Peer Review File

Article Information: http://dx.doi.org/10.21037/atm-20-6201

Reviewer Comments

The authors developed and validated a nomogram for predicting prognosis of patients with uterine cervical adenocarcinoma using SEER dataset. They identified several independent prognostic factors and validated the nomogram internally and temporally. The manuscript is well written and the conclusion is clear. I have several comments that I would like to see the authors address.

Major comments
1. In prediction selection, you firstly identified 7 strong predictors among 15 variables. Could you explain how you selected the 7 variables a little more detail?
2. Table 1: You mentioned the characteristics in the whole cohort. I would recommend to add a brief comparison between development cohort and validation cohort. This comparison might be important for temporal validation of the model.

Minor comments
3. Line 82: Please cite the previous studies that focused on the prediction of prognosis or treatment outcomes of patients with UAC.
4. Line 91: "TRIPORD" should be "TRIPOD".
5. Line 104: The targeted outcomes may be first primary UAC. Could you say "as first primary tumor"?
6. Lines 171-173 and Figure 1: It's not clear to me. The cause of death was unknown for 700 patients? And the UAC for all of these 700 patients were the 2nd or subsequent tumor?
7. Lines 183-193: This part (from "To refine the model…" to "…were selected into the final mode.") seemed to be repeated in lines 255-263 in Predictor selection section of Results.
8. Line 202: "Receiver operating characteristics curves (ROC)" should be "Receiver operating characteristics (ROC) curves".
9. Line 239: "47.6" may be "45.0".
10. Line 411: "grade III" may be "grade III and IV".
11. Line 422: "Table 3" may be "Table S2"
12. Line 432: Please spell out RFS.
13. Table 1: It might be better to show the categorical distribution for age, tumor size, number of LNs examined, and number of positive LNs as in the other tables, because these variables were treated as categorical variable in the model.
14. Figure 6: It's not clear for me why you showed the survival curve (C) by AJCC staging system and (D) by SEER summary staging system.

1. In prediction selection, you firstly identified 7 strong predictors among 15 variables. Could you explain how you selected the 7 variables a little more detail?

Reply 1: LASSO regression analysis via cross-validation was applied to determine the most important predictors from the derivation dataset. In the process, fifteen variables with non-zero coefficient value were identified with corresponding lambda value and likelihood of deviance. Even one standard error criterion was employed, ten variables were remained (Figure S1). To refine the model for clinical use, the predictive power of the models with 5 to 10 variables were compared. The results showed the area under the ROC (AUC) varied from 0.848 to 0.853 when

7 to 10 variables were incorporated into the model, whereas AUC of the model was 0.837 with 6 variables. Therefore, seven variables with the corresponding lambda value were ascertained into further analysis: grade, stage T, stage N, stage M, tumor size, number of positive LNs and surgery of primary site. We have supplemented the contents as advised (see page 10-11, line 189-204).

2. Table 1: You mentioned the characteristics in the whole cohort. I would recommend to add a brief comparison between development cohort and validation cohort. This comparison might be important for temporal validation of the model.

Reply 2: Thank you for your kind reminder. We have supplemented the comparison between development cohort and validation cohort with the raw data in table 1 and with the multi-imputed and categorized data in table S1 (see Table 1 and Table S1). The difference of some variables between the two cohorts showed statistically significant, which indicated that the nomogram was validated with another different population and, in a sense, its extensionality and stability was confirmed.

3. Line 82: Please cite the previous studies that focused on the prediction of prognosis or treatment outcomes of patients with UAC.

Reply 3: We have added the reference in the manuscript as advised (see page 5, line 84).

4. Line 91: "TRIPORD" should be "TRIPOD".

Reply 4: Thank you for your extremely careful review. We have corrected the spelling mistake (see page 6, line 93).

5. Line 104: The targeted outcomes may be first primary UAC. Could you say "as first primary tumor"?

Reply 5: Thank you for your extremely careful review. We have revised it as advised (see page 6, line 108).

6. Lines 171-173 and Figure 1: It's not clear to me. The cause of death was unknown for 700 patients? And the UAC for all of these 700 patients were the 2nd or subsequent tumor?

Reply 6: Thank you for your professional suggestions. At the start of study, all the cases with UAC as primary tumor were enrolled. Among them, 700 patients were found with unknown causes of death, for whom UAC was the second or more of the primaries. As shown in Figure 1, these cases were not included in the subsequent study and we have revised this part (see page 10, line 175-181).

7. Lines 183-193: This part (from "To refine the model…" to "…were selected into the final mode.") seemed to be repeated in lines 255-263 in Predictor selection section of Results.

Reply 7: Thank you for your extremely careful review. We have revised it as advised (see page 14-15, line 275-282).

8. Line 202: "Receiver operating characteristics curves (ROC)" should be "Receiver operating characteristics (ROC) curves".

Reply 8: Thank you for your extremely careful review. We have revised it as advised (see page 12, line 220).

9. Line 239: "47.6" may be "45.0".

Reply 9: Thank you for your extremely careful review. We have revised it as advised (see page

14, line 258).

10. Line 411: “grade III” may be “grade III and IV”.

Reply 10: Thank you for your extremely careful review. We have revised it as advised (see page 22, line 437).

11. Line 422: “Table 3” may be “Table S2”

Reply 11: Thank you for your extremely careful review. We have revised it as advised (see page 23, line 448).

12. Line 432: Please spell out RFS.

Reply 12: Thank you for your extremely careful review. We have revised it as advised (see page 23, line 458).

13. Table 1: It might be better to show the categorical distribution for age, tumor size, number of LNs examined, and number of positive LNs as in the other tables, because these variables were treated as categorical variable in the model.

Reply 13: Thanks for your valuable comments. The continuous variables were non-normally distributed and marked with # in Table 1 (see Table 1).

14. Figure 6: It’s not clear for me why you showed the survival curve (C) by AJCC staging system and (D) by SEER summary staging system.

Reply 14: AJCC staging system and SEER summary staging system are internationally recognized as classical staging systems in UAC patients. We attempted to display the survival conditions for the UAC patients at different stage. As mentioned in your professional suggestions, it seems not very associated with our nomogram. We have deleted this part in Figure 6 (see Figure 6). Thanks for your valuable comments.