Peer Review File

Article information: https://dx.doi.org/10.21037/tcr-21-1037

Reviewer A Comment 1: The authors should add a paragraph to introduce the conventional CT profiles and radiomic features. The authors listed a lot of conventional or radiomic features in Table 1 and 2, which are unfamiliar to most readers. In this paragraph, the authors should address the commonly-used conventional CT profiles, radiomic matrices, and radiomic features.

Reply 1: Thanks for your constructive suggestion. Those comments are of vital importance to us. There are many imaging features used to describe the conventional CT findings of lung adenocarcinoma. Ground glass opacity (GGO) is used to depict lung lesions that appear as hazy regions of increased opacity in the lung with preservation of bronchial and vascular margins, which is an important feature of lung adenocarcinoma[1,2]. There are still some other descriptions, such as bubble-like lucency[3,4], pleural retraction[5], air bronchogram[2]. Radiomic features are divided into two categories: semantic features and agnostic features[6]. Semantic features are those commonly used by radiologists to describe lesions, such as size, shape, location, and necrosis, while agnostic features are those that attempt to capture lesion heterogeneity through quantitative descriptors, such as Histogram (skewness, kurtosis), Haralick textures, Laws textures, Wavelets, Laplacian transforms, Minkowski functionals, and Fractal dimensions[6]. In this new version, we have added these details in corresponding paragraph.

Changes in the text: We have modified our text as advised (see Page3, Line 53-62; Page 4, Line 76-82).

References:

1 Pascoe, H. M., Knipe, H. C., Pascoe, D. & Heinze, S. B. The many faces of lung adenocarcinoma: A pictorial essay. *Journal of medical imaging and radiation oncology* **62**, 654-661, doi:10.1111/1754-9485.12779 (2018).

2 Hansell, D. M. *et al.* Fleischner Society: glossary of terms for thoracic imaging. *Radiology* **246**, 697-722, doi:10.1148/radiol.2462070712 (2008).

3 Gaikwad, A. *et al.* Primary adenocarcinoma of lung: a pictorial review of recent updates. *European journal of radiology* **81**, 4146-4155, doi:10.1016/j.ejrad.2012.08.023 (2012).

4 Yoshida, T. *et al.* Lung adenocarcinoma presenting with enlarged and multiloculated cystic lesions over 2 years. *Respiratory care* **49**, 1522-1524 (2004).

5 Liu, Y. *et al.* CT Features Associated with Epidermal Growth Factor Receptor Mutation Status in Patients with Lung Adenocarcinoma. *Radiology* **280**, 271-280, doi:10.1148/radiol.2016151455 (2016).

6 Gillies, R. J., Kinahan, P. E. & Hricak, H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* **278**, 563-577, doi:10.1148/radiol.2015151169 (2016).

Reviewer A Comment 2: Concerning the perspective or future trend, the authors have to remind the readers about the use of artificial intelligence in this field. For example, the neural network in the classification of gene mutation. Although some of the studies mentioned in this review were conducted by simple machine learning, the use of AI techniques was not stated.

Reply 2: We appreciate the reviewer raising this constructive comment and it will undoubtedly improve our manuscript quality. Artificial intelligence is an irreversible trend of medical development. As a primary part of artificial intelligence, in recent years, radiomics studies have been repeated without new breakthroughs. The author believes that the precise automatic recognition of artificial intelligence will replace the

manual sketch of radiomics, the process that the sketched region of interest submitted to radiomics company for processing will be replaced by one click analysis software, and deep learning, such as the neural network, will replace simple machine learning in the classification of gene mutation in lung adenocarcinoma. Research on the application of conventional CT features and CT image-based radiomics features for predicting the gene mutation status of lung adenocarcinoma is still in a preliminary stage. Although there is some hope that radiomics could be applied in clinical practice, more work needs to be done.

Changes in the text： We have modified our text as advised (see Page 14, Line 310-312; Page 15, Line 313-321).

Reviewer B Comment-Introduction: Please try to address the increasing prevalence of non-smoking related lung cancer in Asian population after the implementation of lung cancer screening program. Therefore, the more evidence about molecular typing of lung adenocarcinomas typing is increasing.

Reply: Thanks for your comments and references. We have amended the introduction part as the reviewer suggested. Many studies have shown that the incidence rate of lung cancer in non-smoking women, and the detection rate of lung cancer, especially lung adenocarcinoma, after the implementation of lung cancer screening program are both increasing year by year[7-12]. Therefore, the research on the molecular mechanism of lung adenocarcinoma at the micro level is further increased.

Changes in the text： We have made extensive modification to the introduction part (see Page 2 , Line 34-38, 43-44).

References:

7 Lin, K. F. *et al.* Propensity score analysis of lung cancer risk in a population with high prevalence of non-smoking related lung cancer. *BMC pulmonary medicine* **17**, 120, doi:10.1186/s12890-017-0465-8 (2017).

8 Detterbeck, F. C. *et al.* The IASLC Lung Cancer Staging Project: Summary of Proposals for Revisions of the Classification of Lung Cancers with Multiple Pulmonary Sites of Involvement in the Forthcoming Eighth Edition of the TNM Classification. *J Thorac Oncol* **11**, 639-650, doi:10.1016/j.jtho.2016.01.024 (2016).

9 Subramanian, J. & Govindan, R. Lung cancer in never smokers: a review. *J Clin Oncol* **25**, 561-570, doi:10.1200/jco.2006.06.8015 (2007).

10 Yu, Y. W. *et al.* [Meta-analysis on related risk factors regarding lung cancer in non-smoking Chinese women]. *Zhonghua liu xing bing xue za zhi = Zhonghua liuxingbingxue zazhi* **37**, 268-272, doi:10.3760/cma.j.issn.0254-6450.2016.02.024 (2016).

11 Wu, F. Z. *et al.* Semiquantative Visual Assessment of Sub-solid Pulmonary Nodules ≦3 cm in Differentiation of Lung Adenocarcinoma Spectrum. *Scientific reports* **7**, 15790, doi:10.1038/s41598-017-16042-9 (2017).

12 Wu, F. Z. *et al.* Assessment of Selection Criteria for Low-Dose Lung Screening CT Among Asian Ethnic Groups in Taiwan: From Mass Screening to Specific Risk-Based Screening for Non-Smoker Lung Cancer. *Clin Lung Cancer* **17**, e45-e56, doi:10.1016/j.cllc.2016.03.004 (2016).

Reviewer B Comment-Method: Please try to address the illustrated figures of radiomic feature and processing for molecular typing of lung adenocarcinomas typing.

Reply: Thanks for your comments. We have amended the part as the reviewer suggested. Radiomic features are divided into two categories: semantic features and agnostic features[6]. Semantic features are those commonly used by radiologists to describe lesions, such as size, shape, location, and necrosis, while agnostic features are those that attempt to capture lesion heterogeneity through quantitative descriptors, such

as Histogram (skewness, kurtosis), Haralick textures, Laws textures, Wavelets, Laplacian transforms, Minkowski functionals, and Fractal dimensions[6].

The process of radiomics for molecular typing of lung adenocarcinomas includes six steps. (1) acquiring the images, (2) identifying the volumes of interest, (3) segmenting the volumes (ie, delineating the borders of the volume with computer-assisted contouring), (4) extracting and qualifying descriptive features from the volume, (5) using these to populate a searchable database, and (6) mining these data to develop classifier models to predict outcomes either alone or in combination with additional information, such as clinical, conventional CT features.

Changes in the text: We have made extensive modification (see Page 3, Line 70; Page 4, Line 71-82).

References:

6 Gillies, R. J., Kinahan, P. E. & Hricak, H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* **278**, 563-577, doi:10.1148/radiol.2015151169 (2016).

Reviewer B Comment-Result and discussion 1: Please try to address the difficulty in the unique vendor setting, CT slice thickness, and different kernel setting in the real-world practice. Therefore, further study with unique CT protocol is warranted.

Reply: Thanks for pointing out this important issue. In the real-world practice, because the CT scanning machine kernel setting and scanning parameters (such as tube current and tube voltage) and the reconstruction algorithm are inconsistent in different studies, the scanning protocol and reconstruction algorithm should be standardized and

generalized. However, there is still no unified CT scanning protocol and image reconstruction standard. We need further study to establish image acquisition and reconstruction protocol standardization.

Changes in the text： We have made extensive modification (see Page 13, Line 281-289).

Reviewer B Comment-Result and discussion 2: Please try to address the heterogeneous study design in studies, and to investigate most important radiomic parameters for molecular typing of lung adenocarcinomas typing. How to clinical application in the real-world practice.

Reply: Thanks for raising these important issues. Most of the studies on gene prediction by CT image-based radiomics are retrospective studies. The heterogeneity of the research mainly lies in the differences of scanning parameters and image reconstruction protocols, and small sample sizes. Ideally, these studies should be combined to provide a meta-analysis, which is what our research team intends to do next.

Among these radiomic parameters, no study has shown which parameter is the most important for predicting the types of gene mutations in lung adenocarcinoma so far.

There are many problems to be solved before applying radiomics for the molecular typing prediction of clinical lung adenocarcinoma. Most of the current studies have a small sample size and lack the verification of multiple center, large sample and prospective experiments. Moreover, radiomic studies lack unified standards

in many aspects, such as scan protocol, reconstruction algorithm, method for the accurate segmentation of regions of interest (ROIs), and testing of the core technology of texture analysis. If we want to apply the results of radiomic studies to clinical practice, it is necessary to formulate standardization in all aspects. Moreover, the specific algorithm of radiomics cannot be explained and the feature extraction part is not transparent. How to make clinicians trust this tool is another problem. Although there is some hope that radiomics could be applied in clinical practice, more work needs to be done.

Changes in the text： We have made extensive modification and additions (see Page 13, Line 281-290; Page 14, Line 291-302)

Reviewer B Comment-Grammar and spelling: Page 4 -Corrected grammatical mistake in the “both which” to “both, which” (line 76)

Page 6 -Corrected grammatical mistake in the “occur along the treatment” to “occur along with the treatment” (line 130)

Page 8 -Corrected grammatical mistake in the “According to the these studies” to “According to these studies” (line 162)

Reply: Thank you for your comments. We have corrected the above three grammatical mistakes in the corresponding position of the article.

Changes in the text： We have made modification (see Page 5, Line 101; Page 7, Line 157; Page 9, Line 190).