Peer Review File

Article information: https://dx.doi.org/10.21037/jtd-23-133

Reviewer A

The authors presented a work to develop and validate a CT-based delta-radiomics model for discriminating synchronous double primary lung cancer from intrapulmonary metastasis using a radiomics approach. The article is well-written and organized, in general. Some issues should be addressed before being recommended for publication.
Comment 1: First, I have concerns about using the term "delta-radiomics" for analysis different from what "delta-radiomics" defines in the literature.

Reply 1: We highly appreciate the Reviewer's constructive comments. We have replaced the term "delta-radiomics" with the "difference of radiomics" throughout the revised manuscript.

Comment 2: Second, there is a lack of explanation of the radiomic feature extraction, including either missing or inconsistent information provided.

Reply 2: We sincerely thank the Reviewer's for the comments. We have supplemented the relevant information of radiomic feature extraction.
Changes in the text: *Heatmap of the 1037 radiomics features by Image Type:93 Gradient features 93 LoG features 107 original features 744 wavelet features By Feature Class: 198 First-order features 264 GLCM features 154 GLDM features 176 GLRLM features 176 GLSZM features 55 NGTDM features 14 Shape features.*

Comment 3: Finally, although the results show some usefulness of the method, the limitations mentioned in this work, besides some other limitations, hinder the applicability in a clinical setting.

Reply 3: We sincerely thank the Reviewer's for the concerns. We have revised the manuscript according to your comments to make up for the limitations, and the main limitations have been addressed in the discussion section
Changes in the text: *Firstly, as our study required an external validation cohort, it was not possible to accurately evaluate the generalizability and stability of the radiomics model, as well as the reproducibility of radiomic features.*

Introduction:
Comment 4: - 79: TNG acronym must be defined the first time it is used.
- 83: Add space between "System" and "(5)"
- 90: Define what "High-resolution" means in CT, e.g. pixel size, etc.
- 92: Add space between "prognosis" and "(7-9)"
- 93: Add space between "studies" and "(3, 11, 12)"
- 99: Add space between "disease" and "(3)"

Reply 4: We sincerely thank you for kind reminder.
(1) TNM acronym has been defined.
(2) Space between "System" and "(5)" has been added.
(3) High-resolution CT was defined as an examination technique for thin slice (1.0~1.5mm)

scanning and high-resolution algorithm reconstruction of images.
(4) Space between “prognosis” and “(7-9)” has been added.
(5) Space between “studies” and “(3, 11, 12)” has been added.
(6) Space between “disease” and “(3)” has been added.

Comment 5: 109: In general, delta-radiomics refers to the analysis of feature variation at different acquisition time points, usually before and after therapy (https://pubmed.ncbi.nlm.nih.gov/34865190). It has been explored in pre-clinical animal studies to study changes in CT images before and after inoculation (https://pubmed.ncbi.nlm.nih.gov/36506838). Studying the difference in radiomic features between two types of diseases (or between normal and disease groups) does not fall under the category of delta-radiomics. The authors should explain the concept proposed using different terminology.

Reply 5: We highly appreciate the Reviewer’s constructive comments. We have replaced the term "delta-radiomics" with the "difference of radiomics" throughout the revised manuscript.

Comment 6:- 132: Add space between “metastasis” and “(21)”
- 141: Add space between “System” and “(5)”
- 147: Add space between “standard” and “(22)”
- 149: Add space between “metastases” and “(23)”
- 141 and 148: IASLC, ATS, ERS acronyms may need a definition the first time they are used.
- 147: CHA may need a definition
- 157: EGFR and KRAS may need a definition

Reply 6: We sincerely thank you for kind reminder.
(1) Space between “metastasis” and “(21)” has been added.
(2) Space between “System” and “(5)” has been added.
(3) Space between “standard” and “(22)” has been added.
(4) Space between “metastases” and “(23)” has been added.
(5) IASLC, ATS, ERS, CHA, EGFR and KRAS has been defined in the revised manuscript.

Comment 7:- 159: A better title for this section may be “CT image acquisition and reconstruction”. The second paragraph of this section should be moved to the following section.

Reply 7: We have changed the title as “CT image acquisition and reconstruction”, and second paragraph of this section have been moved to the “Image evaluation” section.

Comment 8- 175: This section should contain the second paragraph of the previous section. Besides, I would use a different title for this section, such as “Image evaluation” or something like that.

Reply 8: We have changed the title as “Image evaluation”.

Comment 9- 178: What do the authors mean with “performed all image analysis on DICOM”? It does not sound good.

Reply 9: We sincerely thank you for kind reminder. We have revised the sentence as “*Three radiologists with different degrees of experience (HWR, XJ, and YTZ, and with 15, 5, and 2 years of*

experience in radiology, respectively) were invited to perform independently all image analysis from the testing set without knowing of the clinical or pathological findings."

Comment 10 - 197: Add a space between "." and "The".

Reply 10: Space between "." and "The" has been added.

Comment 11- 199: A fixed bin width of 25 is the default of PyRadiomics. How many bins represent 25 bins is the tumor VOIs? Is that enough for a histogram to exhibit the main characteristics of the distribution? Besides, images have been normalized first. Some additional explanation should be included here.

Reply 11: We highly appreciate the Reviewer's constructive comments. For tumor lesions in general, the difference between maximum and minimum densities ranges from 300 to 700, corresponding to bin values of 300~700/25 +1. Considering the size and quantity of the nodules' voxels, excessive segmentation may introduce unwanted noise. Therefore, a value of 25 was selected for comprehensive consideration.

Comment 12- 202: "direct" refers to "original image", right? It should be rewritten.

Reply 12: We sincerely thank you for kind reminder. The word "direct" has been changed as "original image".

Comment 13- 202: Wavelet filtering uses the whole image, not only the region of interest. How was this addressed for comparability?

Reply 13: We highly appreciate the Reviewer's concern. Wavelet filtering yields 8 decompositions per level, which comprise all possible combinations of applying either a High or a Low pass filter in each of the three dimensions. The resultant image obtained through Wavelet filtering, combined with ROI, can be used for extracting radiomic features.

Comment 14- 202: What is the rationale of using logarithmic transformed images, other than it is included in the PyRadiomics library.

Reply 14: We sincerely apologize for the typo. We have replaced"logarithmic transformed" with LoG throughout the manuscript.
Changes in the text: *For each ROI, 1037 radiomic features, including original image, wavelet transformed, loG features, and gradient filtered features,*

Comment 15- 182 and 201: The authors need to be consistent when using ROI or VOI.
- 200: Images were "reconstructed" or "resampled"?
- 208: Add a space between "lesions" and "(20)"

Reply 15: (1) We have replaced VOI with ROI.
(2) The word "reconstructed" has been replaced with "resampled".
(3) Space between "metastases" and "(23)" has been added.

Comment 16- 208: First, reference (20) is about delta-radiomics defined as a change in radiomic

features between different time points, for example, pre-treatment vs. post-treatment, or pre-infection vs. post-infection. The authors use delta-radiomics differently, which is more related to comparing features between two populations with different diseases. Second, the equation in line (209), which by the way should be numbered, uses the absolute value of the difference between the feature at both lesions. Why the absolute value instead of the difference? I mean, -5 is not the same as 5, for example, but the definition used by the authors is not able to distinguish between them.

Reply 16: We highly appreciate the Reviewer's constructive comments.

(1) We have replaced the term "delta-radiomics" with the "difference of radiomics" throughout the revised manuscript.

(2) The present study focused on the difference of radiomics between the two lesions. To minimize the influence of the serial number of the tumor, we utilized the absolute value to evaluate the difference between Tumor 1 and Tumor 2.

Comment 17- 220: In which way inter-observer variability analysis was used to exclude redundant variables? Inter-observer variability analysis helps identify non-reproducible features, for example.

Reply 17: In the present study, Pearson's correlation analysis and least absolute shrinkage and selection operator (LASSO) penalized logistic regression were used to filter unstable (ICC < 0.75) and redundant features (r > 0.8).

Comment 18: Results:
- 238: Add a space between "3.1" and "Clinicopathological"

Reply 18: Space between "3.1" and "Clinicopathological" has been added.

Comment 19: - 257: 1,037 features were considered, and 332 delta-radiomic features were selected to establish the delta-radiomics model. Some discussion about overfitting should be included in the manuscript.

Reply 19: We sincerely thank you for concerns. The five-fold cross-validation least absolute shrinkage and selection operator (LASSO) analysis was performed on the training data set to overcome the overfitting, which has been explained in the revised manuscript.
Changes in the text: *The five-fold cross-validation LASSO analysis was performed on the training data set to mitigate the risk of overfitting.*

Comment 20:- Figure 2: The "Feature Extraction" box says: 93 gradient features, 93 LoG features, 107 original features, and 744 wavelet features, which is 93 x 8 = 744. Besides, 93 (texture features) + 14 (shape features) = 107 (original features). However, line 202 says: "direct, wavelet transformed, logarithmic transformed, and gradient filtered features". Note that LoG is not the same as "logarithmic transformed".

Reply 20: We sincerely apologize for the typo. We have replaced"logarithmic transformed" with LoG throughout the manuscript.

Comment 21- Also note that the rad-score formula in 268 (which should be numbered too) includes a log.sigma.6.0.mm.3D feature. If LoG features were acquired with different kernel sizes, there should be more than 93 LoG features. If not, the authors should explain why 6 mm was chosen.

Reply 21: In the present study, we selected a kernel size of 6 mm. From a clinical standpoint, surgical intervention is generally not recommended for lesions smaller than 6 mm. Hence, our focus was on extracting radiomic information from images of tumors greater than 6 mm in size.

Comment 22- 260: "The four delta-radiomics features (two first-order 90Percentile, and four second order parameters, including two GLCM and one GLSZM) were identified by the LR model." What does it mean "two first-order 90Percentile"? Does it mean 90Percentile feature extracted from two different images? If so, which are those images?

Reply 22: We sincerely apologize for the typo. We have replaced the "two first-order 90Percentile" with "first-order 90Percentile".
first-order 90Percentile: First-order statistics describe the distribution of voxel intensities within the image region defined by the mask through commonly used and basic metrics. Let **X** be a set of voxels included in the ROI, this feature means the 90th percentile of **X**.

Comment 23- 263: Table 2: explain differences in some features in the train and test cohorts.

Reply 23: We highly appreciate the Reviewer's constructive comments. We have added the interpretation of differences in some features in the train and test cohorts.
Changes in the text: *As presented in Table 2, patients with SDPLA demonstrated significantly higher values for firstorder_90Percentile, _glcm_ClusterShade, and _glcm_Imc2 compared to those with IPM. Conversely, the value of glszm_SizeZoneNonUniformity was significantly lower in patients with SDPLA than in individuals with IPM, across both the training and testing sets..*

Comment 24- Figure 3: It would say (d-f) instead of (d-e). The H&E images showing differences in texture between the two tumors may be misleading. While radiomic features are extracted at a millimeter level, pathology images exhibit texture at a sub-micron level, indistinguishable at a millimeter level (https://doi.org/10.1016/j.ejmp.2020.03.018).

Reply 24: We sincerely apologize for the misleading expression. The difference in pathology referred to the difference in the histological type of the two tumors rather than the difference in the radiomics features. We have modified the expression to make it clearer.
Changes in the text: *(d-f): A 66-year-old male with one primary lung adenocarcinoma in the right upper lobe (Tumor 1) and one metastasis (Tumor 2) in the left lower lobe. Tumor 1 appeared as an irregular mass with spiculate protuberance, while Tumor 2 appeared as a spiculate, solid nodule on CT. HE staining showed a similar histological type of predominant papillary patterns for both tumors.*

Discussion:
Comment 25- 298: Add a space between "carcinoma" and "(7, 14, 15)"

Reply 25: Space between "carcinoma" and "(7, 14, 15)" has been added.

Comment 26- 307: Here, it may be useful to explain why shape features were not included in the analysis.

Reply 26: We highly appreciate the Reviewer's constructive suggestion. We have added the related explanation in the revised manuscript.
Changes in the text: *In the present study, no significant differences were found in shape features such as spiculate, lobulated, and air bronchogram, between groups. This study showed SDPLA to exhibit significant differences in tumor location, Δd, and tumor type compared with IPM.*

Comment 27 Conclusions:
- 342: "need to be verified in further prospective multicenter studies". Here, it is not only about verifying the results. Multicenter studies bring additional issues, such as images acquired with different scanners, acquisition parameters, and resolution. Reproducibility analysis must be performed to decide which features are reproducible before being included in the model for evaluation.

Reply 27: We totally agree with the reviewer's comment, and we have revised the sentence as follows:
Changes in the text: *However, due to the single-center retrospective study design, our conclusions must be verified in prospective multicentric studies with different scanners, acquisition parameters, and resolution.*

Comment 28- 331: I agree with the need to evaluate the generalization and stability of the radiomic model. Furthermore, the reproducibility of radiomic features should be addressed as well. For example, wavelet filtering is performed over the whole image, not just the VOI, therefore, how the image is cropped (or not cropped) may lead to completely different values of the radiomic features extracted from the HHH, HHL, ... , LLL filtered images. It is worth mentioning that two relevant features in the rad-score formula were extracted from those images.

Reply 28: The reproducibility of radiomic features have been addressed in the revised manuscript.

Reviewer B

The authors have written an interesting paper regarding value of delta radiomics in differentiating synchronous double primary lung adenocarcinoma from intrapulmonary metastasis. However, major revisions are needed before publication.

Comment 1: There are lots of writing and grammatical errors throughout the manuscript that make it hard to understand.

Reply 1: We sincerely apologize for the errors in the language and have carefully corrected the writing and grammatical errors throughout the manuscript the revised version.

Comment 2: All abbreviations are needed to be clear. (IASLC? PACS? DICOM?).

Reply 2: The abbreviations have been defined in the revised manuscript.

Comment 3: Method section in abstract should be revised with important information and more details.

Reply 3: We have added more details to Method section in abstract in the revised manuscript.

Comment 4: Results section in abstract should be written with more details (numbers, precents, ...)

Reply 4: We have added more details to Results section in abstract in the revised manuscript.

Comment 5: The abbreviation of SDPLA has defined in different manner in second paragraph of introduction (line 95) compared to other parts.

Reply 5: We sincerely thank you for kind reminder. The abbreviation of SDPLA has defined as synchronous double primary lung cancer in the revised manuscript.

Comment 6: In methods section (patients part- line 126) and Fig. 1, 6114 must be replaced with 6117.

Reply 6: We sincerely thank you for kind reminder. 6114 has be replaced with 6117 methods section (patients part- line 126) and Fig. 1.

Comment 7: It is unclear why patients with multiple lung cancer occurrences separated by more than six months were excluded, given the finding that the time between cancer-free periods was less than two years (lines 152-153).

Reply 7: We sincerely apologize for the inconsistent expression. We have revised the sentence as: "...the time between cancer-free periods was less than 6 month".
Changes in the text: "*and (IV) the interval of tumor-free between cancers was less than six months"*

Comment 8: I suggest writing CHA criteria in a separate subdivision of methods with a subtitle.

Reply 8: We have put "CHA criteria" in a separate subdivision of methods with a subtitle.

Comment 9. The radiomics features used in this study were not described in detail, including the dominant features after eliminating others (lines 181-205).

Reply 9: We have added more details about the radiomics features in the revised manuscript.

Comment 10: Based on Table 1, there is a statistically significant difference between 2 groups in terms of sex. While it is written not in line 241.

Reply 10: We sincerely thank you for kind reminder. IPM occurs more often in men than SDPLA in the testing sets (75% vs. 21.1%, p=0.008)

Comment 11: Line 246, lymph node metastasis and distant metastasis…
Reply 11: We have corrected this typo.
Changes in the text: *Furthermore, lymph node metastasis and distant metastasis were more common in the IPM group than SDPLA in the training set.*

Comment 12: The description of tumor location and tumor type are missed in results section.

Reply 12: The description of tumor location and tumor type have been added in results section.

Changes in the text: *Compared with IPM, SDPLA was more frequently located in different lobes (36.5% vs.60.4%, p=0.015) and different Tumor type (*41.9% vs 10.4%, $p < 0.001$*).*

Comment 13: Information about solid nodules is missed in table 1. (Line 250).

Reply 13: We sincerely apologize for the confusing expression. The information of solid nodules was included in tumor type in Table 1.

Comment 14: Discussion part is written organized.

Reply 14: We sincerely thank the Reviewer's for encouraging feedback.

Comment 15. I suggest to use more recent published articles for references.

Reply 15: The references have been updated with more recent published articles.