques.

Intra-observer analysis was also performed, each observer was allocated one patient to re-contour, after a minimum delay of four months. Observers re-contoured the same patient for both the AVG and MC scans allowing a direct measure of the intra-observer variation on both reconstruction techniques. ADMIRE was used to calculate unsigned mDTA between the two sets of contours. A direct comparison was performed of the variation between the AVG and MC contouring for each observer and each structure across all patients, a pairwise Student *t*-test was used to test statistical significance.

Finally, in moving to MC reconstructions, we may find that contours report lower volumes. In removing blurring caused by the respiratory cycle, and increasing contrast, contours may become smaller. Therefore, any associated, volume based dose statistics reported will show differences from using the AVG. Volumes of each structure were compared on the AVG and MC to investigate if this effect is present.

Results

Fig. 1 shows examples of the MC reconstructions, slices highlighting the improved definition of the bronchial tree and greater clarity of the boundary between the heart and the liver.

Results show a significant improvement in using the MC scans for contouring the trachea (paired *t*-test, $p = 0.04$) and some benefit for the remaining OARs, Fig. 2. After editing to remove the uncertainty of the superior and inferior extent of the OAR contours the trachea remains significantly improved ($p = 0.006$), but also the PBT ($p = 0.004$) becomes significant. There is an overall improvement in the mDTA, particularly for the trachea, PBT and oesophagus, OARs which are tubular in nature, are now sub-mm. Figs. 2 and 3 also show the SD across the observer results as the included error bars. The MC reconstructions show a smaller variation compared to the AVG, particularly for edited contours, indicating improved inter-observer agreement. It is worth noting that the brachial plexus showed a significant improvement in Fig. 2. However, there remained a large variation between observers (mDTA of 3.0 cm).

The heart shows little change in mDTA values, both between MC and AVG. The heart is a large organ, with a large semi-vertical border with the lung that is hardly affected by breathing motion. This boundary will not show much benefit from the MC reconstruction and will drive the mDTA results. There may be some

Please cite this article in press as: McWilliam A et al. Benefit of using motion compensated reconstructions for reducing inter-observer and intra-observer contouring variation for organs at risk in lung cancer patients. Radiother Oncol (2017), https://doi.org/10.1016/j.radonc.2017.11.021