Peer review file

Reviewer A:
Comments:

General Reaction: The authors present a nicely crafted review on motion management strategies in lung cancer. The review covers elements of motion management, but in general I found it to be lacking in depth. The organization could use reworking, with more of a focus early on defining the need scope of the problem. How much motion is often seen? What are the resultant toxicities we are trying to mitigate? How prevalent are they? What does the literature from institutions who have implemented these measures suggest in terms of toxicity mitigation?

ANSWER: *we thank you for your comments, which contributed to increase the quality of the manuscript. We reorganized the manuscript according to your suggestions (Introduction: page 3, lines 60-70; Discussion page 4, lines 102-105, page 9, lines 293-298).*

The following comments are offered in the spirit of strengthening the manuscript.

Major points:

I would define the audience. Is this intended for physicians? If so much more info is needed on how to choose between these strategies, and how to implement them in ones clinic. Is it for physicists? If so then much more information on the determination of margins, uncertainty etc is needed.

A review of this nature would be strengthened by figures, either crafted from the clinic of the author or copied with permission from some of the referenced articles.

ANSWER: *it was indeed one of the limits of the manuscript. We re-addressed the audience (physicians) of the manuscript by discussing the clinical results presented in few studies (page 4, line 122-128; page 5, line 144-148; page 9, line 290-298)*

Minor points:
Lines 50-60--I would recommend adding more details re: magnitude of motion here, either from the Seppenwolde reference or likely to also include a few others. This is essentially the paragraph defining "how big of a problem is this?" for the uninitiated reader, and they're left without a real quantitative understanding.

ANSWER: *thank you for your remark. We added the requested data (page 4, line 102-105) and a new reference (page 4, line 122-128).*

Lines 97-104. This nicely breaks down ITV and what is then (line 98) called setup margins. However on line 102 "geometrical uncertainties" are referenced as warranting ITV to PTV expansion. This uncertainty is the "set up margins" but the inconsistency in the verbiage makes that unclear.

I would clarify that these are the same, and highlight the need for daily imaging (usually CBCT, +/- cine tracking at time of treatment setup) to likewise minimize the ITV to PTV marginal expansion. Could build this section out a bit more.

ANSWER: *we corrected the sentence and discussed a bit more the section (page 6, line 177-189).*

Likewise lines 106-117--I think you are describing how we as a field roughly estimate ITV to PTV margins--but this is all very opaque. "To calculate the CTV to PTV margin…" reads as if there is going to be some recipe or instruction, yet it then references "systematic errors" and other factors without describing who measures these, what they are, how they are corrected for. Most RTOG protocols for SBRT use an ITV to PTV expansion of 5 mm--based on your arguments here, are these good expansions? How did we as a field settle on 5 mm? Some concrete examples and sample calculations for (the MD? the physicist?) would be more instructive here.

ANSWER: *thank you for your remarks. We reorganized the discussion of the paper and explained the 5mm threshold in the introduction section (page 1, lines 65-70).*

Lines 172-181--to my knowledge most early CK systems used fiducials, yet this is not clear or described as invasive in this section...yet is followed by lines 180-1 which states you can now do this without fiducials. Please explain the role of invasive fiducials in early CK series more thoroughly.

ANSWER: *indeed, few early series used fiducials (even if early series sometimes do not report it). We detailed the CK tracking system and explained the limits of the use of fiducials in page 8, lines 274-278, page 9, lines 281-298.*

Lines 182-187 needs more explanation.

ANSWER: *we added an explanation in the text (page 8, line 269-270).*

Lines 200-206--this seems more like background/introductory info defining the scope of the problem. What proportion of patients meet these criteria?
Lines 209-214--again, this is background information that would be best introduced early.

ANSWER: *indeed it was more a background information. We moved it up to the background section (page 1, line 60-70).*

Reviewer B:
Comments:

This manuscript reviews the respiratory motion impact on radiotherapy for lung cancer. The aspects of motion quantification, mitigation and management techniques are described. I have following comments and suggestions.

A discussion how respiratory motion is managed during treatment planning, optimization and quality assurance is lacking:

-A lot of work has been done introducing robust treatment planning strategies, mostly for proton therapy, but also in photon-based radiotherapy. For instance studies on optimal beam angle selection towards breathing motion (Casares-Magaz et al. Acta Oncol 2014), and on optimization routines in home-written planning software as well as in the RayStation software. These routines allow the input of data on setup (and range) uncertainty together with individual 4DCT frames, in order to optimize treatment plans which are robust against all these uncertainties.

-The final plan robustness evaluation should also be discussed, with for instance the 4D dose accumulation technique (Ribeiro et al. Radiother Oncol 2019, Souris et al. Medical Physics 2019, etc.).

-In terms of quality assurance, the potential interplay effect occurring due to breathing and delivery timings not being synchronized, could deteriorate dose delivery and should be checked (e.g. Riley et al. Medical Physics 2014).
ANSWER*: thank you for your comments. We deliberately didn't explore the motion management in proton treatment, but we added a sentence and few references in the "Main remaining challenges" section (page 10, lines 429-432).*

Minor comments:
p.3 line 40 'possibility' should be 'possibly'?
ANSWER*: it was indeed, we corrected it in the text (page 3, line 59)*

p.3 line 53 Consider giving some typical numbers of average amplitude and variability observed.
ANSWER*: thank you for your remark. We added the requested data (page 4, line 102-105).*
p.4 Also relevant to mention is the tumour drift often observed during a treatment fraction, see Guckenberger et al. Radiother Oncol 2007
ANSWER*: We added the interesting reference (page 4, 122-128).*

p.4 line 89 There is a fundamental difference between the midventilation and midposition CT approach which should be explained. Another approach could be added: the average CT image which can also be used as planning CT scan.
ANSWER*: thank you for your remark, we reshaped the sentence and added the reference (page 5, line 149-154).*

p.4 line 90 I don't see how the discussion on beam penumbra immediately follows from the midventilation CT explanation. This seems more in line with the penumbra discussion on the next page.
ANSWER*: indeed. We eliminated it because it was redundant with the information on penumbra on the next page.*

p.5 line 105 A study evaluating the underprediction of motion amplitude by 4DCT could be added (Steiner et al. Radiother Oncol 2019).
ANSWER: *We added the interesting reference (page 5, line 144-148).*

p.5 line 111 or 'ITV to PTV'?
ANSWER: *We rephrased the sentence (page 6, line 176).*