Peer Review File
Article information: https://dx.doi.org/10.21037/jtd-24-191

Review Comments

The manuscript “Clinical-radiomics nomogram for the risk prediction of esophageal fistula in patients with esophageal squamous cell carcinoma treated with intensity-modulated radiation therapy or volumetric-modulated arc therapy” aims to develop the predictive model to predict the risk of esophageal fistula in radiotherapy-treated esophageal cancer patients. The authors identified risk factors of developing esophageal fistula after radiation treatment and reported the nomogram using clinical features only and with an integration of radiomics features.
This study was well-designed and analyzed using the standard approach for radiomics and statistical analysis method. The authors demonstrated the methodology of model development and provided user-friendly nomogram. However, the limitations include single institutional dataset, a lack of external validation, and the difficulty in model parameter assessment which might limit the generalization of the model and nomogram.
The manuscript was overall written. On the other hand, I found some of the points and results were inadequately described. I explained my concerns in more detail below.

Comment 1: Line 158: In M&M section, the use of “10-fold cross validation” was described but the results were reported using 5-fold cross validation method (Line 238, 248, 259).
Reply 1: I'm sorry, it was our mistake in writing. It has been rectified according to the correct content. (see Page 8, Line 177).
Changes in text: Page 8, Line 177.

Comment 2: Line 260: The AUC of clinical-radiomics model was reported 87.72% in Table 3.
Reply 2: I'm sorry, it was our mistake in writing. It has been rectified according to the correct content (see Page 11, Line 289).
Changes in text: Page 11, Line 289.

Comment 3: Line 326-328: This sentence suggested that the excessive tumor length was related to the development of EF, however, the results of this study showed that the length of >8.4 cm was associated with significantly lower EF risk with the multivariate OR of 0.049. Some discussion is beneficial for these conflicting results.
Reply 3: I'm sorry, it was our mistake in writing. It has been rectified according to the correct content (see Page 11, Line 357-365).
Changes in text: Page 11, Line 357-365.

Comment 4: Line 359-360: The study design and results did not lead to the conclusion that this nomogram might select the patients to omit radiotherapy.
Reply 4: Thank you very much for your suggestion. We are already discussing adding explanatory content to the relevant sections (see Page 11, Line 303-307).
Changes in text: Page 11, Line 303-307.

Comment 5: There are some issues that need to be mentioned and discussed in more detail.
Reply 5: We are already discussing adding explanatory content to the relevant sections (see Page 11, Line 346-369).
Changes in text: Page 11, Line 346-369.

Comment 6: - The barium esophagogram might not be routinely performed in many centers. This is a critical limitation of the generalization of your model/nomogram.
Reply 6: In China, esophageal barium swallow is a very common examination method, and almost all radiotherapy centers have barium meal examination equipment. Moreover, every esophageal cancer patient undergoes a barium meal examination during diagnosis, which is the most common and convenient diagnostic method.

Comment 7: - Nearly 30% of patients received high radiation dose (>=60Gy) which is not standard dose for esophageal cancer treatment, especially in non-Asian countries. More discussion might elucidate the use of high-dose RT in this study.
Reply 7: Some of the patients included in our analysis had cervical esophageal cancer, and current guidelines recommend a main dose of 60Gy for cervical esophageal cancer. On the other hand, the dose range for thoracic esophageal cancer is generally 50Gy. Although doses higher than 50Gy did not bring survival benefits to patients, relevant studies and our previous data had shown that higher doses could achieve higher local control rates in patients with esophageal squamous cell carcinoma. Therefore, in clinical practice, some patients may receive treatment with a dose of 56Gy.

Comment 8: - Please provide more details or explanations about the RadScore, including the pyradiomics coding or the online tool for scoring assessment.
Reply 8: Thank you for your reminder. We have already provided the coefficients of the selected imaging biomarkers in Table S2. To facilitate the calculation of radscore, we have additionally included details about Radscore calculation in the main text (see Page 10-11, Line 267-275).
Changes in text: Page 10-11, Line 267-275.

Comment 9: - The performance of clinical-radiomics model was slightly improved

compared to the clinical models with an AUC 0.849 to 0.896, which suggests that both are good models. The simplicity and interpretability are the main advantages of the clinical-only model. What are the criteria of recommendation to use the clinical-radiomics model?
Reply 9: Both the clinical model and the clinical-radiomics model indeed exhibit good performance. Although incorporating radiomics brings slight improvement in performance, it still indicates that radiomics features complement clinical information. In the future, we will utilize other artificial intelligence technologies to further explore imaging information, which is expected to bring higher performance improvement. We have also supplemented this point in the limitations (see Page 14, Line 391-397).
Changes in text: Page 14, Line 391-397.

Comment 10: Table 1: Please consider using Fischer's Exact test for factors with a cell that has an expected value 5 or less, e.g. diabetes, concurrent chemoradiotherapy.
Reply 10: We have added data and made relevant modifications in Table 1.
Changes in text: Table 1.

Comment 11: Table 2: The Odds ratio for total dose and ulcerative type was not available.
Reply 11: We have added data and made relevant modifications in Table 2.
Changes in text: Table 1.

Comment 12: Table 3: Please consider adding the Brier score in the table for easy interpretation of the results.
Minor suggestions:
- Please check the consistency of the results between text and tables.
Reply 12: In our study, calibration curve and Brier score were adopted to evaluate clinical nomogram and clinical-radiomics nomogram, which had been shown in Figure 3 and 4.